EXPERIMENTAL PROCEDURES FOR POLITICAL SCIENCE RESEARCH

A Dissertation

Presented to

the Faculty of the School of Public Administration

University of Southern California

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Public Administration

by

William J. Gore

June 1952

This thesis, written by

WILLIAM J. GORE

*under the guidance of h*is *Faculty Committee, and approved by all its members, has been presented to and accepted by the* Faculty of the School of Public Administration *in partial fulfillment of the requirements for the degree of*

Doctor of Public Administration

Date May 26, 1952

Faculty Committee

John M. Pfiffner
Chairman

L. C. Stockford

W. B. [illegible]

John D. Gerletti

Frank P. Sherwood

TABLE OF CONTENTS

CHAPTER PAGE

LIST OF FIGURES

CHAPTER I

INTRODUCTION

The purpose of this paper is to discuss certain of the implications of the scientific method for political science, and especially public administration. In this time when the need for the solution of practical political problems is so pressing, improvements in the methodology of political science may provide assistance in locating solutions for these problems.

I. THE PROBLEM

Statement of the problem. The problem here is to analyze the rationale and techniques of the experimentalists in the three basic social sciences as they relate to research in political science. The problem can be divided into two areas. First, there is the epistemological question. What are the criteria of meaning and explanation which will form the foundation for the sort of scientific method acceptable to political scientists? The second question concerns the methods appropriate to the problems of political science. Is quantification possible in this field? What are the limitations on the use of measurement in political science?

The so-called "cultural lag" might be explained as a failure of various social science disciplines to keep pace with the complexity of the problems to which society seeks solutions. Perhaps most of our pressing national and international problems are the result of inherent weaknesses in world cultures becoming important because of cultural maturation. We are attempting to control institutional processes which we once accepted as necessary evils. It is increasingly clear that control of many cultural processes rests upon reliable knowledge of what actually happens. So far, political scientists have been unable to describe either politics or administration in terms which provide the degree of control which we require. There is little doubt but that our research programs have been largely inept.

One of the serious weaknesses of research in political science is that it has been fragmentary. The political process is group behavior. It is set in a socio-economico-legal context. Most research in this area has been carried out within a single discipline with the result that no researcher has yet had a hold upon the whole problem of politics.

There are many penetrating studies of constitutional law by constitutional lawyers. There are some excellent studies of voting behavior by sociologists and public opinion experts. There are some fine studies of the administrative process by psychologists and anthropologists. Some of these, such as the work by Gosnell and Roethlisberger, have been focal studies which have reoriented the theory about political behavior. We are still waiting for the first study in any area of political science which takes into account both the process and the context.

The writer feels that the traditional function of the political scientist in studying the ethics of political action is of fundamental importance. The expert analysis of the implications of proposed courses of action makes for increased popular control of the political institution since it makes known the implications of these proposals. It would be difficult to say that the high level of skill in ethical analyses of our political scientists is not one of the foundations of our democratic system. The concern of the political scientist with the description of the political process is also of great importance.

The writer has two reservations about the effectiveness of political science today. First, political science research is not extensive enough in its scope; and second, it requires an improved method of description. In these reservations there is no criticism of the effectiveness or the skill of the political scientist as a student and practitioner of political ethics. This is pointed out because the first reaction of most political scientists is that this is what one means when he criticizes the methods of political science.

The criticism of the limited scope of political science research can be remedied through the use of cross-discipline studies which bring the frames of reference of law, economics, sociology, anthropology, psychology and the fields of political science to bear upon an experimental question. The great stumbling block in the way of such cooperation is that it requires that the "experts" from each of these fields admit that there are limitations to their concepts.

It is the second problem, that of the limitations of the research techniques of political science, that is the problem for this discussion. The traditional research method of political science is generally identified as the historical method. It is sometimes called the developmental method, the method of showing antecedents, or the analytical method. Regardless of what it is called, it rests upon the implicit assumption that man and his society are such complex phenomena that there is no method which will reach deep enough to provide us with the detailed interpretations of the behavior of our culture. Those who support this approach often have grave doubts about the value of experimental procedures.

The social scientist has no quarrel with the ends of political science. He does, however, reject this point of view. He rejects it very carefully, being sure to state the precise conditions under which he disagrees. There is no question but that the decision-making process, no matter where it occurs in the political institution, is best handled using the mechanism of the "mind." The person with intelligence and sensitivity and vast experience has the ability to make decisions on an intuitive basis within practical time limits which no system we now have can rival. The social scientist also holds that the brain is incapable of providing a reliable representation of itself or society without mechanical assistance. He may feel that the human "mind" is just like any other mechanism, appropriate to some functions and not appropriate to others.

This is not to say that we have a device which will produce knowledge which is a true representation of nature. Social science does provide a system which produces symbolic representations of nature and society which have a <u>known relation</u> to what exists. Most important of all, it is characteristic of this system that it is self-corrective in such a way that by continual repeating of some of the steps in the scientific method the adequacy of the symbolic representation is increased.

<u>Importance of the study</u>. Contrary to what many

laymen think, there is nothing in the scientific method that will invariably produce absolute truth; absolute truth is not something which can be pursued by means of the scientific method. What one does get are generalizations which have a stated probability of corresponding to what actually exists. The first generalizations derived about a particular area of nature typically have low contingency factors. The probability that they represent something which actually exists is low. The generalization which one derives after repeating a study many times will still have a contingency which makes it far from being a certainty. One may have some confidence that there are ninety-five or even ninety-nine chances out of one hundred that the generalization does correspond to something which actually exists; however, there is no way of determining so far the probability of a generalization derived through the analytical method being reliable. The result is that there is no way of choosing between alternative interpretations of the same set of data and therefore generalizations derived through this method tend to be indeterminate. Early generalizations using the scientific method are indeterminate also. The significant difference is that using this method we have a means of revising the assumptions which are used to derive generalizations. The result is that the assumptions may come to conform more and more to the broad aspects of nature.

It is here that the early portions of this paper come to bear. The assumptions which one places beneath an investigation are of the utmost importance. They determine, in large part, what sort of conclusions will be developed. The manipulation of the assumptions of a study is one of the critical steps. In the scientific method it is very carefully handled. This whole problem is here called the epistemological problem. It is treated extensively in Chapter III and a suggested pattern of steps for handling assumptions is suggested in the first section of Chapter VII.

The second area in which we need to expand research in political science is in the type of techniques used. There are some principles for the analysis of the legal process; and there are some other principles for the analysis of the administrative process; and still others for the analysis of party processes. In almost every case these methods are based upon comparison of descriptions of several processes or in the comparison of the description of what exists with what has been held should exist. The emphasis of these techniques is in evaluating "what is" in the light of "what should be." This is a function which must be carried out, but it is seldom a substitute for the control provided by a representation of what actually exists. The solution of problems, practical or philosophical, ultimately depends upon the identification of the elements of the problem. The "problem" in this context might be some pattern of behavior in society which is not desirable. The solution to the problem is to modify a pattern of behavior.

The fourth and fifth chapters of this discussion deal with methods which may be borrowed from the three basic social sciences to improve the political scientist's ability to represent reality. The ideal form for these chapters would be to lay down a set of methods for the description of any problem in political science. This challenge was intriguing, but considered impractical at this time and the actual content of these chapters concerns only the techniques which political science might borrow. There is no attempt to provide an inventory of techniques here for any conceivable political problem.

The dissertation is organized to discuss the problems which have been outlined here. Chapter II gives some account of the growth of research in public administration to put the problem of research in some perspective. Chapters III and IV outline the epistemological problem and Chapters V and VI deal with research techniques. Chapter VII represents an attempt to outline a sequence of steps which a layman in political science might follow in the development of a research project.

II. DEFINITION OF TERMS USED

A number of terms are critical in the rationale of this discussion. The following terms are central in the dessertation.

Initiation of inquiry. The identification of the problem and the statement of the assumptions and hypotheses of the study.

Exploratory phase. The sequence of steps following and dependent upon the results of the initiation of inquiry. This phase involves the refinement of the hypothesis and the location of variables.

Measurement phase. The sequence of steps, most of them involving statistical techniques, where the data of the study are actually collected and conclusions drawn.

Theory system. A complex of concepts, beliefs, values and the like which are interrelated into a logical system.

Contingency factor. A coefficient, numerical or otherwise, which represents the probability that a given conclusion is dependable.

Epistemic correlation. This term represents the degree of correspondence between the symbolic representation of some behavior in nature and the actual behavior which is being represented.

The following terms are defined only so that the prejudices of the writer will be clear. These definitions are arbitrary and are made only to facilitate the discussion.

Research. Research is the derivation of meaningful generalizations which have a substantial probability of corresponding to what exists. In applied research the generalizations desired are much narrower than those of basic research but the pressure for reliable conclusions is just as powerful as in basic research.

Political science. Political science is the study of the political institution and its processes and relations to other institutions in any complex culture.

Public administration. The study of the processes and devices for the implementation of policy. This includes the study of the relationship of policy making to the political party, to the economic institution and other components of our society.

CHAPTER II

SURVEY OF THE HISTORY AND DEVELOPMENT OF RESEARCH IN PUBLIC ADMINISTRATION

I. INTRODUCTION

The rise of research in public administration is perhaps best characterized as a defensive measure. It has been forced upon administrators and theorists alike as a means of combating our present ignorance about the happenings in organizations.

Research in this field has been largely a means of dealing with the "culture lag" between the social and the physical sciences. There has been developed a body of principles from the experience of administrators and through the logical synthesis of academicians to cope with problems which would be handled in the physical sciences through the use of a system of tested laws. It is widely recognized that there is a great gap between our principles and the laws which are inherent in the day-to-day action of administration.

Professional administrators find that they can not always depend upon the axioms of formal administrative theory. Not only might the administrator find several axioms which may apply to a given situation, but one or more of these may indicate action which would be contradictory to the action implied by some other principle.[1]

This has led to a good deal of criticism of administrative theory. Public administration has reached the point where many professional administrators habitually explain their actions in scientific terms of "span of control," "unity of command," and "channels of communication," but where many of them attempt to base their action upon experience rather than principle. The administrator, faced with pressing practical problems, has found that contemporary administrative theory is not sufficient equipment with which to meet the demands made of him. Some hold that the very fact that our principles are satisfactory as a means of explaining administrative behavior proves their value. The weakness in this point of view is in holding that any principle sufficiently ambiguous to be applied to a large number of situations thereby gains scientific utility. It has something akin with the man who confidently predicts that all elevators which go up must come down, the point actually being whether they will fall down or descend slowly on a cable.

Opinions as to what to do about this situation if

---

[1] Herbert A. Simon, Measurement Techniques in Administrative Research (/n.p./ Civil Affairs Pamphlet, /n.d./), pp. 5-6.

placed on a book shelf would fill a space several feet long. Some hold that much as it may be needed, it is impossible to develop scientific knowledge here and we must "muddle through" as best we can. Others hold that we can develop an absolute science of administration if we merely collect enough descriptive data concerning the manner in which organizations are administered. Some of the more sagacious, holding to the middle ground, feel that we have not yet attempted to create a real science of administration and that an attempt should at least by made to give us some evidence of the probable limits of the field of knowledge.[2]

To state merely the positions of the protagonists in this field leaves out the time dimension. The controversy has persisted some twenty-five years without the precipitation of a definite decision in favor of any of the alternative points of view. Disregarding this controversy a number of organizations moved into the breach in our knowledge in an attempt to fill it as they saw fit.

<u>The rise of bureaus of research</u>. During the last thirty years bureaus of research in the universities, taxpayers' organizations, leagues of municipalities,

---

[2] Leonard D. White, <u>The New Social Science</u> (Chicago: University of Chicago Press, 1930), p. 31.

legislative reference bureaus, divisions of governmental jurisdictions and various other organizations have carried out what they have termed "research" activities in public administration and political science. Dr. White, in 1933, said:

> In few particulars has there been a more remarkable transformation in the last twenty-five years than in the field of research in problems of municipal, state and federal administration. The person involved in governmental operations today is acquainted with an elaborate and active group of research agencies, primarily and often exclusively devoted to solving problems of public administration.[3]

While there were many bureaus present during this period, they undertook only limited types of research activity. A great deal of effort was expended upon development of procedures to enforce accountability. A good deal of work was done on the complex of problems pertaining to fiscal resources. Other typical activities of these organizations included the study of intergovernmental relations and analyses of specific line operations, such as police protection, fire protection and public health.

Gaus pointed out that in 1930 there were a "tangle of influences" which accounted for the research movement which he saw developing as one of the timbers in the

---

[3] Leonard D. White, Trends in Public Administration (New York: McGraw Hill Book Company, 1933), p. 311.

substructure of this field.[4] In his discussion of the role of the universities in the development of public administration, Graham recognizes a research activity as one of the basic parts of the field.[5]

By the 1930's, research activities along with personnel and finance activities were accepted as primary dimensions of administration. This had a tendency to set the field off even more distinctly from political science for, at that time, it was difficult to collect data and analyze them statistically.

The need for research methods. To place the discussion of research methods which will appear in the later chapters of this dissertation in a perspective, it seems worthwhile to inspect the development of research in this field in some detail. It is obvious that the underlying reason for the emergence of research as a major element in this field when it was traditionally shunned in the mother discipline of political science, resulted from the appearance of explosive problems which were specific enough in nature so that the application of taxonomic techniques

---

[4] John M. Gaus, A Study of Research in Public Administration (New York: Social Science Research Council, preliminary draft, 1930), p. 79.

[5] George A. Graham, Education for Public Administration (Chicago: Public Administration Services, 1941), p. 19.

provided insight. But there were many visionaries who braved a storm of ridicule to supplement these "census-taking" techniques with a scientific approach to management. One might very naturally hold that this movement was merely an attempt to apply the scientific method to problems of administration, but this would be a considerable oversimplification. The appearance of the scientific management movement was not merely the result of the diffusion of a relatively new culture complex into one of the periphery areas of social science. It was almost solely the result of a single powerful personality, Frederick W. Taylor. The motives of the scientific management disciples, for they were disciples in a very literal sense, were not those of the scientist who seeks knowledge for its own sake. These people thought that they would be able to eliminate the "human element" from the work situation by the reduction of discretion to a series of formulae. To the extent that this unique approach to management is a part of the tradition in management thought it must be considered.

It is felt here that the activities of the research bureaus, both private and public, developed through the application of techniques for systematic analyses to a newly emerged set of problems. The scientific management approach, on the other hand, was one of those spectacular cultural phenomena, so typical of Western culture, the eruption of a new idea. On the basis of these assumptions the two movement will be treated separately in the following two sections of this chapter.

A third and concluding section will include an evaluation of the state of research in public administration against the present standards of social science methodology today.

II. ORIGINS OF RESEARCH

The origin of research in public administration must probably be traced back to the birth of the field itself. In the discussion opening this chapter it was noted that research was considered one of the integral parts of the discipline almost from its birth. The presence of research as one of the important parts of the new discipline is even more meaningful against the history of political science itself.

<u>Philosophy as a basis for all disciplines</u>. Some hold that philosophy is the mother of all the disciplines which we now know. It is noted that the ancient philosophers were mathematicians, physicists, epistemologists, sociologists and students of law. Knowledge was nothing if not a comprehensive structure of ideas which was rich because of the admixture of points of view which went into it. The

approach of philosophy had traditionally been toward problems of ultimate goals and absolute truths.

Early in the maturation process of our Western cultures a pressure for specific guides to action emerged. The hypothesis of the writer is that as we developed our cultural institutions into intricate mechanisms, we created the need for a body of integrated knowledge. When the Romans attempted to operate their state on a systematic body of laws they soon found that they required epistemological specialists. When education was taken from the home and established as a separate institution specialists soon appeared. One of the interesting dimensions of the medieval church-state controversy is the battle of ideas between the religious and the political philosophers. Fractionalization was resisted here and one of the issues was the attempt to take the subjective religious element out of political theory and place it on a separate epistemological foundation.

<u>Separation of administration and political science</u>. In terms of this splintering tendency public administration can be traced back to political science which sprang from history and law. It might be maintained that there is a qualitative difference between the first splintering and the last one. Disciplines of knowledge have traditionally

served as broad reservoirs of ideas. When the advent of new ideas brought the level of the reservoirs threateningly close to the original boundaries, we have simply divided our ideas into two or three classes and built new reservoirs, each as large as the original one. As noted above, once in a while the maturation of the culture justifies the establishment of one of these reservoirs as a new discipline. These reservoirs fulfilled the fundamental cultural function of maintaining the fund of ideas from which we made withdrawals periodically to meet an emergency or some other emergent phenomenon.

Recently, only yesterday in terms of the time unit which we have been using in this discussion, a new type of discipline has been developing to meet the need for a body of exact, dependable knowledge. A complex culture apparently requires two functions from the educational institution. First, the preservation and creation of cultural alternatives, and second, the development of a body of "facts" which can be used to deal with the now uncontrollably complex cultural processes. The individual-oriented disciplines have long done this. Law, medicine, religion and biology have supplied not only systems of knowledge, but dependable lists of "do's" and "don'ts" which the individual used to cope with various situations he was unable to handle from his own experience. For instance, such facts as what to do to keep from getting frostbite, or religious beliefs necessary to lay up an investment for another life.

Now we are developing the same functions in the group-oriented disciplines, social work, public administration, business administration, educational psychology and others.

The point of view here is that the presence of research in political science is a natural emergence phenominon typical of the larger movement in all of the "group disciplines." The keystone of our maturing culture is organized effort. We have found the key to organized effort to be integrated knowledge--not knowledge of what ought to be, but dependable knowledge of what is. Speculation or even considered opinion has been found insufficient as a basis for organized effort involving a high division of labor.

Beginnings of reform in administration. Political scientists have traditionally placed much emphasis upon the study of the ethics of politics and upon the description of American government. These were, and they remain, of fundamental importance in the development and modification of the broad cultural goals of our society. About the turn of the century a demand, largely based upon the great reform movement preceding this period, for a different type of activity emerged. John Burgess and Herbert Adams, among others, felt this need and began experimenting with new courses patterned after their experience in Germany where there was a good deal of attention given to the practical aspects of government. In 1887, Woodrow Wilson wrote what is becoming a classical article in the <u>Political Science Quarterly</u> where he pointed out the appearance of an interest in problems of administration in several universities, and he held that this was an important trend. Wilson talked of the development of a practical science of administration with research as one of its elements.[6]

His implication was that the nature of these practical problems required a new approach if they were to be solved successfully.

It was Professor Goodnow, however, who stepped forward with the concept to justify the new approach. Goodnow provided the hypothesis that government was not a single cultural process, but two distinct activities serving two rather different functions. In his view there was the political process which amounted to an institutionalization of the intricate problem of the control of government by

---

[6] Woodrow Wilson, "The Study of Administration," <u>Political Science Quarterly</u>, 2:238, 1887.

the public, and also the administrative process which was the institutionalization of the problem of providing governmental services to all with fairness and justice.[7]

Egger points out that Goodnow was the first person to make a substantial reorientation in the study of administration by pointing out the necessity of studying the practical problems of structure and the control of administration separately from the question of the policy goals of the administration organization.[8] This view soon prevailed, though not just in the form in which Goodnow had propounded it.

Public interest in administration. Following 1900 there was a continuing interest in the idea of studying the practical aspects of administration. There is no doubt but what the idea was spreading, though there is little evidence of a concrete nature which can be cited. Farlie published a work on National Administration during this time and Willoughby was studying the problem of governmental reorganization. The whole idea of government reorganization was in one sense a manifestation of the new interest

---

[7] Frank J. Goodnow, Politics and Administration (New York: Macmillan Company, 1900), p. 16-23.

[8] James W. Fesler, et. al., The University Bureaus of Public Administration (University: University of Alabama, Bureau of Public Administration, 1946), p. 12.

in the practical aspects of administration. White notes that it resulted from the long concern with reform. It began just after the turn of the century and was a full scale movement by about 1912, lasting until World War I.[9] An important implication of this movement is the indication of public interest in the independent consideration of problems of effective administration. This concern of the public became concrete in the institution of the first citizen-supported governmental research group and later with the Civil Service Research unit in the federal government.[10]

Thus, by the beginning of World War I the idea of the systematic study of administration was established. This was a significant transition from the conventional approach of the political scientists, even though it was largely the work of "converted" political scientists. It is held here that in this new approach we see the first manifestation of the empirical approach to administration inherent in Goodnow's dichotomy, and the forerunner of the more scientific approach to be set forth in this discussion.

One of the important problems which the pioneers in

---

9 White, Trends in Public Administration, op. cit., p. 323.

10 Ibid., p. 3.

this area faced was the charge that they were overlooking or even escaping from the ethical problems of politics with their intensive efforts to describe government. The controversy raged for several years without producing any agreement and was finally cut off by World War I.

The War provided a good deal of evidence of the necessity for the analysis of governmental operation. It was admitted that no single man, or even a whole staff of experts, could possible explain the processes of government, and hence indicate the means of control. Government was seen as a total process which required a complex, integrated "wholistic" frame of reference for its successful analysis, or in short, a scientific construct. Political science did not provide this. Ethics are purely logical in the same sense that mathematics are purely logical. An ethical system, like a mathematical system, is created upon a set of assumptions and the assumptions underlying political theory have been such that they produced a system of concepts which referred to a very specific part of life, and this only indirectly. The concepts of political science, coming largely from political theory, were therefore only a part of the concepts that are required for the successful analysis of a total institutional process.

<u>The establishment of university chairs of public</u>

administration. During the decade following World War I, Chicago, Columbia, Wisconsin, California, the University of Southern California and several other major schools established chairs of public administration. In 1928 Chicago established a chair of police administration. This is interesting in that it presages the coming fractionalization of the new field. Also, in 1928, the Social Science Research Council established a committee on Public Administration, and in 1934, the American Political Science Association held a round-table on the problems of public administration.

At the same time universities were beginning to establish libraries to handle the materials in the field and case studies of administrative problems began turning up in doctoral theses. The research bureau movement was well under way resulting in a new cooperation between the schools and governmental jurisdictions, especially at the state and local level. Finally, both Willoughby and L. D. White published text books designed for the teaching of public administration in the schools. During the decade of the twenties public administration emerged as an independent field of political science with research as one of its basic sub-fields.

Summary. The emergence of public administration as a field of political science is interesting here because it marks the acceptance by the profession of the need for a dependable system of descriptive terms of reference. Professor Egger discusses the evolution of public administration research as falling into four successive phases.[11] He cites Goodnow's approach as the first phase. This was a new approach in the sense that he first used a new frame of reference in the analysis of administration. For the first time a political scientist studied things such as political corruption, favoritism and gross inefficiency in terms of the behavior which they involved, instead of the ethics which they violated.

Egger's second phase is called the "institutionalization" phase and is concerned with the location of the component parts of the great administrative machine and their relationships. He holds that Willoughby's studies of the problems of top management, coordination and centralization are examples of this.[12] The implications of this development for research are that the sub-fields of public administration were functional, thereby lending themselves to an empirical approach. The importance of the orientation given public administration research in

---

11 Fesler, *op. cit.*, p. 10.

12 *Ibid.*, p. 12.

this period is pointed out clearly by White. He notes that the period 1900-1930 was the formative period in public administration and that it was during this time that the orientation toward an empirical approach was set into this field.[13] White notes that in 1900 there were no organizations in the nation which were devoted to the problem of practical governmental research. In 1930 there were university bureaus of research, governmental research bureaus, and citizen-supported research bureaus, and there was general acceptance of the idea of the purely technical nature of certain governmental problems. This was a denial of the point of view that the all-important problems of government were ethical.

The importance of this period is that it represented the fruition of the split between public administration and political science which allowed for the development of a different attitude toward proper methods and subject matter for study. The methodology of political science was that of history and the law. Neither the historical nor the legal approach provided the control needed for analysis of administrative problems.

Now we are in a period where the approach of public administration research is being questioned by a few and

---

13 White, Trends in Public Administration, op. cit., p. 335 ff.

we are perhaps to face a period when this approach will be questioned by many. In order to clarify the present approach of public administration research, the last two sections of this chapter will outline the research bureau movement and the impact made upon research by the scientific management movement.

III. THE RESEARCH BUREAUS

It is undeniable that the development of research bureaus has been important to the maturation of public administration research. To speak of a "research bureau movement" is somewhat confusing, for actually there were two important types of organizations involved here. On the one hand, there have been the independent research organizations supported by citizen groups or governmental jurisdictions. And on the other hand, there has developed a rather unique type of organization in many of the large universities. Although functionally similar, some of the points we wish to make in the following discussion are not applicable to both types of research organizations. To facilitate discussion therefore, "research bureau," when used below, will refer to those bureaus of research outside the universities.

Independently-sponsored research bureaus. The bureau movement actually began in New York City where, through several large grants, an agency was established to study ways of improving the then cluttered structure of city government. During the time of a sympathetic regime the bureau made analyses and suggestions for the reorganization of city procedures based upon their studies. A change in administration through an election hampered the operation when the city departments refused to allow access to official records, but the persistent researchers turned to such studies as street paving activities and others which anyone could observe. Such tactics ultimately forced city administrators to accept the assistance of the bureau, the result soon being so impressive that bureau personnel were loaned out to other cities in the Northeast.

By 1928, the bureau had made surveys in fifty-three cities, twenty-seven counties and twenty-six states. In almost every case their activities encouraged an attempt to establish a local bureau of research. These attempts led directly to the establishment of five more bureaus in eastern cities.[14] It was after World War I when the activity became so general that it was called a movement. By 1925 there were forty-nine municipal and twelve state

---

14 Committee on Public Administration, The Governmental Research Bureau (New York: Social Science Research Council, 1938), p. 9.

citizen's fact-finding agencies.[15] By 1936 there were one hundred and thirty municipal and state bureaus, some of them having budgets as high as $60,000 a year.

The central activity in all of these agencies was the collection of descriptive facts which would indicate specific administrative problems about the organization under study. This is an activity which could hardly be termed research today, but at this time when the need for reliable information on all phases of administrative procedures was not recognized, it was of great importance. This function was fundamental not only because it gave a trustworthy description of the organization, but because it was the basis of many more refined activities such as reform and reorganization. This was a period when governmental jurisdictions were passing through the "awkward age." Government in this country had traditionally be limited in scope and further limited in funds. It was a low prestige profession. Little was expected of government by the public for they desired a minimum of governmental interference. Emphasis was upon the negative function of protecting individual rights and liberties and holding corruption in check. There was little of the positive service demand which became so strong during the decade of the

[15] Ibid., p. 15.

thirties. Equipped with these traditions, government, on all levels, was suddenly required to expand--explode might be a better term--to meet the needs of a rapidly increasing urban population. The procedures and structure of government, the attitudes and folkways of governmental personnel left our public organizations ill-prepared to meet the need. It tended to prepare the way for increased corruption by reducing or breaking the channels of communication and control. This was the scene in which the citizen-supported research unit entered. It was not a situation where there was great need for precision or sensitivity to subtle problems. The solutions which were required did not have to be new or unique insights into complex phenomena. Government had to regain its ability to cope with its simplest problems.

One of the criticisms of the research bureaus is that as corruption and disorganization were eliminated, the bureaus did not set themselves to these deeper problems. Rather they restricted themselves to the provision of services of the descriptive sort which a jurisdiction could not provide for itself. This left them in a position of providing important, but not essential services to government and reduced the value of their contribution and their influence over the agency. A typical problem might be one where a school board asked a bureau to help it install a financial system. The school system might not have financial officers in its structure. There would be no standard procedures and no qualified personnel. The bureau staff, entering the organization, would make a survey to determine the general nature of the financial requirements of the system, prepare and report its findings to the board and submit recommendations. While the activities of the bureaus cannot be represented by one example, this illustration probably is indicative of the level at which the bureaus operated.

These independent research institutions were created to serve a particular need and they fulfilled it admirably. The fact-finding survey technique which was their primary methodological tool was adequate for the analysis of most problems which they undertook. If one measures them against the purposes which they set up for themselves he cannot help but conclude that they achieved much which they set out to do. In general, it may be true that their methods of data collection were haphazard and there must have been a heavy subjective element in their interpretations of these data, but at the same time, they did produce changes in government. The research bureaus had an important impact upon the orientation of public administration research.

The traditional approach to the study of government, fostered by the attitudes passed down to us by the founding fathers, was ethical. Our yardstick for the evaluation of governmental operation was a set of beliefs concerning the protection of the individual from society. Perhaps the succession of weak presidents leading the United States around the turn of the century was not a happenstance. Perhaps they satisfied the people at that time. Even the great reform movement on the national level was not aimed at effective administration but at the elimination of corruption. The motivation of those interested in government was to eliminate the evil. But World War I showed us that a total effort was impossible without effective government regardless of the philosophical implications of complete government domination of the people.

The research bureau movement manifested certain significant changes in this philosophy. First, it admitted the necessity of government as a positive force in our culture. Second, it manifested the point of view that government could be effective and efficient and at the same time responsive to the will of the people. Third, it established the principle, for the first time in our culture, that the basis of policy-making and decision-making should be the collection and weighing of facts as well as ethical considerations. It is possible to organize and administer small primary groups such as sports teams and work teams and the like on a common-sense folkway basis. It is difficult to organize and maintain a large organization without a method of collection and analysis of facts as the basis for decision-making. Realization of this involved a radical change in the behavior of most administrators. It is a change which many were unable to make because of its threat to their emotional security. The research bureaus had some part in the stimulation of this very subtle but fundamental reorientation of the practice and theory of administration. It is often said that the thing which sets science off from any other institution in our culture is the attitude of mind which it requires. The redirection of the attitudes and the outlook of administrators in our most important jurisdictions was fostered by the local research bureau, in many cases. In a sense it was a recognition of Goodnow's dichotomy of politics and administration.

University research bureaus. It would be difficult to hold that the university bureaus of research were a distinct movement, but they have a uniqueness of function which sets them off from the others.

Cooper classifies the activities of these bureaus under six headings: First, they provide informational services, often to legislators and others who do not have access to or know-how to use the collection of references commonly maintained by the bureaus; second, they act as secretariats for various professional organizations of public officials, this on the assumption that it furthers the professionalization in the government service; third, they offer consulting services to administrators and others; fourth, they carry on specific project type surveys to solve a particular problem; fifth, they conduct training programs where technical personnel of the bureau train persons in their fields of government, and sixth, bureaus undertake their own research projects.[16] It is this last class of activities which concerns us here.

Cooper feels that there are three approaches to the research activity. The research of some bureaus covers problems of great diversity but little depth. Other bureaus tend to limit the area of their research but penetrate deeply in their chosen areas. A third group of bureaus includes those which attempt to undertake research which has some substantial value, even though they are not able to channel much effort into this type of project. Cooper sympathetically points out that the problem of most of these agencies is that they are expected to provide expert solutions for a great variety of common

---

16 Fesler, op. cit., pp. 57-65.

administrative problems. It is difficult to convince those controlling budgets that there is much value in long-range research. It appears to be Cooper's conclusion that the research bureau must compromise to be able to undertake any fundamental research, for it is inherent in the situation that pressure for close relationships with the state administrators will inevitably develop.

Gaus lists MacMahon's study of federal-state relations, Merriam's study of the Chicago region, Vollmer's study of police operations, Ridley's studies of city manager government, and White's history and studies of public administration as evidence that we have entered an era in which we are looking at government through new eyes.[17] The important thing is the wide recognition which these and later studies have received from administrators and the public. Research bureaus sponsored by private interests and by the universities have made a far-reaching substitution of new cultural values for traditional ones. In the end, the fact that these bureaus have resulted from a type of grass roots movement susceptive to subjective interpretation of research results is probably one of their most important characteristics.

---

[17] Gaus, op. cit., p. 89.

IV. SCIENTIFIC MANAGEMENT

During the period when the research bureau movement was developing there was an energetic genius named Frederick Taylor who was just publishing his first works on the scientific approach to management. Professor Gaus has said:

> The great advance which the 20th century is in the way of making lies in the institutionalization, the depersonalization and the testing of hypotheses to the establishment of principles.[18]

Those in scientific management saw themselves as carrying management toward this goal. The goal of scientific management was, "the substitution of exact scientific investigation and knowledge for the old individual judgment or opinion in all matters relating to the work done in the establishment."[19] In the beginning there was no compromise, no admission that there might be problems which were beyond even the methods of science. The goal was bluntly to substitute exact scientific knowledge for more opinion--this before 1900. There were many who considered this a bit stiff.

Division of labor. The scientific management

---

[18] *Ibid.*, p. 22.

[19] L. Urwick and E. F. Brech, *The Making of Scientific Management, Vol. I, Thirteen Pioneers* (London: Management Publications Trust, 1949), p. 11.

movement was first used as a means of making a finer division of labor in mass production organizations. With the advance of the industrial revolution the work of one man in the seventeenth century was made the work of a dozen men in the twentieth. It was found that cost could be reduced if each job was made simpler and rigid controls established to maintain standardization. Early managers were often engineers and they quite naturally turned their attention to reducing the number of operations completed on one machine. For a time it seemed that this process of subdivision could be carried out indefinitely, but soon it was discovered that there was a point where the division and subdivision of labor destroyed the natural coordination of the individuals in the organization and demanded the imposition of a hierarchy.

Taylor felt his formulae for metal cutting would prove that it was possible to eliminate the human element from production by controlling it. What he actually did was to completely control the mechanical aspects of the work situation by disregarding the social aspects present. By the complete control of the mechanical operations involved in production Taylor actually threw the social aspects of work into bold relief. Before his classic metal cutting studies it was possible to hold that the pertinent dimensions of the work situation were all mechanical.

Theorists could disregard the social dimensions because so little was known of the nature of mechanical operations that it was possible to maintain that solutions for all of the problems in production would be provided through increased knowledge of these activities. In short, it was possible to squeeze all production problems into a mechanical frame of reference by the insistence that complete knowledge about the mechanics of production would bring forth a body of knowledge sufficient to cope with them. In his metal cutting experiment Taylor did develop a complete system of mechanical knowledge and in so doing he unwittingly cut the ground forever out from under those of his colleagues who denied the social dimension of the work situation.

He proved that complete control of the mechanical aspects of metal cutting did not provide complete control of a metal cutting organization.

Personnel management. The indirect result of the unequivocal conclusions of this experiment was the creation of personnel management.[20] Significantly, it was not a "watchdog" type of personnel management like that which was growing up in public management, but an attempt to define

---

[20] L. Urwick and E. F. Brech, The Making of Scientific Management, Vol. II, The Hawthorne Investigation (London: Management Publications Trust, 1949), pp. 190-191.

and develop solutions for those social problems of the work situation the Taylorites were able to perceive.

The initial contribution of the Taylor Society was so fundamental that it is often overlooked. Most managers seem to have felt that there was nothing which could be done about the "human element" and they were resigned to the wastage and the confusion which one found in the mass production organization at the time. Implicit in this resignation was the assumption that there was a good deal which was unknowable about organizations. This was certainly the traditional point of view. It was common practice to set up a production-goal, then continue hiring men until in some manner and means the goal was reached. Little attempt was made to build a productive unit on the basis of what one man could produce. It was felt this was almost impossible.

There were a set of conditions present about the turn of the century which seem to account for the realization that a new method of organizing was required. The change was so drastic that it might almost be called an emergent phenomenon, the social equivalent of biological mutation. The change from individual production of the whole product or a complete part of it was one of these conditions. The tightening of competition in the economy and thus the development of pressure for decreased costs to survive was another one of these conditions. The beginning of the break-up of the folk feeling that there was some mysterious reason making it an impossibility to understand people was the third condition. Into this situation Taylor launched himself. He attempted to destroy the grounds for believing that rational study of organizations was impossible. He did not destroy the myth itself, for it is still the most frequent answer a researcher gets from the layman when he is asked about utility of a radical research design, but Taylor laid the ground bare for those who wished to build upon it. He tried to do this by showing the quantitative difference between, say, an army where each individual was expected to defeat one individual in an opposing army through his own prowess and a mass-production operation where there was no individual achievement. Individual achievement had to be eliminated. Folk beliefs about organizations were such that we encouraged competition among members for individual effort. Successful mass-production organizations must suppress this type of reward, along the production line. Taylor did not demonstrate this conclusively, but the point of view was implicit in his writing. Henry Ford is generally considered to have demonstrated the validity of this position.

It seems the first contribution which Taylor made was to provide some ideas for those who were able to feel a difference between traditional and mass-production organizations. Second, he made it impossible to hold that successful organization was a simple mechanical problem. It was essential that the problem of organizing be seen as comparable in size and importance with that of the family, the economy, education and the like. He only indicated these problems, but did not frame them clearly. This came with the Hawthorne experiments.

It seems as if one can almost trace the evolution of the discovery of the social dimension of management from the metal cutting study to the Hawthorne experiment. Where Taylor started with the idea of laying down the mechanical dimensions of management, the Western Electric study took the problem of discovering optimum working conditions. Where Taylor pointed out the grounds for a new frame of reference for managers, the Western Electric researchers pointed out some of the specific things which had to be taken into account such as motivation, communication, attitudes and the like.

The efforts of the scientific management adherents also indicated the necessity for conceptual entities. This marks the second contribution which they made to the development of a sound research point of view. It was not until the rationalization concept was developed that we found any systematic attempt to develop a body of theory.

But the rationalization movement did indicate that such a move was inherent in their earlier work.

<u>Rationalization</u> <u>approach</u>. The scientific management movement is sometimes seen as a part of a larger orientation called the rationalization approach. This is important because it represents an attempt to make a complete science out of management. The doctrines of Taylor were expanded to cover formal organization and research was given an important part. Those in the rationalization movement believed in management research.

In the rationalistic movement, research becomes one of the primary functions. This embodies the fully scientific point of view in that the theories which they contemplated would all have been empirically derived. It is the aim of all science to construct a verified body of theory which is comprehensive and flexible.

In summary we may note that development of research in public administration can be traced from the inception of the field through the rationalistic movement. In the rationlistic approach we see the culmination of the belief that administration is a total problem in itself which must be studied systematically if we are to arrive at solutions for our management problems. The administrator today is familiar with a variety of research organizations.

Many administrators eagerly await each new research development to see if it has applicability to their situation. Other administrators have research units in their own organizations to provide solutions for day-to-day problems.

On the other side of the balance we must note that this development has been more of a "mental revolution" than one which has culminated in the actual development of a scientific theory of administration. Much effort has been expended in tearing down the myths and blocking beliefs which prevented the institution of organized research in public administration. Before surveying the causes of this change in the approach to administration we must note that the change has yet to be manifested in the installation of actual organizational research as a part of day-to-day administration. The hypothesis of this discussion is that we have only recently approached a mental climate where the need for research as an integral part of administration is felt. Now that we are approaching the time of acceptance, those at the head of this movement find we are without some of the research techniques needed to carry out basic administrative research. An attempt will be made in the following chapters of this discussion to point out the specific research procedures available in the basic social sciences and to suggest a pattern in which they can be combined to provide reliable knowledge. It is the faith of the writer that we shall see the day when the rather ambitious goals set down by those who defined the rationalistic approach will be realized.

V. FACTORS CONDITIONING THE NEED FOR RESEARCH IN PUBLIC ADMINISTRATION

The rise of research in public administration was the result of many causes, but it seems certain that the search for knowledge for its own sake was not one of these. Research in this field has seldom, if ever, been in front of the demand for new knowledge. In the physical sciences it is common to find principles and laws which have never been fully applied to practical situations. In public administration we have only recently begun to collect descriptive facts about the administrative process. In general, the development of research in this field has been in response to the need for reliable knowledge. This need has been defined only recently.

A maturing culture. The development of science has always been characteristic of a maturing culture. One reason for this is the non-productive nature of scientific effort in comparison with the directly productive effort of the other elements of our economy. In our culture we have had the economic surplus to support scientific research in the social sciences since perhaps the middle of the nineteenth century. The absence of wide-spread scientific research is best accounted for in the beliefs which we have held about this type of research. One of the strongest beliefs blocking social science has always been the idea that people are inherently too complex to study systematically. No matter how rigorous the method, no matter how comprehensive the research organization may be, many have denied the possibility of systematic social research because they believed that it was impossible to describe the mysteries of the human mind. There has been a sort of floating limit to this belief so that as the body of knowledge concerning people increased, the limit on the complexity of the mind was moved up, leaving the mind only superficially described by our knowledge.

Political corruption in government. Our simple culture, our holy belief in the evil of big government and our complete commitment to the laissez faire economic pattern constituted another complex of beliefs which operated against the development of research. Boas notes that even at this time while we were developing disciplines which were providing some important insights into behavior, that these were unknown to most and often denied by those realizing their implications for administration.[25]

Looking back as far as the Civil War we might characterize the period 1860-1910 as one in which there was something approaching a taboo against government. It seems almost ridiculous to the student of administration now, but during this period the public did not have a primary interest in the processes of government nor in the administration of government except as it impinged directly upon them. The stereotype, widely held by the citizenry, was that the government was in the hands of evil politicians and that the best thing the individual citizen could do was to block any attempt to expand the government. Be it noted that the "evil politicians" had everything to gain by the perpetuation of this belief, for it left them with virtually a free hand.

Lack of interest in the universities. In the universities, one of the institutions in the culture from which we might have expected a real attack upon the problem, faculties forestalled meeting the challenge by maintaining that their proper concern was with the broad goals of man. They said that the analysis of practical problems was not appropriate to the university and

---

21 White, The New Social Science, op. cit., p. 84.

concerned themselves with the questions of liberty, equality, justice and the like. (It should be noted that only in rare cases did they propose any new philosophical alternatives or logical systems. Mostly they limited themselves to "rehashing" the ground first covered by the great Greek and Roman thinkers. In the writer's eyes, study in this area is valuable if it has as its goal the development of a rationale to meet the problems of a growing culture--otherwise it is a wasteful expenditure of the resources of the culture.)

Research in this field cannot be traced back to the politicians, nor to the political scientists, nor to some fundamental creed long present in our culture which produced research as an emergent phenomenon. It will be maintained here that research in public administration is a concomitant of the changing culture, the adaptation of a behavior pattern present in one part of the culture to another area of the culture during a time when government experienced "growing pains."

During the past ninety years our culture has undergone a great deal of stress which has been met, in most cases, by the development of more mature cultural processes. No longer are we considered cultural second cousins of the European and Anglo cultures from which we sprang. We are typified as "industrial," "democratic," "Christian," and characterized as a country of people having great energy, imagination and self-confidence (and an over-developed interest in sex.) It is no longer possible to hold the surly independent farmer up as the "average American." If there were such an average individual, he would be an urban dweller who is a member of a number of highly organized groups upon which he depends for the satisfaction of many of his needs. Organized effort requires knowledge--dependable knowledge which can be used as the basis for planned and coordinated activity. Even the conservative organ of professional administrators has been forced to recognize this condition.

> The compelling aspect of the . . . recent war has shown us that only organized effort can succeed and no activity can be organized effectively unless it is done scientifically.[22]

The term "scientifically" as used here, connotes first a comprehensive system of dependable knowledge and secondly, a set of reliable techniques for its application to actual situations.

Being an "organized man," our citizen is dependent upon society, in the form of a number of productive groups and institutions, for the satisfaction of most of his needs, physical and social.

---

22 Howland Sargent, "Scientists in Government," *Public Administration Review*, 2:340, 1943.

<u>Rise of the "Welfare State."</u> In this new culture, government has become one of our most active as well as one of our most important institutions. Dewey points out the reason for this development well in a simple example. In the days when each man had his own cow he had a personal interest in the quality of milk which it gave. When he became dependent upon dairy operators he often found that the producer had no personal interest in the quality of his milk, and, turning around to see who might take his interest in this case, he naturally turned to his government. So government went into the milk inspection business, the feed inspection business and the railroad inspection business.

Fesler notes that the "welfare state" government with its mass of service responsibilities and its great size has forced us into new methods of studying administration. When we had a limited government, the legalistic and the historical approach provided a certain amount of knowledge in which we could have at least some confidence. The unwieldy organization with which we are now confronted defies analysis through the use of these primitive methodologies.[23]

Many of those who are students of public

---

23 Fesler, *op. cit.*, p. 2.

administration have come to this same conclusion. White felt that the increasing complexity of the culture in which government is operating and the great increase in the size of government itself has become an imperative forcing us to the development of a precise system of knowledge.[24] Gaus holds that governmental research is necessary to arm it to deal with the tasks which confront it.[25] There are indications of a growing awareness that operation in a complex culture requires more predictable controls and more dependable means of coordination.

Economy measures. There were some subsidiary factors behind the rise of research in public administration which should be noted in passing. Gaus feels that catastrophies and crises have often brought forth spurts of research activity.[26] The expense of air war led to a great research project where scientific techniques were developed to pick men to man our aircraft. The evaporation of municipal tax sources during the depression led to substantial research efforts to locate new tax areas for cities. The need to decentralize management analysis led to the development of a whole new technique--work simplification--by the national

---

[24] White, Trends in Public Administration, op. cit., p. 335.

[25] Gaus, op. cit., p. 22

[26] Loc. cit.

government during the last war.

The cry for economy in governmental operation has been another stimulus to research. Before 1910 this was merely a heckling device. With the beginnings of the reorganization movement about this time,- economy actually was advanced as one of the alternatives for the foundation of the whole philosophy of governmental operation. Government, at all levels, was caught between the demand for services and the dictum that taxes be reduced. Often these demands were made by the same portion of the electorate. One solution was to reduce the cost of providing services so that the new functions could be added without increasing the budget. Because of the slack in many administrative operations at this time impressive savings were made through research and problem-oriented research gained wide acceptance. It has since become institutionalized in the research bureaus and the administrative research units.

Large organizations necessitate effective staff operations unknown to small units. This is another factor accounting for the development of research in public administration. First, the staff functions required in the complex organization were different in kind from the staff operations sufficient for the folk organization which we formerly had in government. Secondly, these new staff operations had to be precise to be effective and we simply did not have the administrative "know-how" to design procedures. During the Revolutionary War planning often took the form of conferences between the President, a congressional committee and a department head. The plans might or might not be recorded. In an organization with several hundred thousand people in it, planning could no longer be carried out through face-to-face contacts. Hundreds of thousands of facts have to be collected and collated. Thousands of people have to have the plans carefully explained to them. It was necessary to develop a whole new body of knowledge to determine which facts had to be collected and how the plan should be explained so that everyone involved in its execution knew what part of the plan they were to execute. Although much of this knowledge was developed through trial and error, some research was carried out to develop this knowledge systematically. This process was repeated in the case of internal controls, communication, coordination and shifting from one productive goal to another.

The forces which account for the development of research in public administration are varied. Those forces discussed above surely do not account for the emergence of administrative research, but it would be difficult to deny their importance in the development of this new cultural pattern in government.

VI. WHAT IS NEEDED IN TERMS OF ADEQUATE RESEARCH

Much of the research in public administration (and political science) has been unsatisfactory from the point of view of researchers in the experimental social sciences. William Foote Whyte notes that political scientists have expended an impressive amount of time and energy investigating a variety of problems but that almost none of their research has been directed at actual political behavior.[27] He says that in many cases students in this area have tossed off this criticism with the comment that there are so many informal relationships involved that they cannot be categorized, or even located.[28] This points up the first and most fundamental need for reorientation needed in research in this field. We will not, for we cannot, develop reliable descriptive knowledge until we go to the scene of the behavior we are investigating.

Field investigation. In every science where reliable systems of knowledge have been developed, the first step toward this system was taken when the investigators went into the field and took their research problems out of the

---

[27] William Foote Whyte, "A Challenge to Political Scientists," Political Science Review, 37:692, 1943.

[28] Ibid., p. 693.

life processes which they found there. Modern physics began with the dropping of a ball and a feather from a tower.

To the experimental social scientists the theories of the economist are not scientific because they are not based upon systematic observation, but even these are more reliable than the theories of the political scientist because they are at least based upon observation of actual life processes. Simon put his finger on this feature of administrative theory when he pointed out the hiatus between administrative principles and administrative practice.[29]

Hubbard points this out from a different perspective when he notes that all phases of personnel work need attention in terms of basic research. He says that in the rush of day-to-day administration people often forget that the only means of making substantial advances in personnel practices is through basic research.[30] This is the principal means by which we shall ever make any significant savings in time and money either. His point is that common sense knowledge from the experience of administrators will always be fragmentary and unreliable because of the bias of

---

29 Whyte, op. cit., p. 693.

30 Henry Hubbard, "The Elements of a Comprehensive Personnel Program," Public Personnel Review, 1:14, 1940.

the administrator and his limited opportunity to observe the whole of the process. What we need is a system of concepts which include broad postulates dividing personnel activities into a few comprehensive areas and then a series of behavioral or operational propositions and concepts under these which reflect actual behavior, not the goals of the personnel function.

Warner synthesizes this point of view. He says that if we are to continue the task of social transformation we shall have to provide a great body of reliable facts upon which to build and he notes that these facts must come from the application of mature scientific method to administrative processes.[31] He feels that we can never gain a "surer grasp" on the process of administration until we put a sound empirically-based system of knowledge behind our analyses.[32] He is critical of his own construct and holds that the conceptual framework of the future will be derived by the analysis of a body of factual data comprehensive in scope and intensive in its coverage of details.[33]

The dominant concern of the political scientist has

---

[31] Richard Warner, _The Principles of Public Administration_ (London: Sir Isaac Putnam and Sons, Ltd., 1947), p. 173.

[32] _Ibid._, p. 174.

[33] _Loc. cit._

been with the ethics of politics. There is no denial here of the importance of the study of ethics but it is denied that the study of ethics is sufficient for analyzing and understanding observed behavior. Ethics is the study of goals, more specifically, the study of alternative goals. The goals of a particular type of activity are not synonomous with the behavior expended to achieve these goals. The implication that if we understand the goals of a group is just not borne out in reality. We have no scientific explanation for this at the moment, but a tenable hypothesis seems to be that all group activity is not the result of goal orientation. We need a theory system of operational concepts (literally, concepts derived from scientifically controlled analysis of operations in administration) in addition to the ethical frame of reference of political science and traditional public administration theory.

The abstraction of concepts from experience in an emotional setting, with all the bias and skewing of perspective inherent in the introspective process, can never give us the interpretative power which we require in the management of present day organizational Brobdingnags. As Stene notes, many of the concepts we are not attempting to use might be combined as elements in a series of comprehensive alternative hypotheses in a research program, but they can hardly be defended as satisfactory interpretative concepts.[34] Anyone who attempted to determine precisely the point where communication ends and authority begins would find difficulty because of the ambiguity of these terms.

Adequate methodology. A second major need for re-orientation in public administration research methods is in the area of methodology. Stene holds that we have few methods adequate to our problems and no methods adequate for the analysis of our major problems.[35] Whyte supports this point of view. He maintains that the political scientist is naive who feels he is getting reliable data on the major problems.[36] What accounts for this grave absence of methodological sophistication?

One of the fundamental failures in our research efforts has been our repeated refusal to go to the scene of the studied behavior instead of slumping over books in some library. The collection of data is a very tedious process requiring a long period of planning followed by a considerable period of data collection. It is expensive,

---

[34] Edwin Stene, "An Approach to a Science of Public Administration," American Political Science Review, 34:1126, 1940.

[35] Ibid., p. 1125.

[36] Whyte, op. cit., p. 694.

time consuming and requires a number of highly technical skills. Whyte notes that in almost any case the political scientist considers secondary sources adequate.[37] This may be an exaggeration, but it represents a problem. We require data of the highest dependability. We require data which conform to intricate sampling designs. And data must be tabulated so that they can be manipulated statistically to fulfill a set of demanding mathematical assumptions.

<u>The use of experimental designs</u>. Once this gross defect is overcome there is need then to begin some positive measures which will lead to reliable research conclusions. The first of these is to include the use of experimental designs as a preliminary step in any research. Taking measurements and locating relationships in the social sciences is not as simple as it is in the physical sciences. The dimensions of a cube are universally determined by measuring the edge of the cube. In the social sciences we have not yet even agreed as to the primary dimensions of behavior let alone the manner of measurement. To make plain just what measurement we have taken in any particular case, we must go through a specified procedure

---

[37] <u>Ibid</u>., p. 696.

in a carefully prescribed sequence. This sequence is known as the experimental design. The results of any two investigations cannot possibly be compared unless the experimental designs are such that they include similar assumptions and a common unit of measure.

The cross-discipline approach. A fourth major reorientation needed in public administration research is the cross-discipline approach. As Moulton points out, "Science is a cooperative effort toward a united understanding . . . ."[38] This must be especially true of public administration if it is ever to be a science, for administration is a total cultural problem. Public administration does not correspond to any of the logical levels of nature as represented by the traditional academic disciplines. It is imperative that we bring the pertinent principles from all the social sciences to bear upon this total problem. From the practical viewpoint, this means that research in public administration should be executed by teams of researchers from the other social sciences, especially those closely related to the particular problem under investigation. The steady fractionalization of academic areas into smaller and smaller intensive areas represents

---

[38] White, The New Social Science, op. cit., p. 128.

one of the trends which will have to be overcome if we are to mobilize research techniques of complexity equal to our problems of administration.[39]

Dewey supports the need for the cross-discipline approach by laying bare the heart of the administrative research problem. Administration, he maintains, is not a self-contained natural entity. Administration is above all a process. It often involves group processes, personalities, great cultural processes in the institutions of our society, economic processes and it always involves legal procedures. The sociologist studying groups can disregard the subjective aspect of administrative groups and thus isolate this phenomenon from the rest of the process. The psychologist can restrict himself to the universals of personality and disregard the peculiarities of administrative role. The anthropologist, the economist and the lawyer or judge can likewise isolate that part of administration which interests him. The provincialism of the disciplinarian is one of the basic reasons, according to Dewey, for our inability to cope with many of our unmanageable social problems.[40] No one thing is completely independent of all else. Certainly we shall never comprehend the

---

39 Ibid., p. 55.

40 Fesler, op. cit., p. 14.

administrative process if we insist upon viewing it as a purely legal process. This leads to all kinds of confusion of means and ends. The objectives of administrative action are not the execution of the law but the provision of goods and services. It is this process of providing goods and services with which we must concern ourselves when we are judging the effectiveness of our theory system. The legal procedures must be carried out smoothly if an organization is to be efficient, but informal organization, a strong personality, or a violent change in the economic situation may condition the effective execution of the law.

Frequently a sociologist will note that administration is essentially a group process best studied with sociometric techniques. At the same time there are psychologists who would maintain the most important dimension of the administrative process to be individual behavior. Likewise, the anthropologist, the economist, the law student might present a case for their point of view. White lays down the fundamental proposition of obtaining ultimate value only in understanding all of these aspects of administration.[41] Administration is in essence a process of maintaining an equilibrium gradually.

If we are able to achieve these major reorientations

---

41 White, op. cit., p. 13.

in our research methodology in public administration there will still be those who maintain it is impossible to develop a science of administration. There will also be those who hold that a science of administration is not the answer. Perhaps by then we shall have come to realize that there is no dichotomy between science and art in the absolute sense. We are not faced with an "either--or" alternative. If there is such a thing as absolute science we do not know it yet. Actually our scientific efforts produce results which fall along a continuum. Some of these results are little more than very sound guesses while others surely approach true knowledge. The question is not really whether or not administration should be a science or an art; rather it is the degree to which it can or should be a science. This statement contains the clear implication that in any science there is still a place for intuition and experience as important guides. It is assumed that the rationale and techniques laid down in the following four chapters may be used to increase the proportion of science to art in administration.

CHAPTER III

CONDITIONS FOR A SCIENTIFIC APPROACH TO RESEARCH IN PUBLIC ADMINISTRATION

The quintessence of the scientific method is its universality. It is a structure into which we can fit different sequences of experimental operations, each rather simple in itself, but powerful when combined in a pattern, and is sufficient for the analysis of a great variety of natural phenomena ranging from the very simple to the extremely complex. This characteristic of the scientific method makes it a necessity for the further meaningful study of administration. We are long past the day when man's analytic equipment is sufficient for understanding administrative processes without mechanical aid. We must use some mechanical scheme which allows us to break the process down into simpler parts without shattering the equilibrium of the environment in which our analysis is made. The scientific method allows us to study a variable as if it were outside the complex natural situation in which we see it without destroying the relationships between our experimental variable and the rest of nature.

I. FALLACIES IN TRADITIONAL RESEARCH

One of the imperatives for the scientific method is our inability to get any determinate answer from a mere examination of facts. Many historians, economists and political scientists have spent their lives following Bacon's instructions to collect the facts as the first major step in any scientific investigation. They have been acting upon a common fallacy that any given aggregate of facts embodies a single generalization. Some hold that a generalization is in some way inherent in each unique set of facts and that through the use of some subjective process we must deduce it. Some of the fallacies in the traditions of political science research must be pointed up in detail before we can see the need for the reorientation suggested here.

<u>No objective manner to construct generalization</u>. There are two important fallacies in these beliefs. The first is that there is no known objective manner through which to construct generalizations. "No fact ever told an investigator what law is embodied. The investigator has to dig the generalization out of the accumulative mass of facts."[1]

The point brought out here by Lee is that the derivation of any generalization is a subjective process

---

[1] Harold M. Lee, "Theoretic Knowledge and Hypotheses," <u>Psychological Review</u>, 57:31, 1950.

producing concepts which have no reliable relation to the real world of experience until tested. Generalization is a process of establishing a class of things on the basis of some specific aspect or characteristic they have in common. If we set up classes merely on the basis of personal observation there is a real risk that we may have things in the categories because of a mistake in the identification of some characteristic. The way to increase the certainty of an accurate classification is to take measurements of each object which we classify to be sure it has a prescribed amount of the specified characteristic. This involves the use of a series of operations which are one of the set of operations in the scientific method.

Possible to make more than one generalization about a given set of facts. The second and more fundamental fallacy involved in the traditional point of view is the assumption that it is possible to make one and only one generalization about a given set of facts. Two of the most respected commentators on the scientific method have put it bluntly: "It is an utterly superficial view . . . that the truth is to be found by 'studying the facts'."[2] They point

---

[2] Morris R. Cohen and Ernest Nagel, An Introduction to Logic and Scientific Method (New York: Harcourt, Brace and Co., 1934), p. 199.

out that the study of the facts is superficial because it is not the facts themselves which determine the generalization which can be drawn, but rather, the order (pattern) of the facts; and it is the experimental problem which determines the order of the facts to be drawn. The experimental problem identifies the facts which are to be collated.[3]

> Again and again investigators have plunged into subject matter, sending out questionnaires, gathering a tremendous amount of data, even performing experiments, only to come out at the end wondering what it all proves, and realizing after years of industry and effort that the real difficulty has slipped through their fingers.[4]

These investigators substituted enthusiasm for the careful statement of the problem. Enthusiasm almost always fails as a basis for the collection and interpretation of facts because it is simply not true that any given body of facts presents one, and only one conclusion.

Any set of facts reflecting more than one factor will imply a number of alternative possible relationships. A set of dates at which our states entered the union and their population at time of entry might be used to show that there was a pattern in the sequence in which states entered the union. It could just as well be used to show

---

[3] Ibid., p. 200.

[4] F. S. C. Northrop, The Logic of the Sciences and the Humanities (New York: Macmillan and Co., 1947), p. 1.

there was no pattern in the entry of the states into the union. The whole thing rests upon our definition of pattern, our interpretation of the population figures, etc. We soon find that if we just assume certain things, we can show that the Civil War was caused by the sequence in which states entered the union.

One of the sources of error here is that the body of facts collected is not the right one for the problem under consideration. The writer knows of a City Planning Commission which spent over two hundred thousand dollars for the collection of data for a redevelopment program which was to be a model project without ever collecting a single fact about the social characteristics of the people who were to live there with the result that now, four years after the project's completion, it is considered an undesirable place to live because it is so "inconvenient." Any experimental problem will imply a set of relationships and data will be indicated by these. If we do not define our experimental problem in the light of a theory system or if we use a naive theory system, as was done here, we never perceive the need for all the pertinent data.

Another reason for the derivation of faulty conclusions is the failure to note that the abstraction of any generalization depends upon a set of assumptions. In 1950, a consumer research group specified that the Pontiac 6 was the best all-purpose car for the cost being built in the United States in that year. This stirred up a storm of controversy among producers, dealers and mechanics for they pointed out that this conclusion was based on the unstated assumption that the best all-purpose car was one with low maintenance costs and one with low gasoline consumption in city driving. This year the organization recommends three all-purpose cars on the basis of three sets of alternative assumptions possible here. This organization prides itself on its objectivity because it spends hundreds of thousands of dollars on the collection of facts upon which to base its recommendations to its clients, but it was forced to recognize that the facts could be used to justify the purchase of at least three different cars depending upon the assumptions as to its use. A more complete treatment of the question of inference, which is the technical name for the problem of matching facts and conclusions will be found in the second section of this chapter.

The essential point here, the one which eliminates the possibility of merely collecting facts and then divining generalizations, is brought out clearly by Northrop. He says that there is an infinite number of sets of facts in the universe which could be collected to derive any given generalization, for in the ultimate sense everything in the universe is related to everything else and therefore, if we wish to use the technique of fact collection by itself we must gather every fact concerning the relationship of our problem situation to everything else in the universe.[5] Such a solution is impossible. We are forced to fall back on techniques which allow us to limit our collection of facts to a restricted universe. If this is done intuitively, we have little opportunity to ascertain whether the sample of facts taken reflects the most important relationships or whether fundamental and superficial relationships are intermingled.

Research must depend on common sense knowledge. A third broad fallacy in the research described in Chapter II is that it depends a great deal on common sense knowledge. Throughout the development of our culture, different people in different groups have considered all of the major questions which we face today and they have developed solutions in which they have a great deal of confidence. But, for every major problem, we find we have not one answer, but a number of answers, each of which is held by an impressive (and often highly organized) group of people to be the "right" one. The conscientious administrator who takes the trouble to hear all of the points of view of everyone who

---

[5] *Ibid.*, p. 17.

has any contact with a problem will get a set of inconsistent answers for his trouble. The honest administrator will probably find that he has no reliable criteria by means of which to select one course of action (for he must usually act or be replaced) which seems "better" than another, unless he begins considering the premises underlying these points of view, at which point his personal biases enter into his selection of alternatives. The view taken of common-sense knowledge in connection with research is that it is of limited value except as it provides (1) hypotheses for investigation or (2) alternative goals for the development of new techniques of administration.

<u>Confusion between real and perceived situations</u>. A fourth fallacy in the traditional approach to research in administration has been the confusion between the true or real situation and the one which is perceived. Some investigators appear to be unaware of the fact that when they say an organization is "top heavy" or has too few "channels of communication" that they are imposing a verbal construct upon the existing situation instead of representing the true situation through their verbal symbols. Cohen, perhaps the leading non-symbolic logician of our time, notes that in attempting to discover the solution to some real problem we often defeat ourselves simply because we fail

to represent the true dimensions of the situation and thus fail to provide solutions to fit the real problem.[6] In fact, as the dimensions of the true situation become more subtle there is an increasing tendency to retreat to more abstract concepts. What is required is a technique which produces a definition of our problems in terms that refer directly to the true situation. This, again, is one of the steps of the scientific method: operational definition.

In summary then, we must conclude that methodologically, most research in public administration and political science is naive. This is no damning charge, for it must be remembered that we are evaluating it against something it does not pretend to be, but it does indicate the need for the consideration of alternative approaches. First, there is little awareness of the limitations of an approach based upon the analysis of facts. Second, there is a lack of awareness of the limitations upon common sense knowledge. Third, there is almost no sophistication as to symbolic representation, let alone mere semantic problems. The point of view here is that it is soon going to become essential for political scientists and public administrators to develop techniques for the analysis of some of the

---

[6] Morris Cohen, Reason and Nature (London: Kegan Paul, Trench, Trubner and Co., Ltd., 1931), p. 175.

more fundamental problems in this field and it is suggested that the answer to this need is the scientific method as it is now being perfected in the social sciences. The following discussion will consist of a section laying down the nature of the scientific method and a section describing the specific sets of operations which are imperatives from the scientific method.

II. THE GENERAL NATURE OF THE SCIENTIFIC METHOD

In a discussion of the scientific method it is natural that there should be some attempt to establish the meaning of science. Hartung gives three:

> (1) Science is a body of organized knowledge and a set of rules for obtaining and validating that knowledge. (2) Another meaning confines science to a certain method. (3) Science is simply that system of behavior by means of which man achieves control over his environment.[7]

The first definition will be accepted here since it clearly emphasizes the primary objective of science: the creation of a body of reliable knowledge through the use of a method.

Complexity of the scientific method. The most striking characteristic of the scientific method is that

---

[7] Frank E. Hartung, "Sociological Foundations of Modern Science," Philosophy of Science, 14:70, 1947.

it is not one method at all, but a complex of methods, each method being used for a different purpose. There is one method for the initiation of an inquiry, there is another for the exploration of an area of inquiry, and still another for the derivation of conclusions.

These several phases of the scientific method cannot be laid down with any finality. They merely correspond to qualitatively different operations that are necessary when we develop reliable knowledge. Every scientific investigation begins with a problem.[8] It is the definition of the problem which enables us to select only the segment of facts connected with a given phenomenon. Otherwise we should have to deal with a universe of facts. In defining the problem we "pin point" our interest by the statement of hypotheses which we shall use as the basis for the actual selection of the facts which we intend to collect. In the second phase we have a second complex of methods which we use to collect our preliminary data. The expenditures of great chunks of time and money in the development of instruments is characteristic of this phase. One of the objects of this activity is to gather our data efficiently. Another is to control bias by eliminating the possibility of coloring data with personal bias. Still another is to

[8] Cohen and Nagel, *op. cit.*, p. 392.

determine whether or not our hypothesis is even worth testing in the light of more extensive information about the situation to be studied. This exploration becomes the basis of our planning for the final stage of the scientific method. These three phases are fundamentally different in purpose, in technique used (logical analysis in the first and a fairly mechanical type of procedure in the second), and in result. In the evaluation of the data we undertake a still different type of activity and use a radically different set of techniques, usually statistics, to establish the relationship between the facts and the hypotheses.

Scientific method and the theory system. Cohen points out that a second characteristic of the scientific method is that it is at all times intimately related to a body of theory. In fact, it is fundamental in the scientific method that we have a mature body of theory to work with. The facts which we collect in the second phase are significant only in terms of the theory system we have established in the first. It is one thing to know that there is a reliable relationship between the speed of fire engines and the speed of ocean liners but it is quite another thing to know that there is sufficient meaning in this relationship to give us any power in the control of our environment.[9]

---

[9] Cohen, *op. cit.*, p. 107.

Isolated relationships like isolated facts are meaningless. They gain meaning only when placed in a useful body of theory. We give meaning to relationships by establishing that they are accidental, causal, associational, etc. and this is only possible when we are able to relate them to a mature body of theory.

One of the important results of this intimate tie between scientific theory and scientific method is that it is the means by which we are able to maintain a consistent body of theory. If we were to add concepts helter skelter to our theory systems we should soon find that we had a heterogeneous mass which was of almost no use to us because it would not facilitate inferences about any experimental situation. By maintaining a consistent body of theory new bits of knowledge that we are able to glean are cumulative in effect in that they are related to most of the other concepts in our structure and thereby give us new knowledge about them. The power of the theory system is increased in the same way film with a fine grain increases the power of a camera.

It has been said that the development of the atomic bomb was not carried out in the few years before 1945, but that the whole body of atomic theory had been evolving for some twenty years and that it was the addition of the last formula by the Finns which gave us the final bit of control which we needed to precipitate the bomb.

<u>Scientific method rests upon stated assumptions</u>. Cohen notes that a third characteristic of the scientific method is that it rests upon a complex set of assumptions and that these must all be stated.[10] It is this body of assumptions and the consistent nature of scientific theory which make possible the fourth characteristic of the scientific method--its flexibility.

<u>Scientific method is flexible</u>. A scientific theory system is made of a set of concepts and a careful statement of the relationship between these concepts. When a new concept is added which does not fit into the pattern of relationships of the whole body of theory we immediately inspect the theory system and the assumptions underlying it. We assume that if the new bit of knowledge is valid (that is, if it was derived through the use of a rigorous method), that the theory system must be incomplete in some way or the assumptions must be wrong in some way. Thus, it was necessary to remold the whole body of theory relating to the physical structure of the earth with the invention of the cloud chamber. The paths of particles in the cloud chamber could not be explained in terms of atomic theory.

---

10 <u>Ibid</u>., p. 110.

Although we had never seen either atoms or any other sub-atomic particles at this time, we were able to infer the presence of sub-atomic particles from the pattern of failures of atomic theory. Parodoxically it is the fact that scientific theory is internally consistent which makes it flexible.

<u>Scientific method is standardized</u>. A fifth characteristic of the scientific method is that it is standardized. There is extensive, and often violent discussion of each of the procedures which go to make up the scientific method. As the procedure gains acceptance the conditions of its use, its weaknesses and its advantages are all carefully laid down. The individual researcher would be foolish to disregard these rules in his application of a procedure unless he is attempting to test or improve it in some way. Standardization, therefore, is one of the great advantages of the scientific method.

With standardization of procedures it is possible for one group of scientists to work on one set of problems and another group to work on another set of related problems and for the knowledge in both groups to be integrated into a complete segment of knowledge about a problem. But this characteristic is at the same time a disadvantage. The key to the great progress of science in the last fifty or one hundred years has been the same as it has in our productive institutions--specialization.

One result of the specialization movement in science has been the selection of smaller and smaller problems for research. This has been valuable because of the greater detail of our knowledge. It has, however, introduced a problem. We are no longer testing broad theories which could be the basis of new theory system. We lack the means for the integration of a good deal of this specialized knowledge.

A second result of this specialization is that it has drawn those working on the different aspects of scientific theory apart and thus deprived them of the opportunity for cooperation necessary to the solution of problems in the twilight area between two or more disciplines. (Administration, economic planning, education, and social psychology are examples of these fields).

III. GOALS OF SCIENCE

A second means of gaining insight into the nature of science is through an examination of its goals and purposes. Science has played such an important role in our society in the last half century that it has become known as literally all things to all people. There are those who feel that science menaces our whole system of values and on the other extreme, those who feel that science is the key to our very survival. In answer to the question, "Isn't science a means of pulling ourselves up by our own bootstraps?" an eminent social scientist once told the writer, "Science is a very powerful tool, but don't expect miracles of it."[11] What then, are the actual things which we may reasonably expect from the intensive application of the scientific method?

We may probably state categorically that there is a single broad goal of science and that there are a number of sub-goals which fall together as specific objectives of the larger goal. This broad goal of science is to achieve a systematic interconnection of facts.[12]

> It is the aim of all . . . science to rise above the historical stage and become theoretic, that is, to attain the form of a theory or system in which all propositions are logically or mathematically connected by laws or principles.[13]

There is no ultimate value in isolated facts or even in isolated propositions. As pointed out above, the sum of all of these isolated propositions is common-sense knowledge with all its ambiguity and indetermination. Only a few

---

[11] Personal conversation with Dr. J. P. Guilford, Professor of Psychology, University of Southern California.

[12] Cohen and Nagel, *op. cit.*, p. 394.

[13] Cohen, *op. cit.*, p. 347.

scientists hold that we shall be able to develop a body of knowledge which will be in any sense absolute. The simple reason for this is that we have no way of identifying absolute truth. Our present criterion is closely tied to utility. As we discover new facts which have local utility, our theory systems have to admit them or they are modified. It is just this aspect of science which gives it such great value as a survival technique in terms of the long range preservation of the culture. Science is almost mechanically self-corrective. The fundamental operation of science, the collection of facts, is also the basis of this self-correction for as new patterns of facts fail to fit into the theory systems, we turn to the examination of either assumptions or theories themselves.

Another way in which a theory system can be combined with a research procedure to locate weaknesses in our knowledge is to identify a generalization which seemed universal but which is not applicable in some specific instances. The analysis of the situation and of the generalization will often lead us to a new conception of the true elements of the situation.[14] The result of disturbing any one of the relationships in this three-way equilibrium results ultimately in the propagation of a modified theory which

---

[14] Northrop, *op. cit.*, p. 29.

equips us with new insight into the operation of our culture.

Cohen and Nagel bring this sel-corrective aspect of science out in a slightly different context. They note that the method of science is not a means of progressing neatly step by step from one truth to the next, never to return to the problem again once knowledge is gained about it. Scientific method is essentially circular.[15] It is something like climbing a long spiral staircase. You begin at the first step and after a considerable series of steps and the expenditure of much effort, you find that you have returned to exactly the same position in relation to the building in which the staircase is located, except that you are several feet higher than you were the first time. From this position you are able to look at the same parts of the building that you saw when you started up the stairs, but you have a new perspective. In terms of the scientific method you have a new theory, one which is a more powerful explanatory tool because you have strengthened it by making it compatible with a larger universe of facts.

Hierarchies of generalizations. Boring brings out the first particle of the larger goal of science. He notes that it is a goal of science to simplify our knowledge

---

[15] Cohen and Nagel, op. cit., p. 396.

about nature by bringing it under a small number of generalizations.[16] That is, we must attempt to develop hierarchies of generalizations, some of which are very specific and aid us in handling our problems and others which are of great scope and indicate the relationship between the specific propositions. The reasoning underlying this point of view is brought out by Lee. "There is no other way of rendering anything in experience intelligible but by relating it to other things through generalization . . . ."[17] The process of generalization is a means of determining the relationship between things, and thereby a means of simplifying one aspect of the environment with which we must deal. Generalization is the heart of the scientific method.

One of the functions of this technique is that it is a rather mechanical means of breaking problems down into smaller manageable units without losing control of the interrelations of the smaller problems. The process of subdivision, isolation and finally solution of problems is carried to its application in the process of generalization. The process of generalization overcomes man's inability to react to a whole situation because he is unable to perceive

---

[16] Edwin Boring, "The Use of Operational Definitions in Science," Psychological Review, 52:243, 1945.

[17] Lee, op. cit., p. 35.

both wholes and parts in the same instant.

In connection with the process of abstracting certain aspects of situations which are important to the analysis of a particular problem and then putting these together in a proposition, or generalization, we must deal with this very difficult problem. What are the aspects of a real situation which one must take into account to successfully derive a generalization relative to the relationships of the situation. For example, what are the testable abilities of prospective employees which we must take into account in order to predict their probable success on a given job. We make no attempt to examine all of the abilities of the employee, his personality, his family, or his car. Measurements might show that we do not need to take these dimensions into account to make a successful prediction. Individuals utilizing uncontrolled observation might well feel that these were dimensions which were pertinent. The process of making a generalization is extremely subtle just because of this tendency for apparent aspects of a situation to be unimportant.

Northrop brings out another dimension of this problem of constructing generalizations. He points out that the object of making generalizations is to bring observed fact under our theory system. His position is that isolated facts, as we noted above, are meaningless. What we must do is integrate observed fact with our theories to give them meaning. We must make isolated facts a part of the foundation of a proposition. "Described fact is fact brought under concepts and to this extent under theory."[18] Fact expressed as a part of a proposition is fact brought under our conceptualizations. Therefore, to have new propositions is to have new theories.[19]

Strictly speaking, we can say nothing about a pure fact, that is, a fact which has been merely apprehended, except that we have determined its existence. By definition, only described facts have meaning.[20] Through the connection of a fact with a concept we are able to draw inferences about it from the relation of the given concept with the other postulates of our theory system. Unless we start all over again, we are not able to do this for a fact which is not brought under the theory system. As Northrop says:

> The great importance of concepts is that through their use we are able to introduce unobservable entities and relations into our theories and thus give additional meaning to observable phenomena with which they are associated.[21]

---

18 Northrop, _op_. _cit_., p. 35.

19 _Loc_. _cit_.

20 _Ibid_., p. 86.

21 _Loc_. _cit_.

The compelling reason for the use of theory systems is not that they have any magical significance, but rather that they are enormously efficient. If we can assume that there is a fundamental regularity in nature, then we can construct theory systems to match these regularities and merely by the classification of a fact, infer a great number of things about it from its place in the theory system. The only alternative to this is to build a new body of postulates to show the relationship of a fact to the rest of the universe every time we discover a new fact. It becomes apparent that the critical step in this process, if we accept the assumption underlying the theory system, is the manner of establishing the relationship between a fact or pattern of facts and the concepts in the theory system.

The relationship between a fact or a pattern of facts and a body of theory will here be called the epistemic correlation, after Northrup.[22] While one of the goals of science is the manufacturing of generalizations, a correlary activity is the establishment of dependable epistemic correlations. While this operation is implicit in most of the important commentaries on science it is explicit only in the most profound and unfortunately even these commentators are able to provide little detail as to the nature

---

[22] *Ibid.*, pp. 119 ff.

of the epistemic correlation and the manner in which we establish it. It is, nevertheless, one of the primary goals of science for it is a means of developing a total body of reliable knowledge. The epistemic correlation is the key to the problem of developing a meaningful interrelationship between the facts we are able to locate in a given situation and our theory system, a step without which our generalizations are useless.

Quantification. Quantification is certainly another major goal of science. The physical sciences are almost entirely quantitative while the social sciences have quantification more as a goal than a tool. Quantification is a powerful tool in science because it facilitates precision and thereby brings on confidence. In many areas of the social sciences there is simply no substitute for precision. The tolerances are so fine in the equilibriums with which we are dealing that our norms will slip from the black man to the red man to the white man without our even realizing it if we do not have discriminating measurements. Qualitative measurements often lead to only indeterminant identification of two or more alternative explanations and we seldom have a way of knowing which of these is acceptable, for we do not know the exact extent to which the variables in the situation are operative. This is not to imply that we require or can ever achieve complete quantification in the social sciences. The question of the balance between qualitative and quantitative measures will be taken up in a succeeding chapter.

Creation of a unified body of knowledge. Finally, one of the goals of science is the creation of a unified body of knowledge. It is not sufficient to develop a set of generalizations, and a precise set of measurement operations, and a reliable system of concepts if they are not unified into a theory system. ". . . no science exists without some form of unity. A mere aggregate of facts is never sufficient to constitute a science."[23] Cohen holds that it is the essence of the scientific approach that facts are seen as interconnected parts of a whole.[24]

The ultimate goal of all science is to develop theory systems whose scope is equal in extent to the perimeters of the several phases of nature. When we have this we shall have knowledge of all of the variables in the universe and we shall then have the basis for control over it because we shall have identified all of the variables related to any given process. This is only an ultimate goal, but it

---

[23] L. O. Kattosoff, "Philosophy, Psychology and Postulational Technique," Psychological Review, 46:67, 1939.

[24] Cohen, op. cit., p. 107.

throws light upon the problem of unification of theory systems.

We have inherited a number of sets of theory systems in the social sciences, roughly one for each of the principal disciplines of social science. These are so vast in themselves and there is such a complete isolation between them in terms of inter-discipline cooperation that it is virtually impossible to bring them to bear upon a single significant problem. Since most problems of consequence involve variances which require several of these smaller theory systems we have not been solving any but intra-disciplinary problems in the social sciences. So strong is the provincialism of our disciplines that there are few within the social sciences who are even calling for this type of activity. It is nonetheless one of the most significant goals of all scientific effort. Churchman feels so strongly upon this subject that he opens his work on the experimental method with this statement:

> To the scientist (this) essay has attempted to preach the moral that the process of raising a question and proceeding to reply to it is a process which eventually demands the active cooperation of all fields of research. We can no longer visualize the experimental scientist as the lonely worker who designs and completes his own researches.[25]

---

[25] C. W. Churchman, Theory of Experimental Inference (New York: Macmillan and Company, 1948), p. vii.

IV. USE OF MODELS AND POSTULATED CONCEPTS AS THE FRAMEWORK FOR THEORY SYSTEMS

The point has been made that the use of a theory system is fundamental in the concept of the scientific method put forward here. The function of the theory system will be amplified in this section. The theory system serves two distinct purposes. The first is as a means of controlling that part of the environment of a study which is not related to the experimental question. This will be discussed here. The second is as a means of providing meaning for the experimental results. This will be taken up in a later section of this chapter.

Theory system serves as a means of control. A serious problem which the social scientist faces is that of isolating his problem in such a way that all of the important variables are included within its limits of his investigation and the rest are excluded. As Cohen points out, "Everything is ultimately connected in definite ways with (every other) definite thing."[26] The scientist is caught between opposed principles. The progress of science in our time has been possible through the isolation of smaller and smaller problems, yet how can we find out about

---

[26] Cohen, *op. cit.*, p. 150.

any small problem when it is ultimately related to everything else. The answer is that we compromise in two directions. First, we do not pretend to get absolutely correct knowledge and secondly, we do not study all of the possible relationships. It is the first compromise which allows the other. In each study we make a choice, deliberately or by intuition, as to how many of the relationships involved in our area of study we shall include and therefore the meaningfulness and precision of our results.

To make an intelligent choice we must have some idea of the relationships that exist between the variables in our problem and the myriad of other variables in the study situation. This is one of the ways in which science becomes cumulative, for as we get more mature theory system we are able to infer more about the relationships which are involved and thus run less risk of failing to take account of one which is critical to the results of the study. It is in fact one of the functions of the theory system that through the systematic representation of nature in a unified conceptual structure we are able to tell a great deal about the problem under investigation before we ever collect the first item of data. The use of a theory system is not just a matter of its direct application to the problem of inference. There are a number of misapplications of this tool which are evident when one considers the details of its application.

The use of models or constructs. We use models or constructs to build our theory systems. It is the model which enables us to establish the relationships between the concepts or propositions in our system. In Newtonian physics, the model was established upon the basis of the causal relationships between the concepts in the system. It is the model which makes it possible for us to structure the concepts in the system and thereby represent nature. A model is roughly analogous to the steel superstructure that goes into the giant office building. The superstructure establishes the relationship between the rooms in the building. Once the building is in use, functional relationships may be established between these rooms by the imposition of a social structure, but this addition does not change the original geographical relationship between the rooms. The model which is built using the causal principle has now been named the mechanical model because it embodies such a simple and direct set of relationships.

The problem of models demands brief mention of the assumption about the nature of nature. The model is a representation of a number of the primary relationships which we have found in nature, through verbal symbols. (Topographical psychology represents an attempt to circumvent the almost complete lack of words to describe relationships by the use of a graphic language, but it appears that geometry lacks the subtleness with which the symbol represents the complexity of nature). Newton made the assumption that relationships in nature were all rather straight forward and that we could assume that if there was no causal relationship between two events, there was no relationship which counted in terms of our controlling these events. Mechanics, optics, inorganic chemistry, and until recently, aerodynamics were founded upon this assumption and our control has continually increased in these areas over the past two hundred years (in cases where the field has existed that long). There has been little indication that Newton made any mistake in his assumptions about the nature of physical relationships in these fields, and the mechanical model has been satisfactory.

Newton's model was so successful that others assumed his presuppositions about nature universal and biologists, anthropologists, organic chemists, even some economists and political scientists, accepted it as the principle for their constructs. They all soon found changes taking place which could not be accounted for under the model which they were using. Some explained these problems away, but others began the examination of the reasons behind the failures and came to the position that the mechanical model was not equal to the complexity of the natural phenomenon with which they were dealing.

SCHEMATIC DIAGRAM OF

THEORY SYSTEM

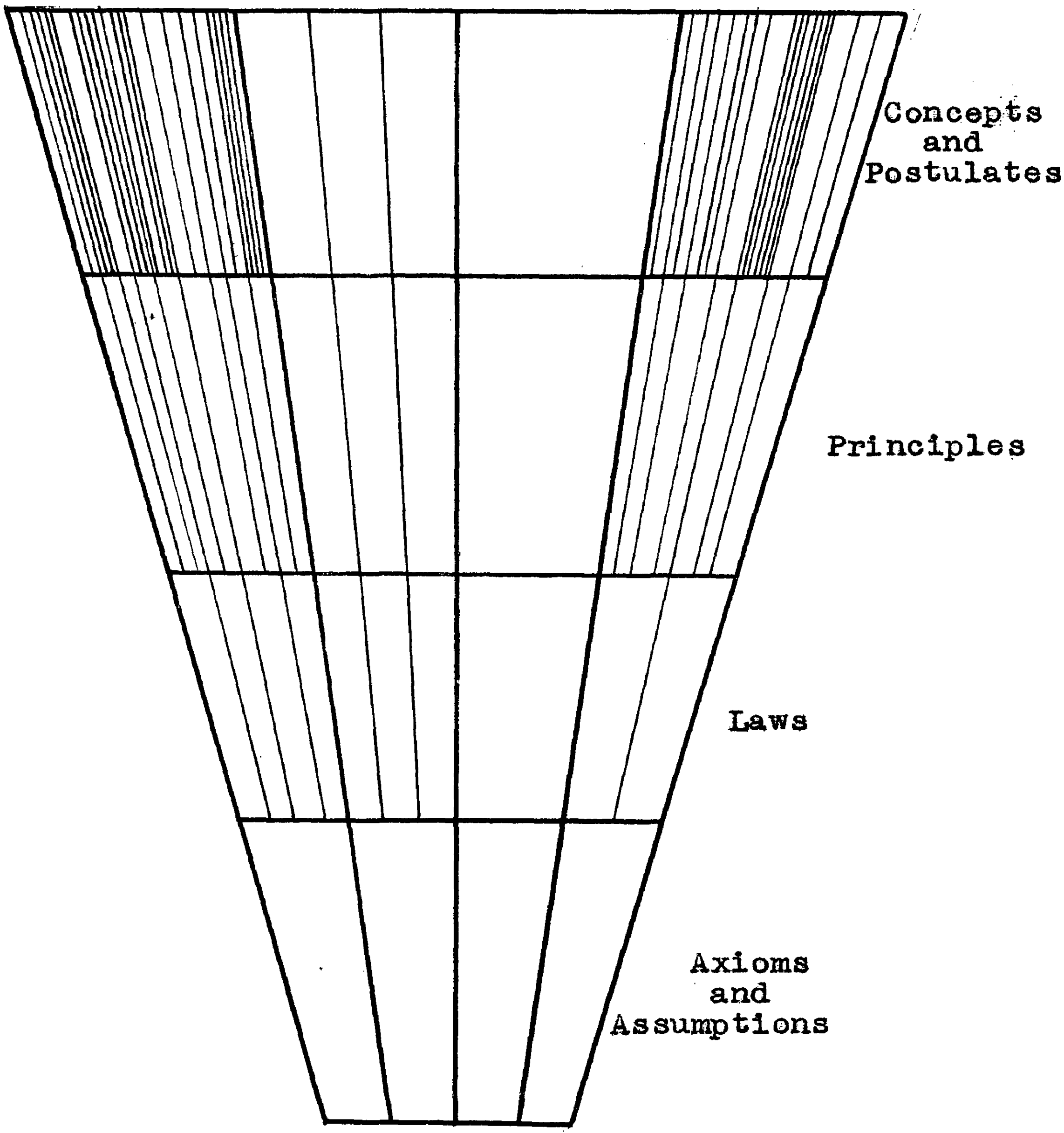

FIGURE 1

We have since found that in dealing with social phenomenon we need not only models which are constructed upon the mechanical principle but models which take into account the different kinds of causal phenomena. In the physical sciences the Newtonian model is now but one of several. With the appearance of the electron, the proton and other atomic particles whose mass was so infinitesimal that it could not be a real "cause," the electro-magnetic field theory became the model for a new theory system--a system created for the research on sub-atomic behavior.

For a time, the mechanistic school of psychology with its simple stimulus-response model made a strong attempt to reduce the ~~behavior of the~~ individual to the mechanistic model. The great controversy in this field is still not resolved, but there are few psychologists today who hold that all individual behavior can be fitted to this model. On the other extreme, there are those who maintain that all individual behavior must be examined in terms of multiple causation. The Gestalt psychologists and the followers of Lewin who hold to the field theory represent the translation of this point of view into practice. If one holds a high regard for the point of view of the "practical man" there certainly are few clinical psychologists who would hold to the simple causation construct. Allport makes this rather caustic comment. "The machine model in psychology had its origin not in clinical or social experience, but rather in adulation of the technological syntax of the physical sciences."[27] He feels that it is inept to attempt the formulation of but one model for the whole field of psychology, holding that both animal psychology and infant psychology require separate models and that the theories of either of these two fields have only limited application to the study of the adult.[28]

Seward takes issue with this, holding that while there are real and significant differences between the behavior of animals, children and adults, these are not qualitative differences, but only differences in degree. That is, differences which he feels could be placed on a continuum of complexity where the behavior of an adult is of the same type as that of a child, but merely more complex.[29] We need not decide this question here, but only point out that it reduces itself to a problem of the selection between alternatively plausible models when considered from the point of view of research.

This question of the nature of the differences

---

[27] Gordon W. Allport, "Scientific Model and Human Morals," Psychological Review, 54:183, 1947.

[28] Loc. cit.

[29] John P. Seward, "The Sign of a Symbol: A Reply to Professor Allport," Psychological Review, 55:293, 1948.

between the behavior of animals, infants and adults points up another characteristic of the model. The model for a theory system embodies the assumptions which we make about nature. The mechanical means which we use to do this if the theory system is written out is to place these at the bottom under the heading of axioms or assumptions. These are the generalizations of the system which form its foundation and we must accept them as true if we are to accept the system. These assumptions are the broad presuppositions upon which the experimental conclusions recorded in the system as laws, principles and propositions, depending upon their scope, rest. Thus, they are, in a very real way, manifested in the minor concepts of the system through the experimental procedure.

Northrop adds a tool for the better understanding of this problem with his dichotomy between factual social theory and normative social theory.[30] The first is theory which is false if it is not in accord with what exists. The theory systems which have so far been under discussion here all fall in this category. The second type of theory concerns what ought to be.[31] Most of the theory which we call political theory falls into this category with its

---

[30] Northrop, _op. cit._, pp. 256-7.

[31] _Loc. cit._

concepts of ideal democracy, justice and equality. As an aside, he points out that frequently people attempt to place these two types of theories in one system. The inevitable result, he says, is nonsense and he cites Marx's whole scheme as the best contemporary example of such a hodge-podge.

The importance of the distinction between normative and factual theory systems here is that in the selection of our axioms or assumptions in the factual theory system, we should probably turn to a study of normative values, for the selection of one assumption over another appears to be based upon the beliefs that the experimenter has about what ought to be.

Northrop points out another dimension of this problem by noting that there is often a transposition of the idea of the truth of a theory system with its proof. Proof is purely a formal within-system phenomenon, achieved by the application of a set of rules and operations to any proposition under the experimental process but it has nothing whatsoever to do with the ultimate truth of the proposition as a representation of some aspect of nature.[32] By a systematic analysis of the normative values which one wishes to put into a theory system the possibility of

---

[32] *Ibid.*, p. 83.

including contradictory or inconsistent axioms would be reduced.

Perhaps nowhere is it so clear that there can be little transference from the physical to the social sciences as in the discussion of models. The physical sciences can deal with the matter in terms of a simple principle of causation while the social sciences must deal with multiple causation in most areas of research. The physical sciences are dealing with an environment which is changing, but the rate of change is so slow that in most cases it is perfectly valid to assume a static environment in terms of the length of the experiment. In the social sciences with the exception of very small, and therefore rather meaningless, cross-sectional studies, it is impossible to correctly assume a static environment. This in itself adds a vast system of controls to social science method never needed in the physical sciences. Finally, and most important, the subject matter of the physical sciences is inert while man has goals and a conceptual power with which to handle these goals and with these goals and his conceptual power he has the ability to put his energy to the satisfaction of different goals as he feels the need. This necessitates the use of a theory system for the social sciences which is much more complex than any theory system the physical scientists have ever anticipated.

The chemical analysis of the body, for instance, provides a very simple explanation of its composition. We are far from having a complete system of theory for the explanation of the behavior of this organism, but already the theory system of the psychologist is staggering in comparison with that of the chemists.

The dangers involved in the use of models are brought together by Slone under the concept "reductionism."[33] He takes the position that in the social sciences we shall have to be prepared for the use of many types of models.[34] The fundamental reason for this (there are several practical experimental reasons) is that in moving from one area to another in the social sciences we reach a point where the model of one area ceases to provide real explanation for the events in another.[35] That is to say, the imposition of a model from one area upon another shows so few of the relationships in the second area that it tells us very little about the events in that area. Slone cites the opinion of many of the leading philosophers of science to prove his point. Certainly we may take this to indicate

---

[33] Eugene H. Slone, "Reductionism," <u>Psychological Review</u>, 52:214-223, 1945.

[34] <u>Ibid</u>., p. 216.

[35] <u>Loc</u>. <u>cit</u>.

that there is at least a strong case for this point of view, until further research gives us sufficient information to make a more definite conclusion.

One practical implication of this for administrative research is that we shall have to use one model for the analysis of group behavior, another for institutional behavior and another for individual behavior in the administrative situation. Whether the systems of sociology, anthropology and psychology can be transferred directly to this task is doubtful.

Levels in nature. Although it is to be given extensive treatment later, the concepts of levels in nature must be introduced here to make clear the problem of reductionism. It is held that there is no single entity "nature" which exhibits a consistent interrelationship of all its parts, but that there are at least five levels or streaks of nature which are relatively independent systems. Among these are the sub-atomic, the biological, the psychological and the cultural. It is of little use to attempt to investigate social phenomena involving groups in terms of biological theory systems any more than it is useful to use a chemical (atomic) theory system in the analysis of the individual. To say that the individual is composed of so many grams of iron, sodium, calcium, etc., gives us no means of understanding his psychological make-up. Likewise, it is of little use to attempt an analysis of a whole culture in terms of drives, interests and abilities, concepts which are a part of the psychological theory system.

Reductionism fallacy. Now one of the goals of science is to reduce complex phenomena to a small number of dimensions or variables so that we may systematically divide these complex occurences into simpler, more meaningful and more manageable events. However, the fallacy of reductionism involves reducing a complex event through a set of concepts which are (1) taken from a level or streak below the one in which the phenomenon appears,[36] or (2) too few in number or so simple in character that they do not adequately describe the phenomenon in terms of its relationships to the other phenomena of that streak or level. (Slone specifies only the first category but the writer feels that the term gains meaning from its extension to include the second one as well). The logicians have long been familiar with a fallacy of this sort and it is technically known as the fallacy of confusion of categories, though they do not use it in the context of nature's streaks.

---

[36] *Ibid.*, p. 217.

Slone holds that this problem arises through the "rather strange illusion of scientists" that they must reduce an event to the smallest possible number of variables or concepts in order to be "scientific." It should be noted on the side of the practitioners that they have long rebelled against this tendency, though without the use of the technical terminology. Slone's definition of reductionism is, ". . . the attempt to explain a complex interrelated whole in terms of simpler elements or parts or in terms of elements belonging to a lower level of phenomena."[37] The six types of fallacy which may result from such an attempt follow.

The first is the fallacy of isolation. This involves isolating a part of a whole and studying its structure or function without taking into sufficient consideration its relations with and dependence upon the whole. For example, the reflex cannot be studied except as a part of the whole nervous system.

The second is the fallacy of mathematical summation. This fallacy consists of holding that because the sum of all the parts of a mathematical situation are equal to the whole, the sum of the parts of some natural phenomenon are the equivalent of the whole. There is a very sound logical

---

[37] Slone, _op. cit._, p. 217.

reason for the distinction between the whole and the sum of its parts. Mathematics has never been presented as an empirical science. It is simply a very precise logical system which uses numbers and symbols instead of words, as does logic, and it has absolutely no basis in experience. Two plus two make four simply because when the system was established, four was defined as the sum of two plus two. Persons unaware of the logical origins of mathematics often feel that because we often successfully impose it upon natural phenomena, this constitutes proof of its correspondence to nature. Such is far from the case as any mathematician will readily point out. It is principally for this reason that statistical techniques provide meaningless results unless used in conjunction with some theory system, where the limits of correspondence are carefully laid out. A central problem in the use of one of the newest and most powerful statistical techniques, factor analysis, is that statistically derived common factors have little or no relation to reality because we take our data from the world of experience and manipulate them in the world of mathematics and are then unable to get back again because of the known difference in the two worlds. The broad alternatives are to develop a completely new theory system on the basis of factor analysis or develop experimental controls which allow us to interpret the statistical results from an empirical theory system.

The third fallacy is termed the fallacy of confidence. This involves thinking of nature as static instead of dynamic. It involves the problem pointed out in the preceding discussion of the differences between physical and social science.

The fourth is the fallacy of origin. According to this view, affects cannot differ from the origin of the cause. Thus, since man sprang from the ape, and the ape is of the genus _animus_, man can be analyzed in terms of the simpler animal theory system.

The fifth is the fallacy of metaphor or analogy. It consists of using terms from a lower streak to deal with phenomena in a high streak with the implication that the connotations of the word at the lower level apply to the new level as well. Slone does not point this out, but the fallacy works equally well in reverse. Thus, reference is often made to animal families, a community of ants and the like. While it conveys a useful meaning, for scientific purposes the lack of precision may be dangerous.

The last is the fallacy of models. This has been discussed extensively above.[38]

In summary, it is important to point out once more

---

[38] Slone, _op. cit._, p. 217.

that there is a fundamental limiting character of models which conditions their utility in the whole scientific process. Models are merely constructs which we impose upon nature to make its many undulations manageable when using our limited conceptual powers. There is nothing mysterious about a model which gives it any indefinable value and models should be rejected completely or modified as soon as new facts or better assumptions warrant such a change. Like organizations, models should be subjected to periodic reorganization programs. It seems that many people fail to question mathematical conclusions they feel to be incorrect because having had a long association with mathematics and understanding so little about their true logical nature, they have a certain awe of these conclusions. We deny the opportunity for the progress inherent in the self-corrective nature of the scientific method if we fail to challenge a model which we feel is weak.

V. CONCEPT OF SCIENTIFIC METHOD

The essence of experimentation is a set of conditions and procedures by means of which we create bits of knowledge.[39] Everyone is aware that the creation of knowledge

[39] S. Howard Bartley, Beginning Experimental Psychology (New York: McGraw-Hill Book Company, 1950), p. 30.

involves, at its heart, a set of experimental operations, but few realize the ambiguity which this oversimplified concept leads to when we begin the specific definition of the scientific method.

As a preliminary, it should be pointed out that there must be different scientific methods for each of the different levels of nature.[40] The dynamic nature of the cultural level demands radically different methods than the relatively static atomic level, as we pointed out above. This fact is taken for granted in the succeeding discussion.

No "all-encompassing" scientific method. Slone holds that all of science has been undergoing a process of reorganization recently because of the realization that there cannot be one scientific method--not even one for a single level.[41] We are coming to know that it is impossible to set down a method which will be suitable for the analysis of any type of problem within a given level. Northrop says that there are ". . . not merely different types of problems, but even different stages of inquiry into the same type of problem, call for different types of scientific methods."[42] He takes the position that people have

---

[40] Slone, *op. cit.*, p. 217.

[41] *Loc. cit.*

[42] Northrop, *op. cit.*, p. 33.

generally come by some sort of a priori concept of scientific method and never looked behind it to see that the actual use of the scientific method precludes the concept of a single all-pervasive method.[43]

We must conclude that there are a great variety of experimental procedures throughout the social and physical sciences which are the basic operations of science in these fields. These are combined into sequences depending upon the purpose of the investigation in that field. To be more specific, these procedures are combined into sequences depending upon the particular problem or hypothesis under investigation. The position taken here is that there is a broad similarity in the attack upon any problem in a scientific manner. To understand the real nature of the scientific method we must use a more explicit conceptualization.

Certainly there are several phases of experimental procedure common to all scientific investigations. All scientific investigation begins with the definition of some meaningful problem.[44] The definition of the problem may take somewhat different forms but it is inevitably followed by the collection of facts or data as a second phase. Following this there is some manner of interpretation

---

[43] Northrop, *op. cit.*, p. 33.

[44] Cohen and Nagel, *op. cit.*, p. 199.

of the experimental question in terms of some body of theory.

<u>Experimental design</u>. For a more meaningful interpretation of the nature of the scientific method we must turn to the concept of experimental design. It is the experimental design which corresponds most directly to experimental purpose or question. It is the experimental design which is used to answer a given question. The experimental design embodies the controls which produce a dependable conclusion. The position to be taken here is that there are definite classes of experimental designs on the basis of the pattern of combination of specific experimental procedures. There is a typical experimental design for anthropology, there are three for sociology and there are several types of experimental designs typical of psychology. These are the only fields of social science which have become experimentally sophisticated to the point where they have well refined and generally accepted experimental designs, but the primary purpose of this discussion is to evolve an experimental design appropriate for the analysis of administrative questions, and thereby to add one more experimental design to the fund of social science tools.

On the basis of purpose it appears that there are at least two fundamental types of experimental designs. The simplest is the one used for the answering of a specific experimental question. The need here is for dependable control of extraneous variables and for great precision in the measurement of the variables pertinent to the question. This design emphasizes the intensive examination of the variables involved. The other broad type of experimental design is used to examine broad questions which have important implications for the whole theory system. The emphasis here is on scope of variables considered. The purpose is to test the relationships inherent in a large pattern of variables. The full implications of these two classes of designs will be cleared after a consideration of the atomistic-wholistic controversy.

One might proceed directly to set down the experimental design which has been developed. An alternative organization of the discussion has been selected, however.

There are three broad problems which must be faced in developing the experimental design. The first of these includes the questions related to the selection of models, and hence experimental presuppositions, and will comprise the remainder of this chapter. The second of these involves the particular experimental operations now a part of social science research which might be appropriate for this experimental design. A rough inventory of these will be made in the next chapter. Third, we must review the statistical techniques applicable to research in public administration, for they influence the whole experimental design. This will be done in Chapter V. On the basis of these discussions, an experimental design will be formulated in Chapter VI.

The experimental design is the device by means of which we integrate the operations which we use to make a breakdown of the total research objective into a logically consistent and determinant measurement of fact. None of these operations alone tell us anything that might be used to answer our experimental question, but each of these operations provides us with a part of the answer. The statistical analysis is meaningful if it has been related to the presuppositions of the first operation and the selection of the sample in the collection of data, and to the formulation of alternative hypotheses. The contribution which a given step makes to the determination of an answer to the experimental question is contingent upon its proper relation to the other steps. The discussion of these three large problems will point out specific relationships which are involved.

VI. MATERIALISTIC-PHILOSOPHIC CONTROVERSY

The development of social science has been dominated by a conflict involving two philosophical questions. The first problem concerned the differences between the social sciences and the physical sciences. There have been many who felt that the physical sciences should be the model for both the procedures and theory system of the social sciences. One reason for this point of view is the exaggerated prestige of physical science in our culture, and the natural feeling that by copying the mechanical pattern it would be possible to attract some of the recognition inherent in it. But it was also largely due to the sincere belief that the mechanical model was ideal for the social sciences.

Once there was agreement that there was both real differences and real similarities between the subject matter of social and physical science, the controversy was confined within the social sciences. The central issue is what are the criteria of scientific truth.

> The boiling cauldron of opinionated clashes between behaviorists, psychoanalysts, and gestaltists boils down to one essential core of disagreement; what are the criteria of scientific truth in psychology, that is, what are the rules for verifying psychological hypotheses?[45]

Must everything be quantified? Must we select only those questions whose scope is such that we can use operational definitions in our hypotheses? Do we have to restrict ourselves to "public" data, and is there no place for

---

[45] Daniel Brower, "The Problem of Quantification in Psychological Science," *Psychological Review*, 56:327, 1949.

introspective data? These and other questions have been the issues in the conflict which has run so deep as to split social scientists into a number of camps, each producing a type of research characteristic of the position taken on these insoluable problems.

The materialist-philosophic controversy has not split the social sciences on the basis of disciplinary boundaries. It has, rather, split the disciplines into parts and these fragments of disciplines have formed themselves into camps. The discussion of the five or six issues in the controversy would be tedious. Therefore, the discussion here will be oversimplified to include only two points of view. The positions taken by the operationists and the position of the so-called hypothetico-deductive group.

As background for this discussion it should be pointed out that the controversy has roots in the discussion of the Greeks. It began in the physical sciences with the introduction of the relativity theory. The consequent destruction of the certainty that there were some things that we could be absolutely sure of such as mass, space and time was swift. The emergence of sub-atomic phenomena upset the traditional mechanical theory systems of the physical sciences. There were those who felt the whole foundation of physical science was threatened. Gengerelli notes that:

> Within the last thirty years this overwhelming confidence in what might be called the naturalistic view of knowledge and the universe has been seriously challenged in the exact sciences by the conceptual obstacles which have been uncovered in mathematical logic and theoretical physics.[46]

Physical scientists are now in the process of "riding out" this upsurgence and many of them feel that their theory systems are richer for the experience. It has had the affect of striping their axioms and assumptions of a lot of "dead wood."

As Gengerelli notes, we in the social sciences are less sure of our conceptual structures and have become involved in a series of interchanges which have involved the examination of the very foundations of social science. Psychometrics and sociometry had just been established when the philosophers questioned the whole basis of the measurement approach to social problems. They could not halt the rise of this movement though, for coupled with the conceptualizations of the behavioristic movement it has dominated the history of social science in the past two decades. Looking back it is easy to see the reason for this rapid dominance. We were able to solve problems with this method which had been beyond solution with an extensional

---

[46] J. A. Gengerelli, "The Dichotomy of Science and Philosophy," Psychological Review, 44:117, 1937.

approach. But the expansion of this movement laid bare its limitations and during the past several years the philosophers have had something they could really "get a hold on."

Operationism. Operationism, the current name for this approach, has been called merely a new name for an old problem by Feigl, who points out that in the end it concerns itself with making the terms of science more precise.[47] Social scientists have always been faced with a whole galaxy of problems which the physical scientist does not face because of the difference in the subject matter of social science--man. The physical scientist can represent his data and his concepts through geographic or topographic systems. Social systems are inherently so complex and of such tremendous variety that it is impossible to use such simple techniques of representation and the social scientist has been forced to rely upon verbal symbols to represent his data and his constructs. Boring feels that this is the heart of our problem and the central difference between the social and physical sciences.[48]

Operationism, which has been characterized as the

---

[47] Herbert Feigl, "Operationism and Scientific Method," Psychological Review, 52:250, 1945.

[48] Edwin Boring, "Rejoinders and Second Thoughts," Psychological Review, 52:280, 1945.

methodological counterpart of behaviorism, was initially an attempt to overcome the ambiguity of conceptual symbols by tying them down to actual behavior, thereby providing concrete referents for all of the important terms in psychological theory. Operationism in the physical sciences, as propounded by Bridgeman, father of operationism, was an attempt to meet the conceptual reservations imposed upon such concepts as length and time by the developments in theoretical physics mentioned above. In the social sciences operationism has been mainly an attempt to solve the problem of developing meaningful theory systems with terms of maximum precision. Usually as the scope of a theory system increases, its major terms lose their referents. The heart of the controversy has been to attack our current theory systems as being needlessly ambiguous and urge a return to the mechanical model and the precision inherent in it.

A representative of extreme operationism says:

> Operationism may be defined as the practice of talking about (1) observations, (2) the manipulative and calculational procedures involved in making them, (3) the logical and mathematical steps which intervene between earlier and later statements, and (4) <u>nothing else</u>. So far, the basic contribution has come from the fourth condition and is, like it, negative.[49]

In his final point Skinner represents the views of many

---

[49] B. F. Skinner, "The Operational Analysis of Psychological Terms," <u>Psychological Review</u>, 52:270, 1945.

operationists in that the approach has not yet devised the specific techniques implicit in its attempt to eliminate the ambiguity which accompanies the hypothetico-deductive approach. The more moderate definition of Feigl represents the views of the moderates who accept some of the limitations which the extremists deny.

> Operationism is not a system of philosophy: It is not a technique but a formation of concepts or theories. It will not by itself produce scientific results. Operationism is, rather, a set of regulating or critical standards. In the light of these critical standards the meaningfulness and fruitfulness of scientific concepts may be appraised.[50]

His six criteria of an operational theory system are: (1) logically consistent; (2) sufficiently definite to be repeatable; (3) empirically rooted to observation; (4) technically possible; (5) repeatable; (6) concepts which are used to allow maximum prediction.[51]

Feigl brings out the striking characteristic of operationism which is at the same time the source of its utility and the root of its limitations--that is, operationism is a means of constructing theory systems which are so closely related to action that they do not contain the ambiguities of the traditional logical construct. There are two core questions here which throw the essence of

---

[50] Feigl, op. cit., p. 258.

[51] Loc. cit.

operationism into relief: (1) what do we mean by the words and symbols which we use, and (2) how do we know that what we say through these terms is true?[52]

These simple questions survey the whole body of knowledge which we customarily think of as referring to the ontological problem. That is, how do we know whether what is said to be true represents that which actually exists in reality? In the broad sense all knowledge is contingent--contingent upon its being such a precise representation of nature that no fuller replica can be found. The operationist declares that his operational concepts should be expected to have a much lower contingency value than traditional concepts because of their direct dependence upon behavior. Bridgeman's classic example of the operational definition of length illustrates this idea. If length is defined as the measurement operations which we use to determine length, then there is almost a one-to-one relationship between our verbal description of length and the operations which we used because we can use words which are so definite that the operations are readily repeatable and we thereby eliminate almost all of the opportunity for confusion latent in other definitions of length.

At its core then, operationism is a new technique of

---

52 Ibid., p. 250.

definition. One definition of a definition is that it is a statement of equivalents: a term which is the equivalent of some event or operation would then be called the definition of that event.[53] An alternative is that a definition is the specification, or the limitation, of the meaning of a term or symbol. In practice these become the rules concerning the use of these terms or symbols.[54] While the first definition indicates the advantages which the operational definition has as a method of representing equivalents, the second points out its advantages in comparison with logical definitions. The operational definition has concrete terms of reference. These may be things, movements, readings of dials or scales, etc. In each case they are observable. An operational definition would never include a term such as "resistance" in describing electricity because this is a hypothetical entity. It does not exist but is merely a ratio of potential to actual conductivity. With the finest instruments and techniques we could never observe the referent of a term of this sort.

In public administration we use terms such as these to define other terms and then we use these to define still

---

[53] Boring, "The Use of Operational Definitions in Science," op. cit., p. 244.

[54] Feigl, op. cit., p. 250.

other terms. It is no wonder that our words like "federalism," "status," and "budgeting" have a good deal of ambiguity. Even if a term such as "budgeting" has concrete referents we run the risk of the fallacy of faulty regression if we do not use the operational definition. Thus, if we define budgeting as the process of planning the distribution of funds for a jurisdiction for a given period we have a number of words which require definition if we are to make an intensive study of budgeting. Which of the processes which contribute something to the budget activity shall be included under the term "process?" When we made the definition we may have had a number of them in mind but the person who attempts to apply our definition traces the regression down to different specific activities and he may come to disagree with our definition merely because of this failure.

The operationist points out that we should take all of the specific operations which go into the process and group them under larger operations. These should then be grouped under a set of still more inclusive operations and these used as our definition. If there is any question as to what our broad definition refers to we merely trace down through the levels of abstraction to the bottom again and offer the questioner the opportunity to observe the same operations which we initially observed.

The definition of terms operationally leads to the following additional advantages. First, it gives us insight into the nominal nature of all definitions and thus eliminates the lack of flexibility in our logical definitions by immediately indicating the need for a change of reference for a term. Thus, as soon as we install a performance budgeting scheme we see that there are different operations involved and we change our definition of budgeting to accommodate this new operation. The advantages of such flexibility are evident in relation to administrative communication concerning the budget process.

Secondly, it is held that the operational definition eliminates the possibility of the concealed contradiction.[55] The concept "liberty" for example, soon leads to a contradiction when we begin discussing the limits of liberty for we come face to face with the problem of liberty versus license. An operational definition of liberty, while a staggering undertaking, would illuminate this concealed contradiction in the beginning.

Third, it is held that the operational definition gives us more precision in experimentation because we are more definite in designating what the term refers to. This

---

55 P. W. Bridgeman, "Some General Principles of Operational Analysis," Psychological Review, 52:246, 1945.

is the most important advantage of the operational definition for it has the additional function of decreasing the contingency of our knowledge by establishing a stronger relationship between our concepts and the natural phenomena which they are assumed to represent. In a field where the central terms include unity of command, communication, coordination, and hierarchy, the experimental advantage of such definition will be great.

The most powerful point which the operationists have in their favor experimentally is that operational definitions reduce the necessity for the assumptions in the traditional theory system.[56] One problem in developing theory systems which describe or represent segments of the cultural process is that we must make assumptions to give our explanations meaning. Thus, in studies in political dynamics we make the assumption that if a person votes the Republican ticket he holds the beliefs of a "typical Republican" in the area in which he lives and we build a theory system embodying experimental conclusions on this basis. If we find upon further investigation that our assumption was partially incorrect or even completely incorrect our theory system is faulty and we must start all over or revise it. This scheme represents the most

---

[56] Feigl, *op. cit.*, p. 253.

rigorous deductive system which we know of at the present.[57] The elimination of axioms and assumptions is a revolutionary approach for it allows us to attempt pure induction. Until now the primary method of science has been to set up a deductive system and then attempt to verify it through induction. Note, for instance, that when we carry out the whole thing by induction there is no need to study normative values as the background for assumptions.

VII. ISSUES IN THE MATERIALISTIC-PHILOSOPHIC CONTROVERSY

One dimension of this controversy is the atomistic-wholistic conflict. The issue here is whether or not we have the meaning of a whole when we know the sum of its parts. It is admitted that the dimensions of a brick wall are clearly a function of the dimensions of its parts and that the whole of a book may be accounted for in terms of its pages, or its lines or even its words. Both sides agree upon this. These who take the atomistic point of view, however, feel that this is all that we must know to account for anything. Those who take the wholistic or phenomenological point of view hold that under certain conditions we can account for things in this way but in

---

[57] Evert W. Hall, "Some Dangers in the Use of Symbolic Logic in Psychology," Psychological Review, 49:153, 1942.

some situations we must look to the functional relationship of these parts for their full meaning. Thus the parts of an engine do in fact account for the whole engine, but when the engine ceases to run, a knowledge of merely what the parts are and what they are made of and any other facts which we might have about the parts would be of little aid in locating the source of engine failure. We must know the functional relationship of the parts, and this functional relationship is said to be something qualitatively different from the sum of the parts. ". . . the wholes exist in their own right over and above the parts or ingredients."[58]

The operationist denies that there is any significant meaning to a whole not inherent in the sum of the parts. Small operations are the obvious parts of larger ones and are in fact, the sequences of the larger operations. To understand the total operation one must merely understand the smaller ones. Operationists have been attacked continually for this assumption.

Public versus private data. A second point of conflict in the materialistic-philosophic controversy concerns the use of introspective data. The operationist holds that we should study only behavior which can be ovserved. This

---

[58] Slone, op. cit., p. 214.

is so-called "public" information in the sense that anyone who wished could observe it. The presence of pain, such as a toothache, a strictly intra-organic reaction, is termed "private" in the sense that it must be described by the subject to be known. Skinner points out that, "The simple fact is that each speaker possesses a small but important private world of stimuli."[59] This is one of the differences between social and physical science. The social scientist is dealing with a complex organism with attitudes and all sorts of other verbal phenomena which cannot at present be observed. The physical scientist has the assurance that with the correct instrument he can observe all the pertinent dimensions of his subject matter.[60] If we insist upon studying only observable behavior, this limitation would mean that we could not study any of the mental phenomena of man directly, but only as they are manifest in action.

There is some disagreement even among operationists on this point. Pratt feels that the issue should be disregarded on the basis that "Every datum in every science starts life as a private experience."[61] He feels that

---

[59] Skinner, op. cit., p. 271.

[60] Boring, "Rejoinders and Second Thoughts," op. cit., p. 280.

[61] Carrol C. Pratt, "Rejoinders and Second Thoughts," Psychological Review, 52:289, 1945.

since all of the data of science have at one time been mediated by the senses somewhere between their observation and their tabulation we might as well accept introspective data. He concludes that since the primary function of science is to provide a meaningful interrelation of facts that such a technical question concerning the data of science is out of perspective.[62]

Skinner holds that the public--private relationship is not real because there is not a one-to-one relationship between them. That is, there is not a public experience for every private perception and therefore there is no way of ever finding out about private experience if we restrict ourselves to the public universe.[63] Skinner flatly suggests that we turn to the causal analysis of private experience, he calls it verbal behavior, through the use of logic. This is in contradiction to the point of view of those operationists who feel that logic should be rejected so that we can eliminate assumptions from our theory systems, but it seems to be of little bother to Skinner in this instance even though it is inconsistent with his total position.

Bridgeman takes the view that the difference between

---

62 *Ibid.*, p. 290.

63 Pratt, *op. cit.*, p. 293.

public and private knowledge is fundamental. It is so fundamental that we actually need two languages to deal with it effectively since we think differently in these different "modes."[64] He feels that the history of man is filled with problems which result ultimately from the lack of a language with the sensitivity to reflect these two modes and that the future of man depends upon the development of a dual language by means of which we can express our different types of feelings.[65]

> Simple observation shows that I act in two modes. In my public mode I have an image of myself in the community of my neighbors, all similar to myself and all of us equivalent parts of a single all-embracing whole. In the private mode I feel inviolable isolation from my fellows and may say, "my thoughts are my own, and I will be damned if I will let you know what I am thinking about."[66]

He feels that our whole culture attempts to surpress the individual mode and enforce conformance to the community mode. The implication of this point of view is that we are justified in hewing to the public mode merely because of its extent and that there is a whole second universe which we need to investigate intensively. Bridgeman denies the implication that we should be attempting to

---

[64] P. W. Bridgeman, "Rejoinders and Second Thoughts," <u>Psychological Review</u>, 52:283, 1945.

[65] <u>Loc. cit.</u>

[66] <u>Loc. cit.</u>

develop a single integrated body of knowledge in the social sciences because there is such a highly significant qualitative difference between the two modes. (It is the writer's opinion that if Bridgeman knew enough about psychological theory--he is a physicist--to be familiar with the role concept, he would never have been moved to raise this objection.)

Quantification of social data. A third issue in this controversy involves quantification of social data. "Quantification refers to the process of expressing our observations of phenomena in terms of numbers or other discreet figures."[67] High among the reasons for the importance of quantification in the social sciences is the prestige which inhers to the use of measurements in our culture. During the period when social science was attempting to get launched in the face of the myth that you cannot ever study people because they are too complex, the prestige which could be placed behind the conclusions that were derived by merely using quantitative methods was important, even critical. Another prominent reason for the place of quantification in social science is the precision which it provides. When we depend upon the uncalibrated

---

[67] Brower, _op. cit._, p. 325.

judgement of some individual we must be content with measures such as good and bad, adequate and the like. When we use tests or scales or some other quantitative device we are able to obtain much finer measures, and in cases where there is a small tolerance, as in the last presidential election, accuracy of measurement can mean the difference between a reliable and an unreliable prediction. With this accuracy comes the capacity to answer more specific questions, and this is another reason for the place of quantitative methods in social science. Ultimately, we shall have to be able to handle specific questions if we are to be able to justify social science.

Bower feels that in addition to these functional imperatives, the spread of quantification in the social sciences has been encouraged by: (1) the way it naturally fits in with our mechanical culture; (2) our unwitting attempt to emulate the physical sciences; (3) the wide spread and still growing use of mental testing.[68] To these, the writer adds the identification of quantification with science and the desire of social scientists to be accepted in the public eye as such.

<u>Limitations of operationism</u>. Operationists have

---

68 <u>Ibid</u>., p. 326.

said that their system of definition and verification reduces significantly, if it does not eliminate, the need for assumptions in a theory system by representing nature directly through operations. This assertion has revolutionary implications for scientific activity if it could be substantiated. We may safely say at the moment it has not been adequately developed, for it is unable to meet the barrage of criticism leveled at it from all sides. One of these charges is that the operational definition will not do what the operationists claim, namely represent nature directly. Hall assails the point that operational definitions are the answer to the problem of representing complex phenomena for he says that the operational definition, while extensive, could never include every one of the details of the operation.[69] Even if we were able to include all of the elements of the operation in the definition we should have to convert these to a greatly oversimplified symbol when we attempted to apply statistical analysis to our data and then we would have gained little in terms of the meaningfulness of our results. Granting even that we could get a true representation through our symbols, Hall holds that we are buying rigor at too high a price. "Deductive rigor . . . , can be purchased at too high a

---

[69] Hall, _op. cit._, p. 148.

price . . . it may be attained at too great a sacrifice of comprehensibility."[70] The writer would certainly agree with this. The ultimate goals of science should never be buried in a mass of Hegelian-like verbage if we are to realize the potentiality of progress inherent in the self-corrective nature of science.

Hall's second charge against operationism is that it simply cannot do what it has proclaimed. He holds that it is impossible to get a set of operations which are not ambiguous.[71] We may have several operational definitions of the same phenomenon, for instance. It is difficult to imagine a single operational definition of coordination, for instance. What is to be gained through the use of several definitions representing roughly integration of authority, function, and communication, respectively? Even if we hold that these definitions are the same or that they represent three things which are actually different we gain little, for there is inevitably overlapping because of the nature of coordination. Statements which mean exactly the same thing are said to be equipollent.[72]

---

[70] Ibid., p. 153.

[71] Ibid., p. 148.

[72] Christian O. Weber, "Valid and Invalid Conceptions of Operationism in Psychology," Psychological Review, 49:65, 1942.

Equipollent definitions are held by the operationists to be characteristic of the conceptual approach; however, is it possible to eliminate equipollent definitions by merely using a different mode of definition if the nature of the cultural process under consideration is itself ambiguous?

We see, finally, that the actual position of the operationist rests upon the assumption that there is no ambiguity in nature and that it is the conceptual approach which is faulty. In the light of our knowledge of real administrative processes, is it possible to maintain that they are actually without natural ambiguity? The writer feels that we must admit that we could develop a tenable position on either side of this question. He prefers to maintain that nature is ambiguous.

It is also denied that we can really represent nature through operational definitions. This is held even by most moderate operationists. Weber holds that there is a dimension of reality which is lost by the operationists.

> I conclude that the object of knowledge cannot be incorporated in the operation by means of which we know it. There is a realm of brute requisiteness which cannot be "liquidated" by the simple device of identifying operations and concepts. Operationism of this sort may ignore the systematic order or lack of it in the things upon which operations are made, but the systematic order or lack of it

> will reappear among the operation. We are always under the necessity of assuming that the interlocking system is there, and that it is our business to find it.[73]

If one accepts the finding of the interrelations among facts and its statement as one of the goals of science, Weber's objection must be considered, for it is implicit in this goal that there is a fundamental order in nature. Is not the whole justification for the existence of science that it seeks to find the order of nature and thus provide man with understanding and control of his environment? The issue, then is what level of order of nature is the proper subject matter for science. That is, are we to limit ourselves to operational terms and therefore, the lower levels of nature or shall we utilize logical tools and explore the upper levels. Acknowledging our lack of sufficient grounds for the selection of the proper levels we take the most inclusive possible answer and maintain that science should provide us with description of all of the levels of nature. Even in the light of this position, Weber's criticism indicates that we should not attempt to place all our eggs in one basket and restrict ourselves to the more superficial levels of nature lest we lose the chance of getting the more meaningful explanations possible

---

[73] _Ibid_., p. 67.

with the more powerful descriptions of the fundamental order of nature. While we may have to qualify Weber's denial that operational definitions represent nature, it appears undeniable that there is at least the danger that we should lose our hold on the larger segments of our culture if we attempt to take the operational structure as a true and complete representation of nature.

Perhaps the most important charge against operationism is that it has gross semeiotic limitations. The great defect of operationism is that it denies the need for a conceptual system which is used to establish meanings.[74] It is admitted that operational definitions may be a means of actually representing nature through symbols, but the charge is that such a system is only a little more meaningful than nature itself and that therefore such a structure has very commonplace uses. It is pointed out that the harnessing of electricity, the subjugation of the atom and the insight into the operation of the mind, all of them practical matters of great importance, have all been facilitated by the use of hypothetical constructs and/or intervening variables. In the development of knowledge about the atom we established sub-atomic particles without ever

---

[74] Bergilius Ferm, A History of Philosophic Systems (New York: The Philosophical Library, Inc., 1950), p. 550.

having seen the atom itself. Such postulations are termed hypothetical constructs. If we find that our postulation is correct by discovering the existence of the thing or relationship postulated we say that we have found an intervening variable. It is characteristic of science that the hypothetical construct of today becomes the intervening variable of tomorrow and that with the discovery of the intervening variables at the sub-atomic level we proceed to the hypothetical constructs of the sub-atomic level. This is considered important--it is, in fact, the very heart of science.

The core of the scientific process is to proceed from level to level because it is in this way that we are able to develop the means of a unified body of knowledge. Cohen states that, "It is of the essence of the scientific method that facts are seen as connected parts of a greater whole."[75] It is this characteristic which differentiates science distinctly from common sense knowledge where we have fragments of reliable knowledge about many events, but we have no reliable interrelationship of these fragments and thus, no broad control.

This point can be illustrated by the development of the periodic table and the system of knowledge that

---

[75] Cohen, _op. cit._, p. 107.

accompanies it--chemistry. Before the development of the periodic table we had knowledge of the actions of gases as a distinct system of knowledge and another system for the description of the action of solids and still another for the description of liquids. With these three systems we were able to control the behavior of these three types of substances as long as they held their original state, but we lost control when they became unstable or when they shifted completely to another state. With the development of the concept of the atom we had a conceptual entity which was universal in the inorganic portion of the physical sciences. The different states of matter were no problem with this concept for we soon found that the only difference between liquids, gases and solids was in the density of the atoms per cubic unit. Our whole approach to chemistry was changed in that all of the different segments of knowledge were related into a total body of knowledge. Perhaps we have the same problem in administration where we must account for the behavior of people in institutions, in groups and as individuals.

The increase in our control over physical bodies with the creation of this unified body of knowledge was spectacular. Now we have the physical sciences rent again with the emergence of sub-atomic particles and while it introduces a good deal of confusion, it offers the possibility of providing the discovery of a new level of nature and the union of organic and inorganic. One of the practical results of such a discovery would be the development of knowledge about the so-called "secret of life"--the explanation of organic cell structure and behavior. One cannot overemphasize the point that this would be impossible without the unification of great areas of knowledge so that we have intimate knowledge of the relationships involved in cellular structure.

Now the point of central importance here is that there may be operationally independent "levels" in nature but we have no way of telling whether there are or not. The only level in nature which we know anything about with certainty is the level in our theory system <u>about</u> nature. We do not know whether electrons exist in nature either, in the sense that we have seen one or more electrons. But we do not need to know of the actual existence of these hypothetical constructs to use them to explain natural phenomena. The hypothetical construct is developed on the basis of a set of conditions in such a way that it conforms perfectly to these and that is all. The hypothetical construct, as far as we know, is a fiction which we develop to represent a set of conditions without any necessity of ever "proving" its existence, for it is a purely functional entity. For example, it seems improbable that except for the most arbitrary definition we shall ever discover a single social phenomenon which corresponds to the broad construct, "morale." We needed a construct to cover a variety of results (conditions) and this term had a long history in our culture and we grouped first one and then another condition under it until we had a great patchwork quilt made up of all the things which could not be explained through our reliable constructs and called this morale. Some unsuspecting researchers have assumed that the existence of the concept indicated the existence of some behavior in nature and they set out to discover and describe it only to find that they had to redefine the concept so thoroughly that they had a different concept in the end for the only behavior which they could find in nature that remotely corresponded to the original concept--morale.

The failure of those who sought to identify morale empirically has not discouraged those who deal with people in organizations from using the term, however, for the term has functional validity for them. This is true simply because morale is one of the concepts in a common sense theory of organization and the theory provides a systematic means of handling the natural phenomena of organizations. Even though morale performs only the function of a trash-box in the corner under the workbench, it enables us to keep the rest of the common sense theory system intact and thus retain our present modicum of control over organizations. The essential nature of morale, i.e., communication, leadership, unity of command, etc., is not its empirical basis but rather its ability to impart meaning to those who impose them upon the complex phenomena of organizations. Regardless of its ambiguity it has been used to develop patterns of behavior which are successful in meeting problems in organizations. It is this feature of theory systems which probably indicates the heart of the controversy between the operationists and those who hold the hypothetico-deductive position.

Scientific theory systems serve two rather different functions. When the practitioner uses the system he imposes it upon the natural phenomenon which he sees during the process of perception and it is a "good" theory if it keys him to perceive pertinent functional aspects of the natural situation. When the scientist uses the theory system he demands that it correspond to all of the different processes that he observes and that it enable him to predict what will happen on the basis of the interaction of the factors which he is thus able to identify. The practitioner is not concerned with the prediction of an event directly. He is concerned with the solution of what he identifies as a problem. The scientist is not concerned with the selection of a particular solution from among a number of alternatives but only with the most precise possible description of the dynamic elements of the situation. This difference seems almost so subtle as to be unreal, yet it is very real to both scientists and practitioners involved in the study of a particular natural process.

A third point of issue in the controversy is the stress on quantification. One of the charges is that though we have an impressive number of statistical techniques for the analysis of quantified data we are often unable to interpret the statistical conclusions in any meaningful way. All statistical conclusions are limited by the degree of correspondence of their mathematical assumptions and the situation in which the data were gathered. In many cases it is impossible to apply the results of analysis because of the failure of the situation to satisfy any of the mathematical assumptions embodied in the statistical technique used on the data. This is true of almost everyone but the professional psychometrician. The only valid charge against the professional is that he selects only those situations for study which do satisfy the conditions of his formulae. Yule makes this strong accusation.

> Failing the possibility of you measuring that which you desire, the lust for measurement may, for example, merely result in you measuring

> something else--and perhaps forgetting the difference--or in your ignoring some things merely because they cannot be measured.[76]

The writer feels compelled to wholehearted agreement here although, as will be stated below, there is no alternative to the quantitative approach when we require precise answers to big questions. Precision demands taking account of total variance and this leads to the accumulation of vast quantities of data which simply cannot be handled meaningfully by the human brain. There is little justification in Yule's terming the demand for quantification which is inherent in the situation, the lust for measurement on the part of the statistician. To the writer, it is more fitting to accuse the statistician of denying the possibility of taking account of critical variance through the use of other techniques. In a complex situation there is little probability that the investigator can simply assess the operation of non-quantifiable variance by intuition. It is justifiable to utilize some of the subjective techniques such as interviewing, personal history, case history, etc., and then make a "best guess" on the basis of this data. It at least represents controlled observation. There are many studies which reflect a certain amount of

---

[76] G. U. Yule, "Critical Notice on William Brown and Dodgrey H. Thomson's The Essentials of Mental Measurement," British Journal of Psychology, 12:107, 1921-22.

preoccupation with precision and exactness where the central problem is the description of total variance. Many statisticians have the tendency to transpose these two on the assumption that they are interchangeable.

Bower makes this fundamental point in discussing the limitations of precision as a test of the value of scientific results.

> It is suggested that numeration *per se* . . . tends to be distortional because it leads to over- or under-determination of meaning. By over-determination we refer to the tendency to over-generalize from limited data. By under-determination of meaning we refer to the accepted narrowing of the horizon of application of specific data. For example, an I. Q. of 55 on the Wechsler-Bellview Adult Intelligence Scale might be misinterpreted in the direction of an under-determination of meaning by observing that this merely shows that the individual has an I. Q. of 55 on the Wechsler test. Or one might commit the fallacy of over-determination by saying that this score *proves* that the individual is feeble minded.[77]

Bower's implication is that a statistical result is easily taken as almost mechanically interpretable while actually we have to exercise the greatest care to get the maximum meaning from a statistical coefficient, and yet not over-generalize. To the writer it seems that one means of doing this would be to carefully set down the assumptions of the technique and the characteristics of the situation and then attempt to determine the degree to which they

[77] Brower, *op. cit.*, p. 326.

matched with as much assurance as possible.

Atomistic. A fifth charge against the operationists is that operationism is essentially atomistic, and therefore does not allow the possibility for the larger meaning inherent in a wholistic point of view. Churchman states this point of view eloquently.

> To be precise, then, we cannot talk of one thing being causally connected with another. One object or event is never by itself the cause or effect of another, and every object or event can be made an aspect of a time-flight which is causally related to another time-flight.[78]

Time-flight here refers to an event which can be assumed to have no time dimension. To get the full meaning from a theory system we must look at wholes as well as the pattern of the parts. Thus, we cannot understand the full meaning of the place of an organization in a community by studying the operation of all of its parts. We must study the place of the organization as a unit in the community. What are the attitudes and more especially, what are the stereotypes of the organization which people in the community hold? It is held that this type of meaning is distinct from any specific attitudes and stereotypes the public holds toward units or persons in the organization.

---

[78] C. W. Churchman and Russel L. Ackoff, "An Experimental Measure of Personality," Philosophy of Science, 14:304, 1947.

Bower points out that the heavy dependence upon quantification is one of the reasons for the atomistic nature of operationism.[79] Numeration is by its nature atomistic since it cannot produce results which are anything but the sum of partials. A median for example, is nothing more than the abstraction of a single sum for a given aspect of a particular set of data. Correlation is by nature atomistic. We have no way of correlating one whole with another. Linear correlation is nothing more than the abstraction of a single figure which we say represents the relationship between a set of individual events or persons.[80] Again, this is not to be interpreted as any indication that we must scrap the use of either the median or linear correlation, but as an indication that we must bolster our interpretation of them by consideration of the whole events or things from which they are abstracted. It would be fallacious to hold that a police unit was composed of persons who were above the median in experience, intelligence, emotional stability and morale was free of administrative problems unless we knew something about the attitude of the public toward this organization, and hence toward the individuals in it.

---

[79] Brower, _op. cit._, p. 326.

[80] Churchman and Ackoff, _op. cit._, p. 329.

Static. Finally, it is charged that operationism is static while we must assume a dynamic environment. Weber states that the motive of the operationist may well be to avoid the shocks to the theory system that result with the emergence of new constructs, new models or new levels of nature.[81] He feels that they are assuming that the relativity theory has eliminated the possibility of further emergences, but this is not true. By pointing out that all theory systems are completely relative to the purpose for which they are framed and the model on which they are built the relativity theory makes it clear that we may have an infinite number of theory systems and shows bluntly how purely pragmatic theory systems are in terms of any imponderable eternal nature we might think they contain.

In summary, operationism has been strongly advocated as a solution to many of the problems that we have encountered through the use of logical analysis. That it is a potent technique is amply shown by the strength of the defense it has aroused and the number of adherents it has gained. It is probably not now nor can it ever become an independent system of philosophy for the reasons amplified above. The principal reason being that it lacks power of allowing inference because of the elimination of broad

---

81 Weber, *op. cit.*, p. 57.

axioms which are a part of every logically derived theory system. On the other hand, it is the only available means of developing verbal symbols which have any of the ability of mathematical and topographical symbols for simple, precise, non-ambiguous representation of natural phenomena. The most telling criticism that has been made of it is that it needs to be supplemented, for the present at least, with subjective and logical techniques of various sorts to overcome the limitations inherent in quantification, atomistic analysis and the shallowness of the operations which we have available to us now. It is to be counted among the two or three most significant processes of scientific method in the social sciences and of equal importance with semeiotics and logic.

VIII. THE HYPOTHETICO-DEDUCTIVE POSITION

The fundamental tenet of those who oppose the operationist point of view is that we simply cannot have a science without the use of hypothetical constructs.

> It is the aim of all . . . science to rise above the historical stage and become theoretic, that is, to attain the form of a theory or system in which all propositions are logically or mathematically connected by laws or principles.[82]

Cohen's implication here is that if we use only

[82] Cohen, *op. cit.*, p. 347.

propositions which relate directly to observable behavior we have nothing but a descriptive theory system. This is in no way a criticism of descriptive theory systems as such but only an emphasis of their limitations. This is perhaps the most important point of contention between the operationists and those who hold to the hypothetico-deductive position. Cohen feels that the whole idea of the scientific method is based upon the assumption of a conceptual system which becomes increasingly powerful as an explanatory tool with the addition of more general and more penetrating intervening variables added through continued experimentation.[83] This is the sense in which the scientific method is circular for we start with a logically consistent theory system and by testing parts of it add propositions which are more general and further removed from that which is directly observable.

The assumption is that a descriptive theory system has great utility in the analysis of our environment but that there is an inherent deficiency in systems which are suilt strictly from sense data. The limitations of our senses preclude our seeing fundamental dimensions of nature.

The great importance of the scientific method is that we can start with pure sense data and infer the

---

[83] *Ibid.*, p. 107.

fundamental nature of natural phenomena by the collection of the specific types of data which we identify through the use of a conceptual system. It is only assumed that inference produces intervening variables which represent nature and the operationists are not willing to accept it for they feel that the price is too dear. The price, as they see it, is the indeterminate nature of the knowledge purchased with this assumption. No one has been able to develop a set of assumptions which cannot be attacked and which do not run the risk of being completely denied at some later date.

Need for different types of theory systems. In the writer's opinion an astronomer (interestingly enough) has offered the best basis for the resolution of this difference in point of view. Weizsacker holds that there are two qualitatively different types of knowledge in terms of function. He feels that most scientists are not aware of this and consequently spend a great deal of time criticizing one another's doctrines. His feeling is that science is confronted with two different types of problems which can only be met with two different types of theory systems. The one he labels the instrumental type of theory system. It is used for the analysis and solution of particular practical problems. It need not be all inclusive or completely unified but it must be highly reliable and reasonably precise. It must be highly descriptive in content for it is its direct relationship to immediately perceivable problems that gives it its value.

The second type of theory is not named but it is used for the analysis of broad problems having some magnitude and complexity in the sense that they involve a substantial number of interrelationships. It is of the nature of this type of problem that we are unable to see any order in it merely through the use of unaided sense perception. The assumption is that either the senses are too simple or the brain cannot handle the great complex of relationships which sense organs send it. Regardless of what produces this block, we all realize its existence. (Note that it is the lack of understanding of this problem which keeps most laymen and many scientists from admitting that we actually can tackle the great social problems of our time). To the writer, this assertion has an impressive amount of "face validity." The charge that "the appeal of operationism is due to its strategic value for avoiding <u>ontological questions</u>" is simply not justified.[84]

The operationists are attempting to fulfill a fundamental cultural need in our society which has traditionally

---

84 Weber, *op. cit.*, p. 55.

been disclaimed by university researchers as "below them." The worst that we can justly say of the operationist is that his great acclaim for filling this need has turned his head momentarily and he has overlooked the existence of these larger problems. Of what direct value is the culture concept to the teacher in a high school class room or the guard at a prison or the driver of a truck? It has no correspondence to the problems which these people face because it does not include interrelationships with which they must deal.

The operationists have come forward with theory systems which apply directly to these situations, because they embody propositions which refer to the relationships which people in practical situations can perceive. They contain propositions which actually refer to relationships at the descriptive level. This does not, however, eliminate the fact that descriptive theory systems, by definition, cannot provide insight into complex problems because of their very limited depth.

It must be noted that although there are two types of uses for these two classes of theory systems they are intimately related methodologically. Behavioristic theory in psychology, a theory system largely developed through operational methods, has gained much from the insights developed by the more general psychological theorists. In turn, the postulates of the theorists have been more precise because of the reliability of the behavioristic concepts. It is seldom that there is one and only one operation for each concept the operationist creates and the operationists have certainly received aid in the choice of operations from the theory systems of the general psychologists.

In addition to requiring two types of theory systems we require not one but several complete theory systems of each type. There may come a time when we shall have constructs of such depth that we can have one complete theory system for all of nature. There is considerable agreement that our present knowledge of nature justifies our having as few as five broad systems of theories. These are customarily referred to as the five "levels" of nature. Larrabie has described them and although there is not complete agreement with the finer aspects of his classification, it is substantially the same as others which the writer has seen.

The lowest level is the sub-atomic physical. Here we deal with isolated electrons and protons and the individual behavior of sub-atomic particles of all kinds. We attempt to determine patterns from which we can ultimately predict the behavior of individual particles. The second level is the gross physical. This is the familiar realm of physical things forming the zone of maximum predictability because we can use the mechanical pattern of the ninety-two chemical elements and the laws of physics to describe behavior at this level. The third is the biological level. This is the level of the organic as opposed to the inorganic. We study the physical aspects of all organisms on the earth and attempt to develop postulates which give us control of their behavior. The fourth level is the psychological. This includes all of the theories about learning, adaptation, drives, etc. That is, all the theory systems related to individual behavior. The fifth and highest level of theory is the socio-cultural. This includes the primary groups, formal organization, institutions and finally nations.[85]

Now these are only postulated levels of nature. We do not know enough about nature to know whether it is continuous or not and we make this arbitrary classification only because we have been able to develop theory systems which can include each of these levels. We have found that in fact these levels are not completely separate at all.[86] Relationships exist between the theory systems that we have developed but the theories cannot be combined yet because

---

[85] Harold A. Larrabee, Reliable Knowledge (New York: Houghton Mifflin Co., 1945), pp. 304-305.

[86] Ibid., p. 305.

we lack the assumptions with which to bring them together.

The relationship between these levels of theory systems is a remarkable one. Each of the lower theory systems serves as a frame of reference for the analysis of partials at the next higher level. Experience shows that these frames of reference are sufficient for the explanation of the behavior of particles at the next higher levels, but they do not describe the action of groups of particles. It is for this reason that we say the cultural level is the highest level of theory. It is built upon the other four.[87]

Thus, we can tell a great deal about any given atom in the gross physical level by the analysis of its subatomic parts, but we cannot predict the behavior of the atom until we know the environment in which this atom is active. Likewise, we can tell a good deal about the typical man with a given intelligence quotient, vocational interests, activity preferences, etc., but we can tell very little about what such a man will actually do unless we know the environment in which this man is living and we do not find this type of knowledge in psychological theory, in sociological and anthropological theories of culture. This is one of the reasons why the future of social science hinges upon cross-discipline cooperation.

---

[87] Ibid., p. 305.

Need for instrumental systems. Supplementing these five fundamental theory systems we require innumerable instrumental systems. Although Weizsacker does not make this point it appears to the writer that we shall need instrumental systems of varying degrees of complexity. These theory systems should correspond to the complexity of the problems they are developed for. There are some single problems of such great scope that they require extensive theory systems of their own. The release of atomic energy might be such an example. Then there are problems which do require a theory system of their own, but a relatively simple one. The theories concerning a given class of insects or plants might be an example of this type. Still other problems can be handled satisfactorily under a common theory system.

These types of theory systems are perhaps satisfactory for within-discipline analysis but they seem to leave the problem of scientific development in the so-called "applied sciences" unsolved. The pertinent characteristic which everyone of these applied fields manifests, and one which remains unsolved, is that cross-discipline theory systems are required to solve the problems in these fields. In the social sciences we have social work, public administration, economics, education. In the physical sciences there are all of the fields of engineering, architecture, biochemistry and others. These fields are not disciplines which have developed because they embody a theory system corresponding to a level or a part of a level of nature. They are fields which have been developed to meet some total social problem. In most cases this problem is a total problem in the sense that it requires the theory systems of two or more whole levels of nature to even begin to make analysis providing real insight.

So far we have been concerned principally with description in the applied fields in the social sciences and so there has been little realization of this problem. We have merely taken the concepts that appeared to indicate behavior which was pertinent in one of the applied fields and used them to make our classifications. Now, at least in social work and public administration, we are thinking of scientific investigations. The very natural solution is to turn to rigorous operationism for we do not have to face the problems of the nature of nature and the making of assumptions. It is undeniable that there is such a great need for systematic analysis in these fields that there is much that can be done merely through the use of these techniques. It seems improbable that we shall ever be able to close the well-known "culture lag" in these fields by the use of the pure operationism when the physical sciences are utilizing the more powerful conceptual approach.

The problem is, then how can we carry out scientific investigations using the hypothetico-deductive approach in these peripheral areas involving two or more of the primary levels, and therefore their assumptions, without the necessity of developing a whole new theory system? We cannot use the theory system of but one of these areas, and its assumptions, without limiting our conclusions to that area. We cannot arbitrarily combine the theory systems of two or more levels, for the philosophers of science maintain that there are "natural breaks" between them which prevent the propositions of one from holding in the other. In public administration there is no doubt but what we must consider at least two levels, the psychological and the socio-cultural. Certainly the biological is at least of some importance. It is literally true that those who maintain that public administration includes all of social science are right.

The writer was unable to find anyone who sensed this problem and recorded his views in the literature. He feels it is a real problem springing naturally from the best thinking of those who have developed the hypothetico-deductive method. It is obvious that the lack of any discussion of the problem might indicate that those who know most about the situation do not consider it a real problem. The writer takes the view that the authorities on this subject have not yet concerned themselves with the applied fields and so have not located it yet.

Common unit for social science. Two solutions seem tenable. The first is that we shall perhaps find a common unit for all social science and the psychological-socio-cultural barrier will be removed. There is nothing in the history of science to indicate that the probability of a unifying concept will be found instead of our discovering another streak in nature. For example, there were those who felt that the periodic table and the concept of the atom would once and for all unite all of physical science and then the sub-atomic particles were postulated and proved, and physical science was split in two. But if we were to assume that the barrier would some day be removed we could arbitrarily combine social theory systems and then revise our theory systems slightly when this occurred.

Development of new theory system. The second possible course of action is to look upon the streak concept merely as a limiting factor, which it is, and then attempt to develop a new point of view of that part of nature which is involved in the total problem of administration. This would be followed by setting down a set of assumptions of our own, postulating a theory system and then setting about to test it. To do this would be new in the social sciences except for social psychology but it has been done in several areas of the physical sciences. An outstanding example is the now almost distinct field of aerodynamics. Meteorology is another example.

One of the requirements for any theory system is a model. The characteristics and the functions of models were discussed in a preceding section. Our new assumptions about the segment of nature involved in the processes of administration will largely determine the model which we select. Next we must fill our model in with concepts or propositions, for a theory system is simply a group of concepts which are related to each other through a model. Concepts are fundamentally abstractions of common characteristics of phenomena in that area of nature which we set down for our study. Knowing the model and the concepts we can apply the theory system to a specific problem by mere inference. We need not have the relationships of each concept stated fully for the theory system is constructed through a set of rules and we derive them as needed by establishing the implication. Larrabie says that implications are the structural relationships between propositions.[88]

<u>Scientific foundations for such a theory system</u>. The

---

[88] *Ibid*., p. 91.

rest of this paper is devoted to the exposition of a method for providing a scientific foundation for such a theory system.

Northrop holds that concepts have two distinct aspects. The first is the aesthetic component which denotes the relationship of the concept to the empirical event from which it was derived. The second is its theoretic component, which is the meaning component which is derived from its association with other concepts, by reason of its being in the theory system, through implication. (Note, that it is the aesthetic component which establishes the epistemic correlation).

It is this feature of the theory system which gives it its utility over the operationist approach. Having concepts with a high theoretic content we are able to look at natural phenomena and abstract not merely common aspects which can be detected through the senses, but through inference, by utilizing some of the rest of the theory system, we are able to infer common aspects and thereby classify phenomena into classes which we never would be able to detect without the aid of the theory system.

> . . . the questions which pure sense data answer are not rich enough to provide us with the presuppositions we must make in order to answer all questions within science.[89]

---

[89] Churchman, *op. cit.*, p. 124.

An example of this is the staff--line relationship. By observation we find that staff assistants have no formal authority, that they have no important subordinates to carry out their work for them and that the successful ones habitually act as people with little authority in their relations with their line counterparts. It is impossible for the observer unequipped with a descriptive theory system to contradict the line--staff principle for this principle was developed just to describe this type of behavior.

However, the observer equipped with a theory system which includes the concepts "power," "status," "prestige," "informal channels of communication," soon finds that he can perceive a very different set of relationships than does the person who is only equipped with the descriptive theory system. What is of more importance, he is able to explain much more successfully the line--staff relationship. He is even able to predict things which the person with the descriptive frame of reference feels are due to some external force. Note that the explanatory power of the theory system does not come from the aesthetic content of its concepts.

In his discussion of the line and staff concept, Simon holds that it is a myth. This, to the writer, is a false classification of the idea. It is a fine and a useful descriptive concept. There is nothing mythical about the differential behavior that we can observe in line and in staff personnel in large organizations. Simon calls for and attempts to establish a theory system which has greater power than the descriptive frame of reference which we have traditionally utilized in public administration.

Northrop notes that we shall never be able to "prove" the validity of concepts which have a theoretic saturation by recourse to the descriptive framework. Our only course is verification through inference by using concepts with a lower theoretic content.[90] This process he calls "establishing the epistemic correlation"--that is, establishing knowing in one way on the basis of knowing in another way.

As we noted above, the essence of the process of getting deeper meaning from a situation is to classify the elements of the situation in a more meaningful way. The risk which we run in doing this is that of faulty classification. Churchman makes the penetrating observation that there are generally four types of classifications which we can make. ~~That is,~~ there are four general bases which we can use in the classification of common aspects which we abstract from natural phenomena. These are mechanical, physical or structural, morphological and functional.[91] We

---

[90] Northrop, op. cit., p. 113.

[91] Churchman and Ackoff, op. cit., p. 308.

might say that a boat, an airplane and a bicycle are of the same class if we were using a functional basis for our classification for the function of all these is that they are used for transport. If we were to classify them physically or structurally we should find them of different classes. A morphological classification might place them in the same class again for they are all of the broad physical class of things which move independently in response to the piloting of man. Mechanically, we might class the boat and the airplane together for the basic mechanical principle involved in the support of both is displacement. We might just as well place them in different mechanical classes for the explanation of the movement of an airplane involves different principles than would hold for a boat. The point is that classification, unhappily for the scientist, is still largely an arbitrary affair.

This is the great risk which we must cope with in the use of the hypothetico-deductive approach, and it is only heightened in the case of public administration where we are dealing with social behavior, for the problems of classification are more subtle here. The purpose in pointing out this limitation here is not to propose the rejection of the technique but to frankly admit one of the problems involved in it. Knowing that we can abstract different aspects of the situation we must take particular pains to abstract that aspect most pertinent to our problem.

The completed theory system has two characteristics which make it possible to establish the implications between concepts at will. The first is consistence. An operational definition of consistency is, ". . . it is not possible to derive contradictory theorems from the same set of postulates."[92] When we establish our theory system we define each of the concepts so that it is consistent with the rest of the postulates in the system. Secondly, the set of postulates must be complete. A complete set of concepts is merely one which accounts for all of the known or anticipated types of variance in a situation.[93] So much for the nature of the theory system.

Methodologically the power of the theory system lies in its role as a body of complex presuppositions for the testing of new propositions. We cannot possibly perceive all of the relationships involved when we are testing a proposition so we use the theory system as a means of making a low risk estimate of the relationships which are experimentally important.[94] The significance of this function is pointed out by Witzemann. He notes that by the

---

[92] Kattosoff, op. cit., p. 65.

[93] Ibid., p. 66.

[94] Churchman, op. cit., p. 124.

time of Galileo there was a great deal of well supported empirical knowledge which could have been used in testing broad theories but it was not possible to use this body of tested propositions for they were not related into a complete and consistent theory system.[95]

If everything is not actually related to everything else, every phenomenon which we seek to test is related to so many other things that we cannot handle the relationships systematically without a theory system. The importance of this knowledge is pointed up by the limited body of scientific knowledge that we were able to develop until the theory system began to be used as a body of presuppositions: it is these presuppositions which determine the pattern of facts which we shall collect and therefore the range of conclusions, from the infinite number of solutions or answers to our questions, which we shall select.[96]

It was not until the advent of the theory system that the use of only partially verified propositions was considered. The fragmental propositions previously used gave such limited control that they required complete verification, as do operational propositions. In a

---

[95] Edgar J. Witzemann, "The Scope, Objectives and Limitations of Modern Science as Seen in the Light of its History," Philosophy of Science, 14:52, 1947.

[96] Cohen and Nagel, op. cit., p. 200.

sophisticated theory system empirical verification of each of the propositions is impossible since many of them are intervening variables which the senses cannot pick up. Cohen notes that we need not verify all of the propositions, though. He says that if we make a careful check of all parts of the theory system that are concerned with observable phenomena we shall be able to hold that the whole theory system corresponds to reality, or is verified, since the higher concepts will be consistent with these we verified.[97]

Kattosoff notes that if we are able to develop hypotheses which produce facts which cannot be subsumed under the theory system we are forced to examine the whole system to determine whether the facts indicate a weakness in the axioms underlying the structure or whether they indicate that one or more of the concepts of the system is ineffective.[98] One of the means of finding adjustments in the theory system is to assume that one of the concepts in the system applies to more than it is said to and then test it for the new scope. If this procedure does not produce a negative conclusion we know that one of the concepts is faulty in scope and we proceed to examine the

---

97 Cohen, op. cit., p. 108.

98 Kattosoff, op. cit., p. 66.

scope of all of them.

By this time it should be apparent that a theory system has been pretty well described. Figure I contains a graphic representation of a theory system which will bring some of these independently described aspects into a total picture.[99] Note that the concepts at the various levels differ in scope as well as in aesthetic-theoretic content. Note also how the whole system is balanced upon the axioms and the major principles of the system. The several levels are labeled axioms, laws, major principles, principles, and concepts in purely arbitrary fashion to bring out the essential difference between the nature and the function of the different levels of abstraction.

It has been mentioned repeatedly that scientific knowledge as it is embodied in a theory system is contingent or indeterminate. Only in the descriptive area can we have anything approaching absolute certainty. Our knowledge is directly contingent first upon a rigorous method. Second, it is contingent upon the degree to which our theory system is verified. Finally, it is contingent upon the assumptions underlying the theory system. It is this third source of contingency that we must point up in more detail at this point.

---

[99] See p. 94.

Cohen holds that one of the primary characteristics of a theory system is that it is founded upon a set of assumptions.[100] One of our primary assumptions is the model for our theory system. If the model is only slightly faulty there may be little or no distortion in the application of the system for a theory system is relatively loose by virtue of its being constructed of verbal rather than mathematical units. A model which is too simple, too complex or too rigid in comparison with the actual nature of the phenomenon to which it refers will, as soon as the theory system is tightened up, show its weaknesses in the data which controvert it. A second set of assumptions concern the primary dimensions of the phenomena to which the system refers. For example, if one assumed that the primary dimensions of administrative behavior were biological and psychological, it is probable that one would soon find indications of the need for a third dimension, group behavior. If, as has been done in orthodox administrative theory, one considered only formal organizational behavior, one would probably soon find indications of the need to consider the social or informal aspects of administrative behavior.

Kattosoff makes this point. There is some

---

[100] Cohen, *op. cit.*, p. 151.

justification for making the fundamental assumptions that there is a regularity and order in the natural phenomena under study if we have continued success in the verification of theory systems which embody such assumptions.[101] Some extreme operationists have implied that these assumptions are not warranted. The writer sees no reason for not making these assumptions if we label them clearly as such. It is worthy of note that the use of these assumptions gives this approach the flexibility so characteristic of it and produces the self-corrective process which is felt to have such great significance.

It has been consistently maintained that one of the prominent features of the hypothetico-deductive approach is that it allows us to get deeper insight into broad natural phenomena than any other scientific scheme now known. Specifically it was said that this method produces deeper meaning. The question arises, what constitutes scientific meaning? This question brings us directly to the ontological problem; how do we know that what we have held to be true in our theory systems is true? Any attempt to meet the problem and provide a single answer is, by definition, doomed to failure for there is as yet no agreement among scientists as to what the criteria for knowing is.

---

[101] Kattosoff, *op. cit.*, p. 74.

The simplest definition of meaning is to say that it is explanation. Hall defines explanation as the development of a deductive system which corresponds to some natural event or situation.[102] This is a very simple, direct, and acceptable concept of meaning, for certainly when we have developed a verbalization which reproduces a natural system we have facilitated its explanation in the everyday sense of the term. If someone can tell you what is happening inside a radio, you will be satisfied that he has explained the operation of a radio, providing you can understand his description. This definition of meaning has a good deal of ambiguity, however, for it does not tell us just what sort of explanation is required; it sets up no standards of efficiency for theory systems. In short, it is not satisfactory here.

Feigl notes that there is no difference between explanation and description.[103] If we violate contexts and connect this with the above definition of meaning, we are led to the position that to impart meaning is to describe. This, certainly, is the position of the operationists. A complete and accurate description of behavior shows clearly

---

[102] Hall, op. cit., p. 157.

[103] Herbert Feigl, "Rejoinders and Second Thoughts," Psychological Review, 52:284, 1945.

the meaning of the behavior. This point of view points up an additional dimension of description. It is not a thing apart or independent from the scientific method or the experimental question. Description is only useful when related to some purpose and some method. In some cases we require simply the means for classification of behavior, as in the case of the law. Motive or cause is a minor concern.

It is apparent that the type of description sufficient in the law is entirely deficient when we are concerned with causes. We might further qualify our definition of meaning by holding that it involves a descriptive system sufficient for the purpose. When we merely wish to operate a radio, the most meager descriptive system will suffice. When we wish to repair one, a more complex system is demanded, and when we wish to design a radio for a new purpose, a most rigorous system is required.

Another aspect of this point of view is that we explain when we show the relationship between a natural event and a theory system.[104] The point is that when we do this we establish the relationship between this event and all other events represented in any way in the theory system and we thereby provide insight into the probable behavior of the objects involved. Certainly this is what

---

104 Larrabee, *op. cit.*, p. 205.

we do in repairing a radio. We know from our theory system that the parts have rigidly established relationships and when the radio fails to work we know that these relationships have been disturbed and we examine each of the parts, checking to see that it has the proper set of relationships. When we find one which does not, we "fix" it so that it does. We do substantially the same thing in "fixing" a sick organization. Our administrative theory postulates a pattern of relationships and we check to see if these exist. If they do we expect to find a healthy organization. Our administrative theory is as yet incomplete and we often fail to explain malfunctioning as easily as does the radio repairman. This concept of meaning, that it involves establishing the relationships between phenomena, especially causal relationships, is most acceptable to the writer, but it, like all the others, is superficial in one important respect.

Only Churchman of all the philosophers of science points out directly that, "The answering of any question of fact or law presupposes a criterion of value."[105] In relation to the problem of meaning his point is just this. There is not one "meaning" inherent in any situation. There is not even one type of thing upon which we can impose

---

105 Churchman, _op. cit._, p. 210.

the term meaning. There are different types and kinds of meaning--as many different types and kinds of meaning as there are different values or value systems. While we can have a scientific method which is completely independent of a value system, we have no goals for our science until we accept some value system. It is this factor which accounts for the situation where two scientists using an equally rigorous method examine the same situation and arrive at different sets of conclusions. The operationists, alone, are able to say that their theory system is not dependent, ultimately, upon a value judgment. They also say that their concepts are independent of any value judgment. The writer feels that this is foolish. There is not, and in any mature culture never can be, one and only one operation which could be used to represent a particular segment of nature. The operationists would be able to eliminate some of the "looseness" from their systems if they too would state the assumptions, ultimately the values, upon which they base their selection between alternative operations.

In the hypothetico-deductive method the place of values is almost directly admitted in the statement of assumptions. It is never merely a case of deriving a single set of assumptions for the theory system. There is always a difficult choice between alternatives, each of which offers certain real "truth value." That is, to the investigator making the selection there is not one of the assumptions which is true and a set of others which are false on the face of it. The selection of the assumptions involves a difficult and fundamental methodological compromise. The set of assumptions which the investigator emerges with are a consistent structure of beliefs which the investigator feels are most true. Upon what basis does the investigator select between these alternatives then? Supposedly upon the basis of their internal consistency, their comprehensiveness, etc. But even after elimination upon these criteria there remain a large number of alternatives. Northrop says that whether he admits it or not the investigator falls back upon his own value system or rationally utilizes the value system which he feels should apply in the situation.[106]

He goes on to note that there are in reality then, two types of theory systems. Scientific theory systems, which provide insight into behavior in natural situations and normative, or value systems, which embody the values of the culture. It is, he says, imperative that we control the operation of value judgments in science as much as possible and the manner in which we do this is to record the values upon which we make our selections between

---

106 Northrop, _op. cit._, pp. 255-259.

alternative assumptions. To the writer this was a traumatic shock.

The implication for those who would do scientific research in public administration is that they must preface their research by a study of normative values (ethics, especially political ethics or political theory). <u>Political theory becomes the foundation of research in public administration</u>, since it embodies the value systems which are the only means we have of attempting a rational determination of the substantive goals of science in this area. Failing this, the constant selections which must be made in the application of the scientific method will be made primarily on the basis of the investigator's mode at the moment and will therefore be uncontrolled. This type of operation is technically termed "bias." While it is recognized and fought against by those who are experts on techniques, the writer submits that there is a tremendous entry of error into all social science through this avenue which most social scientists, on the basis of the literature, are not even aware of.

CHAPTER IV

THE EXPERIMENTAL RATIONALE

We have repeatedly emphasized the need for sound theory systems as the basis for the control of administrative problems. The ramifications of sense data are no longer apparent to the unsophisticated observer. We must have knowledge of what our plans will require when they are executed if we are to continue to increase the size of organizations and yet control them. In this chapter we are concerned with the total picture of the methodology which has the capacity to produce theory systems with both scope and reliability. The core of this discussion is the relationship between theory and methodology.

I. EXPERIMENTAL REQUIREMENTS

The problem has several dimensions. We require a theory system which approaches some sort of an answer to the problem of reality for the researcher. We need a theory system which represents as many of the important variables in administration as possible for the operating official. We should have a theory system which embodies strong epistemic correlations for the benefit of both these groups. Most laymen feel that we should have a theory system which is a full reproduction of the reality which they naturally perceive whenever they find themselves in an administrative situation. This is a naive view having no relation to the utility of the theory system. A theory system gains utility when it contains variables which will explain as many as possible of the variables in the situation which the system concerns. The explanation of variance depends upon the use of predictive factors, not upon the degree to which factors represent all of the emotional color which an individual may respond to in a situation. The reason that a theory system need not represent all the "local color" in any social situation is that apparently when people are living in an organization they soon learn to disregard all but the pertinent details of their work as an energy conserving measure. While the folk belief is that people do react to the hundreds of subtle aspects of an administrative situation, the fact is that they do not seem to do this in practice and so we do not need to represent all of these aspects in our theory system. This is merely another manifestation of the well-known disparity between what people say they do and what they actually do.

The practical implication of this circumstance is that laymen are not in possession of the technical knowledge with which to evaluate a theory system, except as they are applying it and testing its theoretic aspect. Our theory systems must embody the technical knowledge of the experimentalists if they are to have real value.

The real problem in constructing a theory system is the pattern of investigation which is to be used. There are two general methods. One approach is to carefully set down a complete theory system for the existing situation on the basis of a review of the literature and actual experience in the problem area. A whole sequence of small experiments is then executed to test this whole theory system. Actually every concept is not tested, but all of the key concepts are tested and using the internal consistency principle the rest of the concepts are evaluated. The other method is to make small hypotheses which are related to specific problems and test these. This is followed by revision and further testing. After hundreds of the small studies are completed, a theory system begins to evolve from the pattern of the hypotheses which have been substantiated and those which have been rejected.

At first glance these methods appear to represent two contradictory extremes. Actually this is not the case. The first scheme amounts to establishing a model and postulating the factors for a theory system and then testing the whole structure. This is impossible if we have no prior information to found our conceptulizations upon. The scientific method requires the control of extraneous variables through the use of a set of assumptions. The rough control which we should have in carrying out a series of smaller studies would be so loose as to give us only indeterminate results. We must conclude that in an uncharted area like administration this approach could not be used.

The alternative is to make small studies from experimental questions selected without a systematic guide such as the theory system. One soon finds that this too is impossible. In any area as complex as administration, the number of concepts and their interrelationships will be so great that the probability of our stumbling onto them by chance approaches zero. We must have some guiding principle if we are to select exploratory studies which will have value as a means of constructing a comprehensive theory system. This principle is the postulated theory system. Our conclusion must be that these two general methods must be used in conjunction in the initiation of the study of a broad cultural area such as administration.

The question is then, how are these methods combined into a scientific method? This may be answered by locating what we have to do in the initiation of our inquiry.
(1) We have to evaluate our model for the theory system.
(2) We must determine what data are available to us.
(3) We are in need of information which indicates the type of measurement devices that will be required. (4) We

have to reexamine the pertinence of our theory system in the light of the revised view of the experimental problem which we get as a result of the investigations which initiate the inquiry.

As a practical matter it is customary to initiate any inquiry with a review of the literature on the problem or remotely related to it. After some sort of manipulation of this material, a pilot study usually develops and this brings readjustments and a full scale study may finally be launched. What we propose here is a well-defined three phase experimental design which includes a detailed statement of the procedures which should be carried out in each step and the anticipated results of each procedure. The first phase is termed the "initiation of inquiry" and concerns the preliminary steps which are so important in any carefully conducted study. The second phase is the "exploratory" study phase and involves the charting of the experimental area through the location of promising variables. The final phase is the "measurement" process where precise definition is given our variables and their covariance carefully laid down. It is apparent that these phases are capable of including almost any scientific technique and that they have the effect of providing a classification which forms a continuum from the highly subjective techniques used in the initiation of inquiry to the wonderfully precise measurement devices which social science is now developing. This corresponds to the objectives of science which proceed from the location of new dimensions to the point where we have confidence in our dimensions but desire to know how they are manifest in different social contexts. There is a fundamental differentiation of function between the various scientific techniques.

In the initiation of inquiry we search with an open mind for any factor which may be in any way related to our experimental problem. There is no attempt here to develop variables which "fit" with the rest of our theory system. We are happy with any concept which cues us to new perceptions or which represent our new perceptions. There is a deliberate removal of any experimental controls here. At the other end of the continuum, in the measurement phase, we have made our selection of variables and we are careful to control every possible source of variance so that our measurement will be a true representation of the variance or covariance which we are studying. In the exploratory phase we are primarily concerned with the preliminary definition of our variables and the selection between variables which have alternative functions.

The problems of the initiation of the inquiry have been covered in the preceding chapter and will be treated in detail in Chapter VII. In the next two chapters we shall survey the problems of the exploratory and measurement phases. In the remainder of this chapter we shall be concerned with the interrelationships of these phases of the scientific method.

The techniques which are discussed below are taken from the three basic social sciences. The investigator who is interested in research in public administration must turn to other fields to secure techniques which will provide him with interpretative knowledge. The research methods of public administration and political science provide only descriptive knowledge. In the analysis of the mechanical dimension of administration, the division of labor, the development of effective procedures, the installation of internal controls, etc., these methods are sufficient, for the mechanical dimension of administration is inherently so simple a phenomenon that descriptive knowledge gives us dependable control of these processes. In the analysis of the social dimension of administration we require concepts which take account of more fundamental variance for this is such a complex area of behavior that much that is directly observable is meaningless. Anyone who stops to observe the behavior of just one work team will find it difficult to see anything very meaningful in this great mass of minute bits of behavior. To make any generalizations about the behavior of work teams we must refer to the causes of the specific acts that we observe. Analysis of causes requires the use of controlled observation so that we have some assurance that our laboriously derived conclusions have universality. The descriptive methods of public administration have none of these controls either on the samples of observation which are taken or on the bias of the investigator.

The techniques included in this chapter are not considered in the light of their place in social science in general, but as they apply to research in public administration. Thus, testing is given an important place among the methods listed, but it is not given a place of anywhere near the importance it now holds in social research. Testing, the use of paper and pencil or performance tests, is probably the most important of all social science methods today. It cannot play much of a role in public administration research until we have more knowledge of the probable structure of our major social areas and so the case study technique is given a place of equal importance with testing procedures.

In general, techniques which many sophisticated social scientists would never even consider are given a place here because public administration is on the frontier of science and these methods, crude by the standards of the sociometrician or the psychometrician, are appropriate to the exploration and mapping of this area.

The industrial psychologist or the industrial sociologist may say that he has not found recourse to these crude methods necessary to the development of a system of concepts about administration in the industrial scene. The writer feels that neither the industrial psychologist nor the industrial sociologist has developed anything approaching a comprehensive interpretative theory system. In most cases all that has been done has been the adaptation of a frame of reference of the investigator's major field to the industrial scene. While we shall certainly take many of the concepts from these fields, the writer feels that real progress hangs upon our developing a theory system fully appropriate to this important area of our culture. The concepts of the anthropologist are oversimplified, the concepts of the psychologist do not reflect the group dimension of administrative behavior and the concepts of sociology do not reflect the importance of the individual's role in the administrative process. While we shall certainly find that many of the concepts of the basic social sciences have real meaning as tools for the analysis of the administrative situation, we require a model unique to this social process supplemented by new concepts as necessary. The crude techniques included in this discussion are felt to be justified if one rejected the necessity for an independent theory system.

If we are to develop an independent theory system for public administration we shall have to face several problems underlying any theory development. First, we shall have to establish a model for our theory system and then we shall have to hypotheticate a set of concepts to fill the structure. We shall have to develop a set of assumptions which can be used as the basis for statistical analysis of data obtained through our application of these methods and finally we shall have to develop criteria of validity for our concepts. Meeting the large problems successfully will take a great expenditure of time and resources. Small exploratory studies can be organized and executed on small budgets. If an attempt is made at these more important problems in each exploratory study, we should soon be ready for a "focal study," one which will see the emergence of some fundamental answers.

II. BACKGROUND OF RESEARCH

Each research technique implies a theory system and a set of assumptions. Research results will have local pragmatic meaning without a theory system, but maximum research efficiency in the application of a given research technique is gained only when we apply it in connection with a theory system. In a field such as public administration where there is no empirical theory system, we may have only a projected theory system, but this gives the results of the experiment some larger meanings even if the dependability of the results is low. In psychology, where the theory system has been modified and substantiated by thousands of specific research efforts, the theory system is now quite dependable. In any case, whether we have a mature theory system or must rely upon a more tentative set of concepts, we cannot give our research results any but the most provincial interpretation unless we have a theory system behind them.

Theory system provides unique interpretation of research results. The most subtle function of the theory system is to provide a unique interpretation of research results. As was pointed out in Chapter III there is no fact which has meaning independent of some system. It was noted then that the fundamental reason for this lies in the culture pattern which we hold of looking at the total structure as well as at the integral unit of the whole. Of course there is nothing very unique about this and there is a very obvious reason for it: simply that in a mature, complex culture such as ours, we have found that almost everything is functionally related to almost everything else, and therefore if we are not to throw a monkey wrench into the gears of our cultural machine we must look at each cultural pattern both as an independent system and as one of the smaller units in the total culture process. Now in any given case we cannot stand back and look at these larger cultural phenomena all at once if for no other reason than they cover vast geographic areas. A second reason is the terrible complexity of these larger cultural phenomena.

We try to circumvent this by setting up a theory system where we represent the cultural mechanism verbally. The problem arises in the representation of the cultural segments, for there is no single agreed upon theory system which represents the whole organism. We have sometimes had several theory systems, each purporting to represent one of the major units of our culture. This is the reason that it is difficult for the Department of Labor, the National Association of Manufacturers and the labor unions themselves to provide three interpretations for exactly the same set of statistics provided by the Bureau of Labor Statistics. Each of these organizations conceives of that part of the cultural process which they are interested in, in different terms; yet, they all maintain they are talking of the same cultural process. Having a different concept of the whole, each of them has a different concept of the relationship of the parts to the whole, and more specifically, each of them makes a different value judgment concerning the efficacy of the parts. The Department of Labor may see a given set of statistics as indicating a desirable trend and one which is new. The National Association of Manufacturers may look at the same statistics and see only the reoccurrence of an old pattern and one which is dangerous to the culture.

Many experimentalists are critical of the classical Western Electric study because it lacked a satisfactory theory system. In one of the experiments the intensity of the lighting was first increased and then decreased beyond the range required for effective work, and yet production increased. Some take this to mean that the only important dimension of the work situation is the social. Others interpret this as meaning that lighting is very important. Still others deny that it means anything because there were such rough experimental controls on the situation that it would be difficult to say that the lights were associated with the increase in production in any way.

The interpretation which one makes depends upon the theory system he imposes on the result. The results each of these groups dealt with were actually the same set of numbers presented graphically. They presented a constant stimulus to each of the three groups of people involved. There is no question but what the intensity of the light was increased and decreased certain measurable units of intensity and there is no question but what production increased dramatically. No one disagrees with this, but this does not answer the question which is in the minds of everyone who attempts to interpret these data, namely, "What do these results indicate for a government agency, in a can factory, in a taxi company?" The experiment itself has shown what happened when the intensity of the light was increased in the relay room of the Western Electric factory some time in 1926, but the great scientific progress which we have been able to make since the turn of the century has resulted from the fact that in our experimentation we have almost always attempted to get reliable generalizations. By doing this, if we do it, we are able to make the results of an experiment in the Western Electric factory in 1926 give us some insight into the operation of the San Bernardino Ordinance Depot in 1951.

We do this by generalizing from the subject of the investigation to all of those organizations which can be held to fall into the same class as the organization studied. The critical step in generalization, then, is the classification of those organizations which will be held similar to the one where the experiment was carried out and the way we do this is by using our theory system.

There are immediate differences which anyone would perceive between the Western Electric factory and the San Bernardino Ordinance Depot. The critical question is whether or not the two situations are similar on the characteristics which are implicit in the conclusions developed in the Western Electric setting. They were not concerned with whether there was complete correspondence between the two organizations, but with correspondence on the variables related to this type of variance. The criterion of similarity was whether or not we could understand one organization on the basis of a generalization derived in another.

Many social scientists were critical of this procedure for they realized there was a trap here. There are hundreds of thousands of propositions which could have been made about the San Bernardino Ordinance Depot and, undoubtedly, purely by chance some of these would have been found applicable.

Indicates data required. It is a function of the theory system to cue us to the pertinent aspects of two situations which must be matched to allow generalization from one about the other. In all natural phenomena we find only approximate correspondence, never exact congruity. First, then, our theory system must tell us which situations might be matched. Second, it must give us some standard by means of which we can judge whether a match is made or not. It should be pointed out that this is largely a mechanical process, completely distinct from the folk culture and often antagonistic to it.

For example, the writer once heard one of the top administrators in one of the largest city police departments in the world, arguing with a social scientist that I. Q. scores for Negroes and Caucasians could not be compared because of the obvious differences between what to him were two distinct types of people. Whenever we match two organizations or two people or two situations, we must be aware of the fact that we are only matching one or two or millions of possible dimensions of these two situations. Many laymen expect exact similarity when the researcher says we may generalize. The point is that the scientist has determined that these are not differences which account for any of the variance in the phenomenon which he is investigating. This brings out the second function of the theory system.

Provides complex controls. The theory system should provide us with an objective set of concepts for looking at cultural phenomena which have a degree of complexity approaching that of the behavior in which we are interested. If we are ever to create a body of reliable knowledge, we must supplant the emotional or folk knowledge, what is here called common-sense knowledge, with a body of knowledge of great complexity. One of the reasons that we are unable to understand much of the behavior which we are now attempting to analyze is simply that this behavior is tremendously complex and we are attempting to use over-simplified stereotypes to treat processes which are very intricate and very sensitive. The situation has something akin to measuring the pressure in a steam boiler by tapping the safety-valve. In the cultural realm the result of using such crude instruments is that social situations keep exploding in our faces because we do not have the precision which is required to understand and control them.

Our folk knowledge was developed for and it is appropriate to the adjustment of the individual to his society and this is attested to by the great majority of people who make successful adjustments using it. It was not designed for and it has never been appropriate to the analysis of social situations involving great numbers of individuals. Psychology became effective only after introspection was dropped as a technique of scientific investigation. Psychologists found out the hard way that reliable knowledge is possible only after we take the investigator's ego out of the experimental situation.

Reliable knowledge in public administration and political science waits upon this very same shift.

Experiential knowledge is reliable only by accident; the accident that the individual stumbled upon the reliable concept by chance. It is as if the investigator were throwing dice with concepts on all six sides. If by chance he throws that one combination in thousands that is reliable, or represents reliable knowledge, he is able to add that bit of knowledge to the culture. The method of science, on the other hand, is to attempt to approximate reliable knowledge on the first throw, and then through a process of successive approximations, prove the dependability of the knowledge through careful control. Once in a while using the scientific method we make a mistake on the first throw of the dice, but the probability of our making a correct throw is hundreds of times better than that of the experiential investigator. One of the reasons for the high probability in favor of the scientist is that he has a theory system which approaches the complexity of the social phenomenon he is interested in and one which enables him to develop experimental questions which approximate the complexity of the true situation.

Theory system embodies assumptions. Experimentally the most important reason for a theory system behind every research technique, is that each methodological procedure embodies a set of assumptions. These are carried in the theory system. Actually there is substantial similarity between the assumptions involved in many procedures, so that we really get classes of techniques which embody nearly the same assumption. We have many of these sub-classes. The correct interpretation of the results of a given class of techniques requires that we have a theory system which embodies the same or nearly the same assumptions. One of the assumptions implicit in any research technique is the unit of behavior which is to be used in measuring the activity under study. The anthropologist is often concerned with the cultural pattern. This is the nature and the sequence of behavior related to a given cultural process, such as the production of foodstuffs. He is not directly interested in the whole individual or in the whole family or in the whole society, as such, when investigating a particular cultural pattern. He takes parts of each of these units of behavior as he needs them and as they are related to his pattern, and he builds a whole structure from these. When the anthropologist uses this technique, it is impossible to interpret his results in the light of psychological theory, since there is only an accidental collection of information about the individual here. The individual is not the unit implicit in the theory here. Since the unit of investigation used here was not the unit implicit in the psychological theory system, we are unable to generalize about individuals from anthropological data. For example, if the anthropologist shows that an individual in an Indian culture is aggressive in his relations with his father, we could not generalize from this and say that the individual will probably have an aggressive pattern in the other roles he plays. The reason being that the unit involved here does not reflect the same type of data as the data we get applying the standard psychological unit.

Interpretation. The theory system can also be used to tell us what portion of the total problem we have investigated when we complete a study. Our exploratory and measurement techniques are designed to measure limited variances. We need to know what variances are still to be studied.

For example, the well-known sociometric technique of sociograming developed by Moreno tells us only about structure of a group. It is not designed to tell us about the goals of the group, characteristics of the individuals of the group, or anything else. We cannot fully interpret a sociogram nor get its full meaning nor understand the limitations of its meaning, unless we have some theory system about the nature of groups. In this connection it is interesting to note that L. L. Thurstone, one of the most meticulous of social scientists, published a book,

<u>The Vectors of the Mind</u>, setting forth his theory system about individual abilities, before he published his work on, <u>The Factor Analysis Technique for Determining the Dimensions of Individual Ability</u>.

The coefficients obtained from the application of the factor analysis technique are still interpreted within this theory system, only it has since been revised and modified in the light of the results obtained using factor analysis. It is this requirement that makes the adaptation of general social science research techniques to public administration problems a long, expensive undertaking. These techniques were developed for the analysis of specific classes of problems in one of the basic social sciences. Public administration does not present exactly the same problems as one finds in these sciences. It seems improbable that we will develop a theory of administration which merely combines elements of the theory systems in the three basic social sciences, anthropology, sociology and psychology, since the problems in administration are different from those in either of these sciences. Yet we wish to borrow their techniques in developing our theory system. There is no easy answer to the problem. We shall simply have to recognize that inaccuracies are going to be present in our results because of the incongruities between their methods and our theory systems and we shall have to develop research procedures which are adapted to our theory systems.

It is hoped that these requirements point up clearly the need for the development of a theory system in administration which will embody the assumptions implicit in the experimental method. The marriage between theory and method will be seen to be the foundation of scientific research, for the reciprocal relationships here are so subtle and so intimate that they cannot be divorced without destroying the reliability of our knowledge.

It is necessary that we lay down at least some broad idea of the model for the theory system which is anticipated with the research technique that will be discussed below. We must set forth the broad assumptions underlying it and give at least a general idea of the relationships which will be implicit in it.

This bond between theory and method makes necessary the exposition of the broad outlines of the model for the theory system which accompanies the method to be set down here. We must set down the assumptions for the model, also.

III. PROJECTED MODEL FOR THEORY SYSTEM

One of the functions of the assumptions of any theory system is to establish the uniformities in the complex social phenomenon under study so that we can represent multi-variant parts of the total situation as logically precise units with limited complexity. A person who is actually looking at social phenomena carefully will inevitably perceive hundreds of separate variations. If we are to deal with social phenomena we must restrict the number of variables with which we must deal to some manageable number by assuming the actual situation is made up of a small number of causal parts. They are our hypotheticated intervening variables. In effect, it is the assumptions that are used to select among the many characteristics of a social situation those we choose to call the pertinent ones. Following are the assumptions we lay down relative to a theory system for administration.

Assumptions. Our fundamental underlying assumption is that administration is essentially a group process which has important individual dimensions. It has the following more specific manifestations:

1. The source of energy which is the input into an organization is the internal physical and social drives of the individual members of the organization. (The implication is that organizations, like machines and biological organisms, require an input of energy to produce activity and the only energy which is put into an organization is individual energy. There is no such thing as group energy

or mechanical energy in an organization in terms of coordination, control, leadership and the other organizational processes. The only possible way we can conceive of mechanical electric energy being a part of the source of an organization's energy is the electrical energy which goes into communications by teletype, radio, or telephone, or television. As this is only a facilitative device it has no positive impact on the organization, only negative or limiting effects.)

2. The basic unit in organizational research should be a functional, or sometimes a productive unit; that is, the unit which accomplishes a distinct operation, whether it be social or productive. In most cases this will be a work group of one size or another, but in isolated instances it may be an individual, as for example, a lawyer in a small organization or an accountant in a small organization.

3. By definition, all organizations manifest the division of labor principle, either formally or informally. Any group which does not manifest a division of labor in some way is by definition a group in a sociological sense.

4. The pertinent social dimension of individual activity in organizations is the adjustment of the individual to organizational requirements in such a way (a) that he achieves his own life goals through channels which (b) combine the appropriate sources of behavior and

and channel this behavior so that organizational goals of the agency are achieved at the same time the individual needs are satisfied.

5. There is never complete identification of personal goals with organizational goals, so that we find every individual playing a dual role in the organization to one degree or another. He plays (a) one role in relation to the formal objectives and the division of labor and (b) another role in his relationship with the social organization and other social manifestations of his own goals.

6. So many of the goals which the individual in an organization holds cannot be achieved except through the cooperation of others, that almost every individual in a large organization is a member of one or more groups which are organized about the common goals of individuals. The stability of these groups, the basis of membership in these groups and the activities of these groups can be explained in terms of the individuals who make them up and the dynamics of goal compromise in establishing group objectives.

7. Individuals strive to fulfill goals. Goals represent the integration of physical drives and social drives of the individual with the taboos, status symbols and other rewards and penalties in the social situation in which the individual exists. Organizational adjustment could be measured by the degree to which the individual

was satisfying his personal goals through and in accordance with the reward and penalty system of the sub-culture in which he lives. It would be assumed that attitude, interests, values are examples of the intervening variables which mediate between the basic drives of an individual and his goals. It might be possible to study individual goals in an organization without immediately going into extensive research on these intervening variables since they are the proper subject for psychology and social psychology.

The study of organizations might almost be called the study, on one hand, of group adjustment, or the adaptation of individuals to groups, and on the other hand of the mechanical problems of breaking down and re-synthesizing organizational objectives. If one wishes to call this the study of informal and of formal organization, and mean this by it, the writer feels that it would be difficult to deny the existence of these two fundamental dimensions of organization.

<u>Model for the theory system</u>. The fundamental problem in organizational operation has these two dimensions. It is necessary first to elicit the expenditure of energy on the part of members of the organization and, secondly, it is necessary to channel the energy expended into the formal division of labor so that organizational objectives

are achieved. For this reason it is held that the model for an empirical theory system for public administration should represent the problems of maintaining the equilibrium between the objectives of the organization and the personal or social objectives of the people in the organization. In many cases, personal and organizational objectives will be identical or will supplement one another, but there are other cases where they will be in opposition. (Status and high production are an example of the first, and empire-building and coordination are an example of the latter.) The core of the theory system should point out the sources of energy, the classes of personal and social objectives that the people in organizations have, and finally, the procedures and activities by means of which we can maintain a balance between these two in the light of requirements to shift production goals, to slow down or increase production.

The equilibrium concept is used for two reasons: first, the social objectives which are to be kept in balance are in constant flux. Individual social needs may change without any action on the part of the organization at all from some influence on the individual from an outside source, such as the family, a church or recreation group. Individual goals will change because of maturation if for no other reason. Secondly, the outside pressures playing on the organization change the objectives of any organization regardless of actions of administrators. The demand for new products, depressions and inventions make it necessary to change organizational objectives merely for the survival of the organization. The external influences have little relation to the efficiency of the administration of the organization.

The whole concept is one of a living social organism. New goals develop. New cells are added to keep the organism alive and functioning. New cells are added to allow changes in functions. New skills are added in the face of new techniques. The essence of the whole concept is a dynamic process. This concept reflects a time dimension, change, that is one of the primary aspects of organizational operations.

The question is what, specifically, are the factors to be kept in equilibrium? We probably could not give any indication of the final factors that will appear in the theory system since they will come after careful analysis and hypothetication of the theory system. As illustrative examples, organizational objectives and standard production objectives are one constellation of factors on the formal side. Adaptation to new production goals or adaptation to new production techniques for the same goals could be another constellation and perfecting internal management processes, such as coordination, communication, leadership, internal controls, budgeting, selection, would be a third constellation. On the personal or social side, personal goals such as security, a living wage, pleasant working conditions, physically, and pleasant working conditions socially, identification with some larger and enduring objective, and the like, would form this constellation. It may well be that the method of integrating these goals or setting them in equilibrium will be through a set of intervening variables, such as status system, incentive, participation, consultation, two-way communication, grievance systems and the like. Note that in a sense, the real situation as we know it now probably corresponds rather well to this. The theory system contemplated here will be both instrumental and theoretical, having practical as well as research application.

We are not using these concepts to mollycoddle individuals, for if the culture is to endure we must place at least equal emphasis upon the organizational objectives as against the achievement of personal goals. Immediate emphasis should probably be on the development of intervening variables which are effective in mediating different patterns or combinations of these two sets of goals. But note that in this model, we put the emphasis upon the patterns of goals and not upon the intervening variables alone. Research would undoubtedly take into account the environment of the administration as it has its impact upon the development of organizational objectives, and the environment of the individuals who live in organizations, if it has an impact upon the development of their personal goals. Finally, the internal impact of the organization on its own goals and on the goals of the people who lived in it must be accounted for.

Continuum of method. The research techniques from the three basic social sciences which follow are appropriate to research in public administration in the sense that they have been designed to investigate subject matter which one confronts in public administration as well as in the science in which they were developed. They have been designed with a specific theory system in mind and insofar as the assumptions of an empirically based theory system for public administration may differ from the theory system for which they were developed, they are not completely appropriate to public administration. These research techniques are all methods by means of which we can develop conclusions which give us interpretative power. These techniques will provide us with generalizations about administrative behavior which identify either causal or partially causal behavior or behavior which is consistently associated with particular administrative techniques. They are radically different from the traditional descriptive research techniques of public administration and political science. It is the difference that was pointed up in Chapter I between scientifically reliable knowledge and common-sense descriptive knowledge.

The research methods which will be discussed by class, are classified into two categories: exploratory techniques and validation techniques. This classification is based upon the proposition that there is not one homogeneous function which we can call scientific investigation. Rather, in all scientific investigation there are three almost distinct but not unrelated phases in the development of concepts which will ultimately take their place as reliable knowledge in a theory system. The first of these functions is the initiation of the inquiry involving laying out the problem and its boundaries and the most significant characteristics of the area of behavior. In this phase of investigation we determine which relationships are to be considered external to the behavior which we are studying and which relationships we will investigate. Also, we collect a great deal of invaluable background material which can be used for framing the objectives of research in the second phase.

In the exploratory phase of research we attempt to identify as many of the variables and the relationships between variables in the area of behavior we have selected, as possible. It is during this phase of the research that we attempt to select those concepts and propositions which best represent the behavior involved. (The broad scientific function of this exploratory phase is to develop the scientific model for the theory system.) The general function of the exploratory phase is to fill in our theory system with concepts which represent all the major dimensions of our area of behavior and give us as many detailed indices as possible. The validation phase is the last step in the development of reliable knowledge but it is never complete because the precise measurement and the rigorous controls used lead us to see disjunctures in our theory system and we are forced to begin the second phase of research all over again with this new knowledge.

It is in the validation stage of research where we use the precise quantitative techniques. Validation is unknown at present except through these techniques and so, in effect, we have been unable to validate concepts which are not measureable through our present quantitative techniques.

In the social sciences, at least, we find all three phases of scientific investigation in progress at the same time. The only difference between the three basic social sciences is the extent to which one or the other of these phases is used. We might almost typify anthropology as being preoccupied with the first, sociology as placing heavy emphasis on the second and psychology as claiming the third. To the very limited extent that characterization is correct, the second phase will include sociological and anthropological methods and the second, psychological methods.

CHAPTER V

EXPLORATORY TECHNIQUES

The development of an empirically based theory system in public administration waits upon the development of interpretative knowledge through the application of exploratory techniques. These methods are no substitute for the speculative hypothesis of the creative thinker, but are designed to supplement them. They may either help the creative thinker reduce a terribly complex phenomenon, such as administration, to a small number of variables, or they may help him see variables or relationships which he would not ordinarily conceive because the steps in the method include looking at the phenomenon from points of view which he does not habitually take. It is not implied that these methods have been developed solely for the discovery and charting of new areas of behavior. Many of them may be used satisfactorily in two or more phases of scientific investigation and where a specific method has significant value in two areas, this will be noted.

This chapter will provide an inventory of techniques and their relationship to the two categories which were established. It will also indicate the general degree to which they are qualitative or quantitative, since one of the important characteristics of any method is whether it provides us with data which can be handled statistically without a complicated process of classification into categories which are, at best, somewhat arbitrary. Most of them do not allow rigorous interpretation in the light of the mathematical assumptions underlying statistical procedure. The qualitative--quantitative dichotomy is a convenient basis for ordering the techniques to be discussed in this category. The qualitative techniques are discussed first, since they usually precede the application of quantitative techniques, if only because qualitative techniques are uncommonly used, where quantitative techniques are available.

I. HISTORICAL METHOD

This is the only interpretative method among those traditionally utilized in public administration research. The historical method is based upon the assumption that a deep time slice is necessary if we are to identify basic variables. The unique justification for the historical method is the assumption that everything has an antecedent or a set of antecedents. If we are to understand the city manager movement at the moment, for example, we must have data about the development of the city manager movement covering a period of 10, 20 or even 50 years.

There is an important difference between the use made of the historical method by historians and by social scientists. The historian customarily provides a chronological record of growth and change in terms of events, while the sociologist, for instance, provides us with specific details and rather reliable data concerning the behavior of individuals or primary groups. His unit is infinitely smaller than that used by the historian, and therefore his data provides much more detail. In fact, it appears that the two methods are coming to be seen as independent. The historian's historical method is being called the descriptive historical method and its application by sociologists and other social scientists is being called the analytical, synthetic, or sometimes the culture history method. It is the analytical historical method which is of interest here.

Young points out that in many studies (which we are classifying as exploratory) fundamental dimensions of behavior will be overlooked unless we take account of the history of the group or individual involved.[1] In a study of a religious sect, the Molokans, in Los Angeles, about 1924, it was found that membership in this group was almost unchanging so that measurements of group stability were exceptionally high. She feels that it would have been

---

[1] Pauline C. Young and Calvin F. Schmid, Scientific Social Surveys and Research (New York: Prentice-Hall, 1946). p. 210.

impossible to account for this abnormal coefficient of stability unless it was possible for one to find out that this sect was made up of people who were formerly members of the Russian Orthodox Church. They had been persecuted in Russia and forced out of the country by the Revolutionists in 1917. Seven years later, members of the group still clung together on the basis of this past experience of persecution. Young cites this as an illustration of the necessity for taking a deep time slice when engaging in a pioneering investigation, lest we lay down superficial concepts which are based merely upon an analysis of the situation at a given moment.

Variables which become the best predictors are those variables which represent fundamental dimensions of a situation. Therefore they are less susceptible to chance fluctuations than the more superficial variables. If we attempt to account for all of the superficial variance in a situation by making a prediction which utilizes superficial variables we are bound to have variables which are influenced by momentary variation and our predictions cannot help but be less satisfactory than they would be ordinarily if we had fundamental dimensions which were relatively unchanging. Young maintains that the significant prediction of future behavior often is impossible unless we have background information from a deep time slice, because only then do we locate fundamental variables.[2]

Typical of the sources of historical data are documents and other formal writings in the culture. Among these would be judicial writings, statutes, treaties, government records, records of proceedings or various government organizations, and the like. Another source of historical data is the verbal material of the culture, such as geneologies, chronologies, annals, biographies, public addresses, letters, the literature, poetry, newspapers, periodicals and radio and television scripts, etc. Also there are the personal sources of authentic observers or witnesses, such as personal letters, diaries, autobiographies, and others. Finally, we can study the artifacts of a culture, such as the money, utensils, objects of art, and recreational facilities.

Few social scientists regard historical data as any basis for the adequate explanation of social phenomena. This is simply not a function of historical data. Likewise, historical data lacks objectivity and it is not comprehensive enough to provide for prediction by the standards of the social scientist. It is appropriate to the location of social forces, such as the relationship between institutions in the culture, like that between church and state, or

---

[2] *Ibid.*, p. 212.

economic and political. For this reason it is probably true that the most effective use of historical data in social science research is to provide a description of the setting in which the particular behavior patterns under study take place. If we anticipate that factors in the environment of the behavior we are studying will be significantly represented in the prediction formula, this will be an essential technique to use during the exploratory phase of research.

The limitations on the historical method are obvious. The most grave limitation is our inability to control the data which we use and therefore the probability that it will be truly representative. A primary characteristic of scientific research is the collection of all the relative facts. This is impossible using the historical method. The second major disadvantage of the historical method is that we are almost unable to control bias. Using letters, diaries, records of conversations, proceedings and the like, we are at the mercy of the bias of the individual who provided the record we are using. One means of reducing that portion of unreliability in historical data which is due to bias in reporting, is to cross-check sources when we have two or more sources reporting the same event. But even then we do not have a control on bias, only a comparison of the bias of the reporters. In conclusion we can say that the historical method is the weakest of all of those methods which will be reported here.

II. ANTHROPOLOGICAL FIELD TECHNIQUES

There are a set of research techniques, most of them developed by anthropologists, which naturally group themselves under this heading. Some of them will be discussed elsewhere, where they fit more appropriately into the organization of materials. The anthropologists have developed a whole system of research techniques for describing and analyzing primitive cultures. Since the people in most primitive cultures are not literate, few of these techniques require subjects to respond in writing. Information is gathered through talking with people or merely by observing them. The cultural anthropologist may make a linguistic analysis and an analysis of artifacts. He may make a map of the village or geographical area which contains the cultural unit he is investigating, or he may analyze the whole physical setting, including the art, the architecture and the like. On the other hand, he may, as a participant observer, live with a family or some work group and analyze the personalities of the individuals he contacts in this setting. He may also analyze personalities by using the case history, by extensive interviewing, by the use of informants, or merely by observing people.

There are studies where some of our projective tests have been used in an attempt to analyze the personality of primitive people.

As a result of applying these research techniques, the anthropologist arrives at a complete description of the primitive culture. He will lay out the social structure, the belief systems of people in the culture, the operation of the major institutions, such as the church, the family, economic institutions and the like, and the dependence of this culture on previous cultures and surrounding societies. It must be said in favor of the anthropologist that by living with the people that he is studying, by learning their culture, by eating the food they eat and partaking in the recreations that they enjoy, he gets a "feel" of the situation which it is admittedly impossible to gain through any of the more precise techniques which we will discuss below.

In modern society, where as many as 98% of the people may be literate and where the culture is often completely dependent upon verbal symbols, many of the anthropologist's techniques are just too simple to accurately represent the social milieu. For one thing, primitive societies which are non-literate, gain only those culture traits from preceding generations which can be remembered and transmitted by the elders in the tribe. In literate society, it is possible to preserve and transmit to each new generation useful culture traits developed at the time the technique of recording was created. We, for example, use a system of numbers over two centuries old. Most anthropological methods take a time slice only as thick as the amount of time the investigator spends studying the culture. For this reason, most of the techniques listed above are not directly applicable to the analysis of behavior in our culture.

In recent years a small group of anthropologists have begun to develop a specialized field known as applied anthropology. These men are interested in the analysis of modern culture--namely, our own and Great Britain's, and especially the ethnic and industrial problems which are present. Perhaps the most widely publicized attempts in this area have been in the Office of Indian Affairs in the Federal Government. Impressive results were also obtained in the administration of internment camps by the War Relocation Authority. In these two agencies anthropologists were able to develop running analyses of the people being administered, with special attention to those aspects of the situation related to the decisions which the agencies must make. Some of the techniques which have been successfully used by the applied anthropologists will be discussed below.

The participant observer. This is a method where the investigator attempts to make himself a part of the situation which he is studying. The general pattern of investigation of the participant observer is to "work himself into the situation" very carefully and to spend his time when he is actually in the situation asking questions about a predetermined set of topics, and then when he is out of the situation, recording very carefully each day the data which he has collected.[3]

The fact that the investigator is actually participating in the behavior which he is studying is said to eliminate the tendency of subjects to change their behavior when they know they are receiving the attention of the investigator. Also, it allows the researcher to develop not only sound rapport but in many cases a high degree of intimacy with the subject so that it is possible to get information about taboos and other sensitive subjects. This information simply could not be collected directly in any other manner. The more competent researchers all reject completely the possibility of using a participant observer who is unknown to the people in the situation, because it would take a long time to get the kind of

---

[3] Luther C. Fry, The Technique of Social Investigation (New York: Harper & Bros., 1934), pp. 41-59.

familiarity that would allow the participant observer to ask questions on all kinds of diversified subjects without being suspected and by this time the investigator would become so enmeshed in the culture that he would lose his capacity to perceive many of the subtle details that he would supposedly be getting. Participant observers should be taken from the study situation and trained in data gathering procedures.

One of the most important advantages of the participant observer technique is that it allows us to get information which is much more accurate because in using a questionnaire or an interview or a life history approach, we change the situation in which the individual is reacting and this is certain to lead him to give us some responses which are not the ones he would actually give us were the situation unchanged by the intrusion of the technique. On the other side of this same point, the observer can perceive the information which he receives in its true context instead of in an artificial questionnaire or interview situation. While the participant observer actually changes the situation slightly, the impact which he has on a situation is not significant in comparison with the interview situation. It is obvious that the only research technique we have which will pick up gossip and other informal communication, is the participant observer approach and many times this type of information will reflect behavior which could never be located through direct questions.

A responsible investigator, George Lundberg, notes that actually there is a good deal of advantage in being identified as an outsider when an individual is playing the role of participant observer. There are many people who will talk to an outsider about features of their behavior which they never think about ordinarily, for no better reason than they are flattered by the prestige they see in being so important that someone wishes to study them. It has been found this is almost universally true, regardless of the sophistication of the persons who are being investigated. A participant observer who acquires the confidence and the trust of his subjects can soon find that he can ask them almost any question and get a well-organized, complete frank and honest answer from them.

There are certain very definite disadvantages to the participant observer method. The most obvious is the great expense involved in making any extensive analysis through this method. A second important disadvantage is that it takes a great deal of time to collect the information and once collected it is very difficult to quantify this data. A more subtle problem is the fact that it is extremely difficult for the participant observer not to become identified with some one of the groups which he is investigating and thereby become suspect by members of the other groups in which he is interested losing his supposed advantage with them. In conclusion we can say that the participant observer method is a means of collecting data which is so expensive and so time-consuming that it can be used only as a last recourse.

Informants. Informants are merely persons who live in the situation being studied with whom the investigator develops rapport and interviews extensively. The only conceivable reason for using informants at a time when we have perfected the depth interview and the focused interview is to get at information about taboos and other unaccessible material which people either evade or refuse to talk about. It is rarely possible to get a representative sampling of people for informants and although informants may give you a flavor of the culture and although they are undoubtedly a rich source of information when you know nothing whatever about the situation under investigation, the information obtained in this way is so unreliable and so time-consuming to get, that it probably would have only the most limited use in public administration research. The same information could be collected much more effectively and much more efficiently, and be much more reliable, using any one of a number of other techniques.

Charting. One of the established methods of the anthropologist is to map the situation which he is studying. While this is not directly applicable to public administration, some applied anthropologists have determined a method of charting the movements of people inside and outside of a building, which shows the patterns of physical contacts which the persons in the organization using these facilities have, and indicating therefore the layout that will provide the most effective pattern of contacts in carrying out work. First, a complete scale drawing is made of the building in which the organization is housed. The sidewalks and streets surrounding them, the elevators, the stairs, the layout of desks and work places within the building, and the location of essential equipment such as typewriters, files and such are all charted. Then a list of contacts which each person makes during the day is obtained on a form which will be described below. Charting this information gives us an indication of those people who function as centers of information, those who function as transmitters of this information, the informal communication system, those who function as blockers, and the like.[4] Although this application was not developed by the

---

[4] "Field Methods and Techniques," Human Organization, Vol. 9, 2:32, 1950.

anthropologist, it seems probable that this chart coupled with a sociometric diagram, would probably represent the whole informal communication system in an organization graphically. It would also be a means of checking upon the spread and the transformation of planted rumors.

The principal problem here is in determining some reliable means of sampling the organization at different times, so that we could have some assurance that the pattern represented is a stable structure rather than a chance phenomenon. Using the time sampling schemes to be discussed below, it is very possible that the reliability of this method could be made acceptable.

Quantitative efforts in anthropology. One of the most interesting developments in applied anthropology is the attempt to develop quantitative measures of culture. This movement appears to be largely the result of cross-stimulation between sociometricians and applied anthropologists for there are very few studies in this area, most of these studies having been done in cooperation with sociologists. Two studies of this type will be reported here as examples of this approach. The first concerned an attempt to study family structure in a rather complex cultural situation. The second concerns the study of inter-personal contacts in an industrial situation.

In an anthropological study of the family in a Mexican community, data were gathered on both the personalities of persons in the family and inter-personal relations in the family situation, and on the relationship of the family to other families and other institutions in the culture.[5] This is a divergence from the sociological approach to the family. The sociologists restrict themselves almost exclusively to the study of patterns of family behavior and have given very little attention to the relationships of the family to other institutions in the culture and to the structures of families in the culture. The site of the study was a village in Mexico with a population of 3,500 people. These were divided into seven traditional family groups which had come from seven tribes which long ago came together to form the village. The technique of collecting information was to use participant observers. There were seven people on the staff. Each of them took one family from each of the seven family units in the village. The family to be studied was selected on the basis of sampling procedures. Several weeks were spent in analyzing the local census figures, which were only a year old, and families were stratified on the basis of the

---

[5] Oscar Lewis, "An Anthropological Approach to Family Studies," The American Journal of Sociology, 55:468, March, 1950.

number of people in the household, income, occupations of people in the household, and so forth. Then a stratified sample of seven families was selected, using random methods of selection within each of the strata. After the family had been selected, one each of the seven members of the staff went to live with one of the family groups selected, and autobiographies of individuals and a case history of the whole family were collected. Projective techniques were also used to collect information. The completed reports for each family ran about 250 typed pages.

Commenting on the use of sampling here, the author said:

> The traditional anthropological reliance upon a few informants to obtain a picture of the culture and the people, though perhaps feasible in a small, primitive, tribal society, was inadequate to this situation. The question of sampling and procuring data and informants representative of all the significant differences in the village, was just as pertinent here as in the study of the modern urban community. Sampling and quantitative procedures were therefore employed wherever possible, as were census data, local government records and documents, schedules, and questionnaires.[6]

The traditional justification of the procedure of using informants has been that there is an essential homogeneity in primitive cultures and that therefore there is little need to use sampling procedures. This has led to a

---

[6] Lewis, loc. cit.

real methodological weakness, however, in anthropological studies of cultures. "An account of the culture based upon a few informants is bound to appear more universal than it really is."[7] It was felt that the case study technique and the projective technique were highly appropriate to this type of exploratory study because they gave some idea of the tremendous range and variety in customs and family life in this relatively homogeneous (modern) peasant community.[8]

The importance of this study in its implications for public administration research is that there certainly will be many areas where we shall wish to use the case study technique in the investigation of large, complex organizations. Even though we may collect data which cannot itself be quantified, the reliability of our conclusion will be significantly increased if we can at least show that the data which we obtained provides a representative picture of the situation under investigation. There is little doubt that a sample of seven families is an unduly small N upon which to generalize upon some 500-600 families, but certainly in almost all organizations we shall have data available by means of which we can set up a stratified

---

[7] Ibid., p. 472.

[8] Ibid., p. 432.

random sample which will assure us at least this much representativeness in our data.

The second technique concerned a method for the determination of interaction patterns in an industrial organization. The instrument was called a contact questionnaire. It is designed to locate patterns of contacts and to differentiate face-to-face contacts from those which occur by telephone or memo. As can be seen from the chart, the contact questionnaire provides us with the name of the individual, the date on which he records his contacts. Knowing the name of the individual, we can, of course, then go to personnel records and obtain a good deal of personal data which could be used in correlational analysis. It should be noted that names are used instead of titles of position, for it has been found that this provides more accurate identification in the eyes of the person who fills out the questionnaire than does the use of job titles. Also note that all the names of the people whom the person might normally contact in the execution of his job duties are shown in two classifications at the extreme left of the chart, and then there is a space for the individual to add the names of those who could not be anticipated when the question was framed. The rest of the columns are self-explanatory, but it should be noted that the most unreliable data which we get in using this questionnaire

WEEKLY WORK CHART

NAME____________________ DATE____________________

NAME	How many contacts per day with each person		Average length of contact	Proportion of contacts you initiate		Do these figures hold relatively constant through the week?	
	Face to face	Telephone		Face to face	Telephone	Yes	No
(This column contains a list of all persons who those in the organization might contact under three headings. 1. Names of everyone in the organization. 2. Names of those frequently contacted in agencies having working relationships with the organization 3. Names of those in other organizations with whom members of the organization come in regular contact. 4. Space for fill-ins.							

FIGURE 2

comes in column 3 concerning the average length of contact which is made.

If the data obtained are to be at all representative, the questionnaire must be submitted to personnel in an organization when the agency is in a relatively stable and normal period.[9] If such a period cannot be decided upon, the questionnaire can be administered three or four times, say on a different day in every third week. Those who have used the questionnaire recommend that it be administered at least twice so that we have a two weeks' sample to base our data upon.[10]

One of the unique advantages of this technique is that it has an internal check, for every time one person checks a name on the left of the questionnaire, someone should be at the same time checking his name on the questionnaire and making an estimate of the length of the contact, which can be cross-checked, and an estimate of the number of days per week that the contact is made and this can be cross-checked. By administering the questionnaire at least twice, then averaging the four estimates which we get for each pair of individuals who have a

---

[9] "Field Methods and Techniques," Human Organization, Vol. 8, 2:22, 1949.

[10] *Ibid.*, p. 23.

contact, we have a composite picture which undoubtedly has more reliability than any of the comparable measures we are using at present. If we desire a really accurate estimate of the time, all of those estimates which do not agree or which indicate only chance agreement can be checked through a personal contact. The data which we finally obtain will have a good deal of accuracy. After the contact questionnaire has been administered a number of times in the same organization, nearly all of the outside contacts would have been recorded and there would be little use for the D category in the extreme left-hand column.

The author suggests that to get a check upon the validity of the data obtained through the use of this questionnaire, a running record of the contacts which an individual makes during the day by 15-minute intervals could be made on a desk-audit basis for a stratified random sample of people in the organization before the contact questionnaire is ever administered, and these two sets of data can be compared.

To the writer, this technique is particularly applicable to the investigation of public administration problems. One application which seems immediately apparent is that this could provide an empirical measure of delegation of authority, simply by adding another column to the questionnaire where a record of the content of the contact would be made. Such an instrument would indicate the degree of over-supervision or under-supervision in terms of the number and content of contacts made between supervisor and superior and between supervisor and subordinates. A second application of this technique would be to check the effectiveness of the channels of formal communication through the coupling of the content column with the length of contact column and the nature of the contact. A third obvious application would be the discovery of the effectiveness of the operation of upward channels of communication. Also, since interaction patterns are almost the heart of leadership, a chart of interaction patterns superimposed upon a sociogram would increase our ability to interpret the distribution of real power in an organization a good deal more than just two-fold.

The thesis of the above discussion is that the significant techniques from anthropology which will be of interest to public administration researchers are those developed by the applied anthropologists. This point is brought out explicitly in the following discussions on the limitations of the use of anthropological techniques in the study of modern culture.

Limitations on the anthropological methods. Bierstedt holds that the major limitation of anthropological methods is that they were designed originally for use in non-literate, primitive societies. Some anthropologists are now recognizing this fact by making a distinction between the concept of primitive culture and civilization.

> The contrast between literacy and non-literacy is an extreme one, a contrast whose ramifications extend to every sphere of social life; for, while non-literate society has a language, it has no alphabet; it has customs but no laws, legends but no literature, techniques but no science, art but no esthetics, religion but no theology, a weltanschaung but no philosophy.[11]

Anthropological methods have been designed to detect language, customs, legends, techniques, art and religion, but the techniques have not been developed so that they will pick out the formalization of these phenomena in a complex culture. Case studies, observation by participant observers or non-participant observers, the use of informants and the like, and projective techniques may well disclose the customs of a people, for example, but to determine the laws of a civilization we must turn to a study of its institutions and use more penetrating methods of analysis. It is not held that these techniques fail to supply reliable data. The contrary is true. It is asserted, however, that these methods do not provide us with data with enough scope to support generalization.

---

11 Robert G. Bierstedt, "The Limitations of Anthropological Methods in Sociology," The American Journal of Sociology, 54:23, July, 1948.

In the study of primitive society, the use of informants is essential, for these people have no records or documents which can be examined and if the anthropologist is to procure a sufficiency of data he must use informants. The position of the social scientist in a modern culture is just the opposite of this. The problem of the social scientist is to select from the great variety of sources of data available to him, those which are most effective in providing him with generalizations. In addition to the great number of written manifestations of our culture, such as newspapers, novels, historical works, poetry, periodicals and the like, the social scientist has many scientific studies which he must review and examine to determine the state of maturity of the system of knowledge in a study area. The modern social scientist is dealing with a society which, to begin with, has a certain amount of knowledge, however unreliable, about itself. This makes his situation completely different from that of the anthropologist, who in dealing with a primitive society is dealing with a culture which has no knowledge about itself because it has no science. The very fact that we are literate allows us to have both a history and a science.

Bierstedt's second limitation is that anthropological methods are not designed to take account of the time dimension. A primitive society has no history so that the anthropologist who gets a time slice in his research which corresponds only to the length of time that he actually lived in the culture has as good a time slice as anyone could select. If one accepts the point of view that every culture has a history, whether it is available or not, and that many manifestations of the culture at the moment can only be understood in terms of historical antecedents, then the anthropologist, by definition, cannot provide acceptable explanations.[12] While there are obvious exaggerations in his criticism here, it is probably true that the anthropologist's methods are not designed to take full account of the dynamic dimensions of modern culture. In a complex society, where we have a history and a science, those things the people know about history and scientific knowledge have an impact upon the development of sophistication in the culture. There is a direct connection between the historical and scientific discussions carried out in the culture and the nature of the institutions which arise in it. This is the antithesis of the process of cultural development in the primitive society, where the institution tends to evolve almost without the members of the culture realizing it.

---

[12] Ibid., pp. 24-25.

Take just one example. The labor union movement was recognized as an important element of the economic institution in our culture as early as 1934 and 1935 when we passed legislation which dealt with it. Labor unions have received intensive study by social scientists since this time and the development of the labor unions themselves has been influenced by the conclusions which the social scientists reached. In fact, recently social scientists have written books on the philosophy of unionism. These books attempt to describe the present functions of the institution in our culture and will undoubtedly provoke a good deal of discussion as to the appropriateness of such functions. It may be that the people who are members of this institution will themselves change the function of their institution on the basis of their stimulation from these books.

Bierstedt holds that because the anthropologist does not take adequate account of the time dimension he misses many of the fundamental causal factors in the culture.[13] He feels that the anthropologist is likely to become concerned with small, trivial or subtle aspects of behavior which can be easily observed and infers that things exist while actually his data do not warrant this inference.

---

13 Ibid., p. 25.

> To assume, for example, that toilet-training--a cultural device which currently fascinates a number of anthropologists--can explain the rise of Hitler, the waning of the British Empire, the conflict between Hindu and Moslem, and American presidential elections, the power of Petrillo, . . . or any of the general type of phenomena of which these are specific cases, is to illustrate how ludicrous some of these approaches have become and steer perilously close to nonsense.[14]

This gentlemanly butchering of Margaret Meade is probably unwarranted, for she makes no attempt whatsoever to practice rigorous experimental methods and states it. Bierstedt is exhibiting a provincialism very typical of the experimentalist in denying the value of the concepts developed by a person who uses poor research procedure. The value of anthropological methods as a means of exploration and development of new concepts should not be denied merely because it is not appropriate to the discovery of fundamental dimensions. We should recognize it for what it is. In the primitive society, the only significant structuring is the kinship relationship and the anthropologist has developed techniques for reliable determination of this structure. In modern society the individual seldom plays only one role, as his role in relation to any given institution may change as his total social status changes and so the anthropologist's techniques for determining structure in modern, complex society, are not satisfactory. For that

---

[14] Ibid., p. 26.

matter, we do not have any techniques for the determination of structuring in our society outside of the sociogram and this is still of unknown validity. It would seem that wisdom would point to the anthropologist potentially being just as important as any of the other social scientists in this area.

Bierstedt admits that,

> Anthropology has done most to lay the ghost of ethnocentrism, the Baconian "Ideal of the Tribe", which in an earlier period imperiled the objectivity of social research.[15]

But he says that anthropology is currently giving rise to the modern place counterpart of this threat to science in Temperocentrism, which might be defined as the unexamined and largely unconscious acceptance of one's own sensory perception of one's own area of one's own lifetime.

It cannot be denied, and no anthropologist would deny, that the criticisms raised by Bierstedt are comprehensive and true. It is debatable, however, that the concepts of anthropology must be discarded, any more than the concepts of philosophy, political science or economics must be discarded because they were developed by people who did not use rigorous scientific methodology. Professor Bierstedt's discussion points up the danger of scientific

---

[15] *Ibid.*, p. 27.

provincialism as much as it points up the failure of anthropological techniques. The works of Warner, Leighton, Meade, to mention only three of a long list of anthropologists who have made important contributions, cannot be denied merely because they use a different method. In the writer's opinion, we disregard the concepts and the methods of anthropology at the risk of holding back the progress of science.

<u>The case study technique</u>. Wallis defines the case study technique in this way:

> The term "Case Study" as employed in the field of anthropology, economics, history, law, political science, . . . usually connotes a relatively intense analysis of a single instance of the phenomena being investigated.[16]

When the investigation focuses attention on the developmental sequences or the causes behind the development of a phenomenon, it is often referred to as a case history or life history technique. When the object of interest is a primary or a secondary group or even a whole primitive society, we find the term "case study" used in a more restricted sense.

The case study technique actually embodies a number

---

[16] Paul Wallis, <u>Supplementary Study A</u>, p. 183. Paul Horst, editor, <u>The Prediction of Personal Adjustment</u> (New York: Edwards Brothers, Inc., 1941), 455 pp.

of techniques for the intensive analysis of complete episodes. Each is designed for the investigation of episodes involving different numbers of people. In each case we are using the term "case study" to connote a research procedure. The case study technique is also the basis of prognosis in both medicine and psychiatry and the basis of diagnosis in industry, government and social work.[17] The case study has come to have an important position in prognosis because its outstanding characteristic is that it naturally brings out the pertinent factors related to a given problem. Where some techniques give us a representation or a sample of the dimensions of a total situation, the case study technique tends to obscure the proportions of the larger picture. Factors which appear to have the greatest importance in a case study are those most strongly related to the problem under investigation. The case study technique cannot be used to get an objective perspective of some particular class of behavior. Its research value lies in the fact that it does not do this but rather that it aids the investigator to perceive factors which may have even a remote relationship to the problem under investigation. It is for this same reason that there is controversy as to the practicability in using case studies

---

[17] Wallis, *loc. cit.*

for prediction. (Wallis, however, denies the inappropriateness of using case study data for prediction when a representative sample of case studies pertaining to a given phenomenon type of behavior is collected and the data from them are quantified.)[18] Many social scientists take the view that any major advances in predictive effectiveness in the social sciences will depend upon the optimum combination of case study techniques and psychometric procedures. Insofar as the case study techniques provide us with a methodology which gives us certain mechanical aids in perceiving new dimensions of situations, the writer is in complete agreement.

> In the end the individual appears as a person, as a microcosm of the group features of his culture. It is possible that detailed studies of the lives of individuals will reveal new perspectives of the culture as a whole which are not accessible when one remains on the formal plane of observation.[19]

> Most social scientists look upon the life history as a way-station through which one may or must pass in the search for valuable hypotheses which can then be verified "scientifically".[20]

Function of case study. Foreman classifies the function of the case study as follows: (1) Concept or

---

[18] Ibid., p. 184.

[19] John Dollard, Criteria for the Life History (New York: Peter Smith, 1949), p. 4.

[20] Ibid., p. 7.

hypothesis development; (2) hypothesis testing; (3) Prediction or post-diction; (4) methodological testing or refinement; and (5) illustration.[21]

The most important research function of the case study is in the development of new classifications of information or even of entirely new types of information. The identifying characteristic of the case study method is the free response situation of the subject. In a sense, the "person who knows more about the situation than anyone else", because he has lived in the situation, is assisted in telling the investigator everything that he knows or feels about the situation under investigation. In this way each case study, while it is in the broad sense a microcosmic representation of the larger society, also provides us with a slightly different set of variables and interrelationships of variables so that there is a good possibility of the experimenter discovering a new point of view and therefore perhaps creating new variables or new relationships of the variables. It is this same characteristic of the case study method which helps us develop testable hypotheses. By making case studies of organizations we are in effect framing hypotheses from a number of

---

[21] Raymond C. Cattell, "The Dimensions of Culture Patterns by Factorization of National Characters," _Journal of Abnormal and Social Psychology_, 44:410, 1949.

points of view in an attempt to find the most effective point of view. Each of these may be used as a point from which to view the experimental problem.

It is undoubtedly true that we could make a rough test of a hypothesis using the case study method in spite of what some psychometricians may say and this will be discussed at greater length below. The real question is whether the case study method is an efficient means of testing hypotheses. If we know enough about a situation to frame worthwhile hypotheses it is doubtful whether the tremendous expenditure of time and money required to make case studies is justified when we compare case study results with results obtained for the same expenditure of time and money using psychometric methods.

On the other hand it is very likely that there are certain hypotheses which require the collection of subtle, non-quantitative data. The case study method is the only practical means of testing these hypotheses. At the same time, it is probably true that with the collection of a little more data many of these hypotheses can be reframed so that they could be treated psychometrically.

The writer sees great value in the use of the case study technique in the development of research procedures. In the application of any research procedure to a social problem it is always possible to get worthwile suggestions on the specific way the research procedure should be applied by consulting those who are to be the subjects. If we are using anything but a laboratory situation it is probable that case studies of the application of any technique would give us detailed knowledge which would increase our ability to develop reliable results.

The methods of case study. The principal characteristic of case study techniques is that they do not structure the response situation of the subject. These techniques can be classed under the case study heading: (We use Horst's classification.) (1) The orientation questionnaire. This is the open end question schedule for procuring biodata that is typically referred to as the "Face Sheet." It is a form for obtaining data such as age, education, occupation, income, place of birth, religion, number of children, and marital status. "The more knowledge about an individual that the investigator can obtain this way, the more efficient probably will be the collection of subsequent data."[22] The purpose of this information is to give the investigator a preliminary orientation to the situation he has to study so that he will have some sense of the direction in which to project his questions. Each configuration of these items may be a clue to a whole set of

---

22 Wallis, op. cit., p. 187.

new items of data for the individual who is experienced in using this questionnaire. (2) The standardized tests--tests of aptitude, ability, personality, characteristics, intelligence tests and neurotic inventory. (3) Projective techniques. The Rorschach test, the Murray Thematic Apperception Test and the psycho-traumatic spontaneity test are all examples of this type. (4) The oral interview--the traditional method of the case study. Wallis feels that to use the oral interview technique most effectively it should succeed the application of these other three types.

No interview can be completely unguided or the participants will merely jump from subject to subject in a random fashion. The problem of the interviewer, then, is to know which of the items of data which he wants he may possibly get through the interview of a given individual and the tests mentioned above are often effective guides to this objective. The major methodological importance of the interview is that it allows the observer additional opportunity to develop inferences about the respondent's perception of the problem by intensive follow-up of any information which the respondent provides that appears meaningful.[23] It is the flexibility and adaptability that is the primary advantage of the interview. The interview

---

[23] *Ibid.*, p. 190.

permits the investigator to search for the configurations of factors in a particular situation instead of having to anticipate them. (5) Written reports by subjects. These may include autobiographical sketches, complete autobiographies or so-called biograms. For the person who has difficulty verbalizing or the person who has no time available for the interviewer, it is often effective to use the written report as an alternative to the interview. To be effective, the written report technique should embody a set of clear-cut, brief, explicit instructions as to the topics in which the investigator is interested. There is definite indication that the spontaneity and the unplanned nature of this type of data may be most revealing. (6) The participant observer. In each of the above methods, the full cooperation of the subject is assumed. When it is impossible to get the cooperation of the subject, we may revert to the participant observer method in which the investigator places himself in a position, unknown to the subject, and observes his everyday routines and interactions.

Participant observers may be friends of the subject and they may be paid by the investigator. One illustration of the use of the participant observer technique was in an industrial organization where it was felt that the workers had an agreed restriction on output but because of the cleavage between management and workers no data could be collected by investigators openly since they had to get the permission of management to enter the company and this would identify them with management in the eyes of the workers thereby setting up a block. Therefore, three participant observers were put into the organization, one at the management level, one at the supervisor's level and one at the worker level in the organization. In this way a sample of data was gathered on the nature of the split in the organization.[24]

Wallis points out that the written narrative and the oral interview are the only traditional methods associated with the case study.[25] He admits that the closed question-naire method and projective method of collecting data cannot really be classified as case study techniques because of the close restrictions which they place upon the response the subject can make. The questionnaire, by allowing the subject to give his own responses and set his own context and establish his own proportion of importance of various factors, will often point up new relevant factors not previously recognized as pertinent to a hypothesis when the investigator was forming it. Wallis notes that in

---

[24] Orvis Collins, Melville Dalton and Donald Roy, "Restriction of Output and Social Cleavage in Industry," Applied Anthropology, Vol. 5, 3:1-11, 1946.

[25] Wallis, op. cit., p. 193.

hundreds of interviews people have complained to him that the responses provided for them in the questionnaire were not satisfactory for expressing the opinion which they actually held and he notes that these were questionnaires where careful pre-testing of response patterns was carried out and an analysis of these pre-tests was made. His conclusion is that it is simply impossible to provide closed end response patterns which can completely represent the opinion configuration of all individuals.[26]

Foreman defines the case study technique as including only the personal documents, such as personnel records, life history, autobiographical writings, participant observer records and personal oral reports. This is the more widely held concept of the case study technique.

All of the first techniques listed by Wallis assume that we will be able to outguess the classification that the individual makes of his own behavior and it is simply impossible to do this in an area where we do not have a well laid out system of knowledge. To get indications on probable classifications we use tape recorded interviews and make analyses of emotional contexts in which information is offered. We have interviewers record their feelings as to the meaning of inflections of the voice and

---

[26] Ibid., p. 194.

gestures on the part of the respondent, and we make content analysis of the words that the respondent uses. None of this is possible using the standardized response data gathering techniques.

The writer rather questions whether there are many situations where it would be worthwhile to use all of the standardized tests which Wallis suggests we use before we use the free response technique, and it is questionable whether it is worthwhile to use any of them in most practical situations. The case study is typically used to push forward just one portion of the frontier of ideas in a given theory system, so that we are actually working from a scientific construct to begin with, and unless we are in a completely uncharted area, the information which we can get from these tests is probably readily available to us.[27]

Criteria of satisfactory case study data. Because case study datum is successful if it points out true

---

[27] In connection with the biodata technique, we may say with some assurance that the data obtained from this means are useful for orientation purposes alone. A number of studies have been conducted in an attempt to measure basic variables this way and it has been found in almost every case that significant measures come only when we go directly to the more fundamental and more immeasurable variables themselves. G. E. Manson and G. L. Freeman, "A Technique for Evaluating Assembled Evidence of Potential Leadership Ability," Educational and Psychological Measurements, 4:33, 1944.

relationships or if it points up a promising hypothesis, the standards against which we evaluate applications of the case study technique are much less concrete than those which we can use for some of the more precise methods. Three people have discussed the criteria of satisfactory case study and although there are some important areas of agreement there are some very important areas of difference in the discussions of these three authorities. Actually one of these is Dollard's set of criteria for the life history, which is a more specific type of case study, so some disagreement is to be expected.

One point upon which there is complete agreement is that the case study data should be relevant to the problem under investigation. In securing data from subjects you must have some principle of selection as a basis for the inclusion of pertinent data and exclusion of the data which are not related to the problem.[28] Foreman makes the observation that we require not just one principle for the selection of pertinent data, but a whole set of principles or, in fact, a set of assumptions which amounts to a theory system.[29] In preparation for making a case study,

---

[28] Wallis, op. cit., p. 198.

[29] Paul B. Foreman, "The Theory of Case Studies," Social Forces, 26:413, October-May, 1947-1948.

the investigator should collect background information through personal observation of the situation, review the literature and develop a set of hypotheses about the situation which amounts to a set of principles for the direction of interviewing in data collection. These principles provide the springboard from which the interviewer poses questions and objectives and toward which the interviewer guides the respondent. By sophisticated analysis of the set of principles before the interview and several times during the course of conducting the interviews, the investigator should locate the gaps and the inconsistencies in his set of principles and these should indicate areas where he should be especially careful to encourage full participation by the respondent. We immediately see, then, that there is nothing miraculous about the case study method as a means of exploration. It is, and cannot be anything else, but a set of procedures which will help the investigator gain insight. In the end, the case study rests upon the adequacy of the theory system presupposed by the investigator. To the extent that the investigator is able to sense gaps and inconsistencies in his theory system and handle his respondents in such a way that he fills in these gaps he gains insights.

That the oral interview is the most important case study technique in a purely exploratory study is apparent, for where written reports are analyzed, and in the case where crucial documents are reviewed, the investigator does not have the opportunity to carry out this probing. It is entirely possible, for example, that the investigator may give the respondent non-verbal cues which will bring out points which could not possibly be pulled up without the rapport which face-to-face contact of respondent and investigator provides.

Our second criterion of the case study is supplied by Wallis. There should be a sufficiency of relevant data.[30] There are actually two parts to this criterion; that is, there should be a sufficiency of relevant data for each point and there should be a sufficiency of major points. The problems connected with getting a sufficiency of information, once the respondent has mentioned the point, are not difficult to handle if one is aware of them. The first point concerns the subject. He may lack motivation in terms of his ego-involvement with the problem under investigation. If the investigator realizes this the obvious thing to do is to divert the interview back to the preliminary stage of selling the individual on the purpose of the study. Secondly, the respondent may lack the ability to get insight into the behavior of which

---

[30] Wallis, _op. cit._, p. 199.

he is a part. If this is the diagnosis, there is very little that can be done except to discard the respondent. Third, the respondent may lack the ability to verbalize. There are several solutions possible here. One is to attempt to develop rapport through a series of shorter interviews, until the individual has had a time to become accustomed and develop an identification with the interviewer. If this does not work, the interviewer can then attempt to get the individual to write the report for him. If this is done the attempt to bring on verbalization of the interviews will not have been wasted, for the individual will have learned what the problem is through them. Finally, an individual simply may not be able to understand a problem. Many individuals lack the ability, for instance, to comprehend high-level abstractions. If the problem involves a situation which can only be explained in these terms, the investigator will have to drop the respondent.

One thing which makes interviewing an art is that each individual who is faced by the interviewer is unique, and the first problem of the interviewer is to diagnose the degree to which he is motivated and the degree to which he has insight and the ability to comprehend abstractions, and then he must decide, as quickly as possible, whether it is worth spending a good deal of time attempting to motivate an individual having some assurance that he will have the requisite insight and ability to comprehend abstractions. The fact that individuals vary in these abilities introduces error into the case study data by definition. (Note that the same characteristics introduce error into test data unless the test has been pre-tested with controls on these factors.) While this type of error is not too important if the case study is being used in an exploratory role, this is the primary reason why predictions using the case study technique are very difficult. Standardization is virtually impossible.

The second reason for insufficient data is the investigator. The first problem which arises in every interview is developing motivation and the means which we use to develop motivation is to get identification of the respondent with the problem under investigation, and in most cases this means that we must get identification of the respondent with the interviewer. Unless an individual has an existing identification with a problem, it is very difficult to build up identification with such an abstract image, merely during the course of an interview. This means then that the investigator must be the type of person who can be sympathetic and interested in people. He must be a person who is visibly wrapped up in the problem he is investigating. He must be a person who has an understanding and a "feel" for people which enables him to respond to the non-verbal as well as the verbal cues of the subject. The writer can report from personal experience that this also requires that the investigator have a tremendous reservoir of energy, especially when the subject comes from a background which is very different from that of the investigator. Studies have shown that age, sex, social roles, class positions and related factors have a differentiating effect upon the ability of an interviewer to secure rapport.[31]

Another very important factor in developing motivation is the degree to which the interview embodies variety in terms of the respondent's interests.[32]

An all-important factor in obtaining a sufficiency of data is the total amount of time which is taken up in interviewing. It is a simple matter of probability that the longer an interview runs, the better the possibility that the respondent is going to provide information in all of the areas of interest to the investigator.

A third criterion for the case study data upon which both Wallis and Foreman agree is that the materials must reflect true occurrences.[33] From the scientific point of

---

[31] Ibid., p. 202.

[32] Loc. cit.

[33] Foreman, op. cit., p. 413.

view, of course, this is the most important of the criteria we have yet set down and it is characteristic of the whole case study method that this is the one that we are least able to fulfill. In the sense that we are never able to completely represent a true behavioral situation verbally, the case study technique and the psychometric techniques have the same shortcomings, but in the sense that we can control our bias and control our perception, and establish accurate measurements of the phenomena under investigation, the case study technique falls far short of the psychometric technique on this criterion.

An important means of insuring some validity in case study techniques is to develop a maximum amount of rapport and therefore motivation on the part of the subject.[34] By psychometric standards, of course, this is no control at all. Wallis provides the only real criterion available here in his concept of representativeness.

Having hypotheticated a theory system which operates as a set of principles for guiding and listing information in the interview and having anticipated the gaps in this set of principles, the interviewer should attempt to gain from each subject a complete response to each of the principles which he has and each of his gaps.

---

[34] Foreman, _loc. cit._

The object of the investigator is to get a valid picture of each individual in relation to a problem and since we cannot possibly get a complete picture of every individual involved, our next objective then should be to get a representative set of data about the individual rather than just a great deal of information about one particular relationship to the study problem. The control which we use to get a representative set of data about the individual is rigid conformance to and balanced probing on each of the principles and into each of the gaps in the theory system.[35] Again, the principle of having as long an interview as possible applies, for the longer the interview the more possibility that we will be able to get representative sampling of information about the individual.

> . . . even in the longest case study, some selection occurs, so that the investigator must always be on his guard that his data represents the person as he actually is rather than as he may have been upon one or two special occasions.[36]

Perhaps Dollard provides us with the most satisfactory criterion of validity in the case study in what might be called situational control. This involves the collection of data, not only upon the individual's description of his own behavior and perception of the social

---

[35] Wallis, op. cit., p. 202.

[36] Ibid., p. 203.

situation, but upon the social situation itself. We compare the individual's description of his perception of the situation and his behavior in response to the situation with the true situation and the behavior which would seem appropriate to it.[37] Other controls that we can use here are to check interview data with personal documents and with the data we obtain from friends or associates of the individual.[38] Another control is to interview people who are equally familiar with or were involved in the same situation and compare their perception and their description of the individual's behavior, with his own. Foreman suggests the rather imaginative control of simultaneously interviewing two people both of whom were involved in the study situation, after they have been interviewed independently, and have them compare notes and cross-check each other.[39]

We can also check for internal consistency in the data that each respondent gives us.[40]

Wallis suggests a rather unique check of having someone observe the interview and selectively respond to

---

[37] Dollard, _op. cit._, pp. 32-33.

[38] Foreman, _op. cit._, p. 414.

[39] _Loc. cit._

[40] Wallis, _op. cit._, p. 204.

the non-verbal cues which the subject provides and then checking the description of this second interviewer with the verbal record which the first interviewer gets for internal consistency.[41] He feels that the skilled interviewer will quickly note tone and inflection, gesture, hesitation, flushing and so forth. While they are not constant and do not have the same meaning in all individuals, they do have meaning as non-verbal cues in most individuals and can be used as a very important internal check. In the writer's personal experience one interviewer can do this merely by using a system of pluses and minuses, placing a plus before those statements where the non-verbal and verbal cues agree and a minus before those where they do not.

A second problem in validity in case study data, is to provide some assurance that the interviewer perceives the data supplied him in the same context as does the person who provides it. This is partially controlled when the interviewer maintains reciprocal identification with the respondent. Another means of control here is to provide at least two leads for each of the points the interviewer wishes to cover and then check these for internal consistency through the interviewer's separate responses.

The final criterion of an acceptable case study data

41 Ibid., p. 204.

is that there be a clear differentiation between the raw data and the interpretation of this data.[42] Foreman feels that very often the trained interviewer is more subject to this than the inexperienced one. The interviewer will perceive the raw data and record an interpretation almost unconsciously. He maintains, and the writer agrees completely, that if it is desired to make interpretations during the course of the interview, the only way of doing this without corrupting the data is to have two interviewers, one to elicit raw data and the other to interpret it.

There are several mechanical devices which can be used in connection with case studies. If we desire to make an internal check of inflection of voice, hesitation, the rate of talk, and similar factors, we can use tape recording equipment and make this analysis separately and away from the pressure of the situation, after the interview is concluded. Then there are a number of techniques, situational factors which must be considered in developing rapport and motivation, which need only be mentioned in passing here. The interview should be conducted in private, in pleasant surroundings with which the respondent is familiar; he should understand who is conducting the interview and its purpose. These and many other minor

---

[42] Foreman, _op. cit._, p. 415.

points can be picked up from any standard case study commentary.

<u>Special types of case study techniques</u>. The life history is a written account of the growth of a personality placed into an interpretive theoretical framework.[43] It is a technique for studying the individual in relation to his environment, especially his social environment, over an extended period of time. It has been used for studying family adjustments, group adjustments and the like, and it is included here because it appears that it could be appropriately applied to the study of the adjustment of individuals to organizational living. In the life history study the individual appears both as a creature of the social milieu in which he lives and as an explanation of the unique form of the segment of society in which he lives.

Dollard holds that an adequate life history document will point up the fact that every individual is born into a cultural setting. It describes the setting in which the individual is born and shows its impact on his early development.[44] At the same time the life history document will show the forces which molded the individual's development. It describes the development of individual

---

43 Dollard, <u>op</u>. <u>cit</u>., p. 3.

44 Ibid., p. 15.

characteristics which, after a span of years, account for change which individuals collectively bring about in the environment which at one time shaped them. The technique embodies the assumption that the individual interacts with his culture and it shows both of these dimensions as a means of giving an adequate picture of the individual. A second criterion of any life history is that it points up the relation of the individual to the specific institutions in the culture, especially to the family, the school, church and the work group.[45]

Third, the life history shows the development of socially acceptable modes of behavior and describes adjustments and the development of non-acceptable modes of behavior. It also provides a description of the sources of the individual's tensions and conflicts.[46] It is apparent that the life history document is heavily slanted toward the genetic approach in that it continually attempts to explain what is in the light of what was. In our discussion, the term "history" in the name of the technique is extremely appropriate for it literally assumes antecedents to every significant dimension of the individual's personality in his experience.

---

45 *Ibid*., p. 23.

46 *Ibid*., p. 25.

Dollard provides a detailed conceptual scheme for guiding the life history interviews which makes the use of this technique much more nearly a mechanical operation than any of the others.

Dollard's conceptual scheme has not been developed for the analysis of adjustment to organizations but to the culture as a whole, so it would have general applicability to the analysis of adjustment to administrative organizations but would require extension to include a great many specific concepts as to what constitutes adjustment to organizational living. From the standpoint of exploratory studies, the life history technique is probably one of the techniques upon which we can place the greatest reliance and one which will be an extremely rich source of concepts and hypotheses for empirical testing.

A second specific application of the case study technique is the biogram. It is essentially a short, written life history document. Where Dollard uses the interview technique, the biogram is a record written by the respondent in accordance with instructions.[47] It is a technique appropriate where one wishes to make intensive case analysis of a limited phenomenon and take into account

---

[47] Theodore Abel, "The Nature and Use of Biograms," American Journal of Sociology, 53:112, 1947.

a large number of people. The typical length of a biogram is 50 to 100 double-spaced typed pages, where the case study may run to six hundred double-spaced typed pages. As the author suggests, the biogram is an efficient rather than a highly revealing source of information.

The best biograms are written by persons who have become emotionally involved in the purpose of the investigation.[48] One of the means used to get this ego-involvement on the part of the respondent, is to have a well known group which includes most of the members of the population which you wish to study, sponsor and actively direct the procurement of the biogram. Another is to conduct a contest and this has been found a very effective technique. Fifty, one hundred or even two hundred dollars are offered for the best biogram relating to the problem. The best biograms are submitted by people who are readily able to verbalize. This is an obvious point and if the investigator uses some control, such as a short test, to check this, he will get better biograms.[49]

The biogram is exclusively an exploratory instrument having very little demonstrable validity. It has been found that if anonymity is guaranteed validity will be

---

[48] _Ibid_., p. 116.

[49] _Ibid_., p. 117.

higher. Also it has been found that if an appeal to a cause can be worked into the introduction of the contest or the technique the biograms will be more worthwhile.

This is admittedly a superficial technique for relying completely on the willingness of the subject to cooperate in a situation where little or very little is known about the social structure and the beliefs and attitudes and mores of the people involved and where we have no reliable picture of the sub-culture. It seems that the biogram data might be very effectively used as a means of gathering background information to frame our theory system for the application of more refined case study procedures.

Predictions through case studies. Wallis points out there are three general patterns that we can use in attempting to predict through case study data. The first is prediction by factors. In this approach the investigator attempts to identify the factors which the situation contains and to match the factors which the individual or the group possesses to these.[50] In effect, he sets up a theory system which he feels represents the factors which the situation requires and then he uses the theory system

[50] Wallis, op. cit., p. 209.

to measure the degree to which each individual possesses these factors. We attempt to collect data which directly reflect these factors but if we cannot collect such data the investigator must infer the degree to which the individual possesses the requisite attitudes, tendencies, motives, wishes, level of aspirations, and background. The critical point in this type of investigation is determining whether the data which we gather indicates the presence of one or more factors.[51] This is a problem of classification. It is not always readily apparent what type of data indicates a particular factor. It would seem that this is a mechanical procedure but it is not. Data obtained from an individual often appears to reflect every factor which we are looking for on the basis of the situational analysis. We may know that this is impossible.

The second method of prediction from case studies is from type. When we establish what he calls an empirical type we are going through essentially the same process as in the first type of prediction except that we do not attempt to make a direct match between the individual factors and the data. We make a match between the configurational patterns of the data and the situation instead.[52]

---

[51] *Ibid.*, p. 210.

[52] *Ibid.*, p. 216.

For instance, in an investigation of prisoners we set up several classifications of prisoners on the basis of the typical configuration of their behavior. Then we match these to the different degrees of prison security. In this situation the individual is not considered as a whole but against the established type pattern.

The third procedure for case study prediction is from the unique case. In this scheme we attempt to develop a total picture of the individual's personality and then we compare him in terms of individual differences with the configurational pattern of others.[53] Using this predictive technique we may not have much success in prediction but we will push out the frontiers of our theory system as we develop more and more comparisons. Wallis thinks that we should be able to develop a nearly perfect theory system through the ultimate application of this technique because it almost mechanically cues us in the perception of pertinent variables.[54]

One of the problems in predicting from case studies and, in a larger sense, handling case study data at all, is that it is not quantified in raw form. An experiment was performed where five professional people, including a

---

[53] *Ibid.*, p. 217.

[54] *Ibid.*, p. 218.

social worker, a psychiatrist, an experimental psychologist and several researchers, attempted to take recorded case histories and categorize the data contained in them. The general procedure was for eight of the nine people to sit down twice a week over a period of ten weeks and develop a set of categories which they thought were pertinent to the problem under investigation. They developed a detailed set of instructions so that the ninth person, who did not meet with them, could perform the same operation.[55] Then all nine listened to the tape recorded case history and categorized the data contained in it. Following this, statistical comparisons were made to determine whether it was possible to get reliable categorization of case study data. Their conclusion was that this is a very promising technique.[56]

In the first place, there were a number of categories where there was complete agreement between all nine people and it did not make any difference whether they had attended the conferences or not. There was no significant difference between the categorizations of people because of

---

55 Malcolm G. Preston, Emily Hartshorne Mudd, William L. Peltz and Hazel B. Froscher, "An Experimental Study of a Method for Abstracting the Content of Social Case Records," Journal of Abnormal and Social Psychology, 45:633, 1950.

56 Ibid., p. 645.

differences in sex, age, professional background or the like, as long as they thoroughly comprehended the instructions. It was found that most of the disagreements on categorization could be explained better in terms of weaknesses in the system of categories than in the people who were doing the categorizing and in several cases it was found that the information provided in the case study itself was contradictory and therefore did not lend itself to categorization. Their conclusion is that this is a promising technique whenever we wish to handle large numbers of case studies using categories we can perform statistical manipulation--at least a limited number of statistical manipulations.[57]

If one accepts the proposition that in exploratory studies we do not have to provide rigorous experimental controls, then the greatest limitation of the case study technique is that it depends upon the theory system which is used. Wallis says "The major determinants of the kind of data sought for in case studies, are, of course, the assumption and point of view of the investigator."[58] It might be added that the theory system the investigator uses reflects the appropriateness of the functions and

---

[57] *Ibid.*, p. 646.

[58] Wallis, *op. cit.*, p. 196.

point of view of the investigator. In practice, the comprehensiveness of the case study is actually limited by prior scientific knowledge of behavior which we are investigating.

The second great limitation on the use of the case study technique is that interpretation is often indeterminate. Two competent investigators will interpret the same data as indicating two different conclusions. There is nothing we can do here but admit indeterminancy. Even so, in the development of the empirically based theory system in public administration it is the writer's feeling that the case study technique is indispensable. Where the immediate problem is to open up a field for research; where the problem demands further conceptualization of factors; where the problem demands emphasis on the pattern of variables involved and where the problem is to determine a particular pattern of factors significant in a given case, the case study is not only indispensable but it is almost the only tool which we have for scientific investigation.[59]

The technique of clinical psychology. Another application of the case study technique is the pattern of study of the clinical psychologist, even though it can

---

[59] Foreman, *op. cit.*, p. 419.

hardly be applied by the layman. The approach of the clinical psychologist is to study past experience in terms of its developmental influences upon the individual as a means of understanding the present behavior of the subject. The clinical psychologist studies both the basic factors in the individual's adjustment to others, his general sociability and his emotional maturity; his physical history; and his adjustment to the major institutions of our culture, that is, the family, school, his vocational adjustment and his relationship with community and recreational organizations.[60]

At least one of the dimensions of the research in administration should be an attempt to measure the adjustment of individuals to the institution in which they spend their working hours. The clinical psychologists study the adjustment of the individual to his family in terms of pride in his family, his attitude towards the accomplishments of the family, the degree of intimacy of family ties, loyalty, the attitude of the parents toward each other, the presence of the type of affection demonstrated in the family group, the incidents of jealousy,

---

[60] Edward M. Westburgh, Introduction to Clinical Psychology (Philadelphia: T. Balakiston's Son & Co., 1937), pp. 271-275. Also, T. G. Andrews, editor, Methods of Psychology (New York: John Wiley & Son, 1948), p. 577.

nagging, irritability, nature of violent emotional outbursts, the homogeneity of interest and hobbies of members of the family group, and many other similar dimensions.[61] Their hypothesis is that the individual plays a definite role in the family group but that roles differ from one family group to another and that some of the roles are more healthy than others. At the same time the clinical psychologist describes these roles at a number of different time slices during the individual's development, giving, in effect, a representative set of data about the individual at a number of representative times during his life. The linking together of these common dimensions is one of the things which the clinical psychologist uses to account for the individual's adjustment.

This has at least two fundamental implications for the investigation of administration. The first is that the case history technique could be used to locate the pertinent dimensions of the administrative role. The mentally healthy individual develops a complex, internally consistent role in the administrative situation in the same way that the mentally healthy individual has a role in the family situation. Our problem then is the development of a theory system to be used as a set of assumptive principles

---

[61] Westburgh, op. cit., p. 271.

for the analysis of individual adjustment in a given situation in a given institution. Secondly, once such a theory system is begun, its refinement can be carried on through the use of psychometric techniques. A set of concepts to train supervisors to be effective clinicians would be of great practical value. The development of an astute clinical approach to supervision hinges, among other things, upon the development of such a theory-system. We need to have a standard terminology and a basis for the uniform classification of given types of behavior by all supervisors.

Another basic implication of the clinical approach for research in public administration is that the individual, as our basic scientific unit in social science, is a whole, if he is mentally healthy, in spite of the fact that he plays a role in several different institutions. To look into the subtleties of adjustment to organizational life, necessitates the study of relationships between these different roles. It is often held that violent emotional outbursts displayed by individuals in a work situation are the results of conflicts between their domestic role and their work role.

Although we have no clear-cut attempt to actually do anything like this as yet in any administrative research, the writer feels that a good deal of thinking of Cattell and Likert holds implication of this type of effort and it is only a matter of time until some research is carried out in this area. Worthwhile studies by laymen could be made in the initial phases of this work by a review of pertinent literature in social psychology and clinical psychology, establishing a very tentative theory system for rough case history attempts. Then analysis should be made of these and the more refined case history theory system established, the process being repeated again and again.

Interviewing. The most important single means of gathering data in exploratory studies is the interview. There is little need here to review the mechanical details of interviewing, but some of the more fundamental aspects will be pointed up.

The interview is essentially a meeting of two or more personalities.[62] It is a situation where two individuals meet face-to-face for the prupose of discussing some problem. If it is to be an effective meeting of two personalities, the subject should perceive the interview as a genuinely reciprocal situation where he will be talking with someone who holds common views about the problem

---

62 R. C. Oldfield, Psychology of the Interview (Cleveland: The Sherwood Press, 1941), pp. 6-7.

under discussion. Actually the interviewer may play a probing role and the subject only a responding role.[63] This applies to the interview where we are looking for information. In the clinical interview, of course, there may actually be a reciprocity in the interaction of subject and doctor.

The limiting factor in the interview situation is the respondent's inability to provide information for more than an hour and a half, or two hours at the most, so the interviewer must decide upon a balance between the scope of the interview and the intensity with which each point in the interview is to be covered.[64] For maximum efficiency, the interview should proceed from point to point as fast as possible, but the rate of the interview will be dependent upon the intensity with which each point must be covered to get from the respondent the data of the required depth.[65] This assumes adequate motivation on both sides.

An interview consists essentially of a cycle of phases of alternate stimulation by the interviewer and the cognition and response by the subject.[66] But the underlying

---

63 Ibid., p. 67.

64 Ibid., p. 89.

65 Loc. cit.

66 Ibid., p. 15.

relationship upon which the successful interview depends, is an interaction of the attitude level between the interviewer and the subject. The verbal stimulae which pass from the subject to the interviewer can never be a complete and true representation of the behavior which is being described. However, it is this complete behavior in which the interviewer is interested. Therefore, says Oldfield, the thing which the interviewer must do is to attempt to infer the subject's attitude structure directly from the verbalizations which he provides us.[67]

With this in mind, it is seen that successful interviewing involves a deeper interaction than mere conversation. It involves the exchange of symbols which bring out the attitude structure of the subject. To stimulate this type of response efficiently the interviewer should frame his questions from a set of terms which represent his anticipation on the attitude structure of the subjects involved.

In observing an interview the sequence of conversation might make almost no sense, and it may appear that the interviewer is not even talking about the problem which he is supposedly investigating. Garrett points out that this type of interviewing is impossible on a one-shot

---

[67] *Ibid.*, p. 89.

basis.[68] The normal person to obtain social acceptance in a variety of situations has learned to mask his basic attitudinal structure, therefore being able to make a superficial adjustment to a great variety of people. The interviewer must slice through this reserve before he can get rapport with the individual on this basic attitudinal level. Since it is probable that most respondents come to the interview situation the first time with some apprehension, or at least with some unstable anticipatory feeling, the process of cutting through this reserve and the defenses which the individual will naturally erect in the interview situation, successful interviewing cannot be completed in an hour or even two hours.[69] Most people cannot voluntarily lay aside their own reserve during a one or two-hour interview, and the skillful interviewer soon locates their limit and restricts himself to this area on the first interview, rather than run the risk of getting a strong reaction from the subject and eliminating the possibility of deeper rapport thereafter.

There is a good deal of indication that at least a minimum amount of planning produces a great increase in

---

68 Annette Garrett, Interviewing, Its Principles and Methods (New York: Family Welfare Association of America, 1942), p. 121.

69 Oldfield, op. cit., p. 14.

the efficiency of the interview.[70] It also seems that increasing refinement of the interview plan eventually provides additional restrictions on the course of the interview which decreases the efficiency of the interview. It may even threaten rapport.

What should this planning include? First, it should include some anticipation of the set the individual will have upon entering the interview. Second, it should provide a set of alternative leads which the interviewer can use to bring the subject face to face with the problem upon which information is required. It is a fundamental principle of interviewing that we start with the subject where he is at the moment he enters the interview.[71] Third, the plan should provide a set of alternative ways of explaining the objectives of the interview to the respondent. In some cases it is sufficient merely to tell the respondent what is desired, and in other cases there will have to be an interchange of ideas as to what is required, and in still other cases the respondent may have to be asked to summarize what is required after it has been explained to him.[72] Fourth, the information which is anticipated the

---

[70] *Ibid.*, p. 70.

[71] Garrett, *op. cit.*, p. 14.

[72] *Ibid.*, p. 122.

respondent can actually provide should be identified so that the interviewer will have some idea of the reasonable objectives of the interview. One common fault, even among experienced interviewers, is failure to hold the topic under discussion to the practical objectives of the interview. This preplan should include a series of techniques for developing the basis of interaction between the participants in the interview on the basic attitudinal level, rapport.

Some of the techniques of developing rapport are:

a. Agreement--disagreement. One of the really effective means of guiding an interview is by agreement or disagreement with the respondent according to whether or not the interview is going in the desired direction or whether the respondent is providing us with the depth of information we desire. A sensitive respondent can often be induced to explore deeper meanings when the interviewer agrees with him, while a more aggressive respondent may be stimulated equally well by almost consistent disagreement. The normal subject will respond in the same way to some combination of agreement and disagreement.[73] Partial agreement and disagreement is probably the most subtle of all means of locating the general structure and intensity

---

[73] Oldfield, op. cit., p. 82.

of the attitudes of the individual.[74]

b. The exhibition of surprise, or astonishment. This probably amounts to nothing more than a subtle form of flattery but it may be effective with almost any type of personality if used sparingly. Typical place where this technique might be appropriate is when an individual says he believes something but that he is afraid that everyone disagrees with him. To spur the respondent on, the interviewer might point out that this is something really new to him and is illustrative of the basic type of data he is searching for, and ask the respondent to amplify his views.

c. Seriousness and earnestness. The interviewer should maintain a serious and earnest set as long as the individual is cooperating along the lines in which he is interested. By carefully controlling his degree of seriousness and attentiveness the individual can be kept from rambling away. Note that this is a non-verbal means of control.

d. Frankness or recklessness. With the strong personality, frankness and aggressiveness can be an effective and not too subtle means of regaining control of the interview, if agreement or disagreement fails to provide this. The frank retort, or a slightly reckless expression

---

[74] _Ibid._, p. 84.

of views, when carefully used, is often a means of controlling the direction of the interview. The basic principle involved here is that when the subject expects you to respond to him in one way and you actually respond in a quite different way you throw him off balance momentarily.[75]

e. Humor. Humor is the means of reducing tension and a means of building confidence in the respondent. It is also a means of controlling the dominant personality through subtle ridicule.[76]

f. Sympathy. This is a very important technique for handling the person who is unable to cut through his own inhibitions. The principle upon which sympathy operates is that it gives the individual motivation to cut through his inhibitions by offering him release from inner tension and therefore, relief.[77]

g. Interest or lack of interest. These are devices for controlling the direction of the interview.[78]

h. Formality. It is often necessary to open an interview under rather formal tones to respond to the subject's preconceptions of the interviews and therefore

---

[75] *Ibid.*, p. 85.

[76] *Ibid.*, p. 86.

[77] *Loc. cit.*

[78] *Ibid.*, p. 87.

set the stage for rapport. Having begun the interview on a formal level, it is often possible to get a good deal of rapport with the individual by immediately shifting to an informal air, and perhaps by calling the respondent by his first name.[79]

i. Affecting to misunderstand. One of the easiest means of getting the subject to provide more details is to feign misunderstanding and if used with some sincerity this is an effective technique with almost no threat to rapport.[80]

It should be noted immediately that the techniques described above can never be applied mechanically. They must be applied sincerely within the context of the interaction pattern of the interview. As rapport comes to be developed on deeper levels of attitudinal interaction, the subject gets more and more insight into the behavior of the interviewer and unless these techniques are used sincerely, unless there is a sincere disagreement or a sincere

---

[79] The writer feels this is especially true of university-conducted research where the subjects anticipate the university personnel to be different in kind than themselves and when the interviewer attempts to build rapport with the subject by adapting himself to the subject's attitude structure. It is typical that there is a sudden irrational loss of orientation toward the interviewer which may often go past this and take the form of either superficially disguised or even open aggression on the part of the respondent. *Ibid.*, p. 87.

[80] *Ibid.*, p. 99.

misunderstanding, for example, albeit prompted by the interviewer's reviewing these techniques in his mind, this rapport may be broken and destroyed and perhap irreparably.

It will be seen that many of the techniques discussed above can be used to control the rate at which we proceed from point to point. For maximum effectiveness, the interview should proceed from point to point as fast as possible, getting the required detail on each point at the time we are discussing it. Note then that this requires in the preplanning stage a rather clear cut idea of the amount of detail desired on any given point to provide sound data. The interviewer must keep continually in mind that the respondent will tire by the time the interview has proceeded to the stage where there is interaction on a basic level. Even the strongest respondent will not be able to maintain this rate without some letup in the pace. If we continue the interview to the point where the subject is unduly fatigued, he may well get a reaction after leaving the interview situation, and repeated interviews will be more difficult, rapport being reduced.

The greatest single threat to rapport is insincerity on the part of the interviewer. Enough has already been said of this. A second common threat to rapport is the interviewer attempting to put words in the mouth of the respondent. This leads almost universally to an aggressive reaction against the interviewer. A third threat to rapport is the interviewer trespassing upon what the subject feels to be his private personal views. This actually amounts to little more than the interviewer getting anxious about the pace of the interview for ultimately when complete rapport is established almost any subject will talk about any of his so-called private attitudes as any "good listener" already knows.

A final part of the plan should be a series of alternative means of closing the interview upon a positive tone so that the interviewee goes away with the feeling of having contributed something.

Matching interview technique with problems. The effective collection of data through the interview method requires the selection of an appropriate interview pattern for our problem. There appear to be at least two factors involved here; whether or not strong rapport will be necessary to obtain the desired data and the degree to which we know the probable responses that the individual will give. If we are making such a completely exploratory study that we are still attempting to discover the probable responses, the problem is really quite different from a study where we are attempting to find out how many of each type of anticipated response occur in a given population and what configuration they assume.

The nondirected interview is a method where the interviewer tells the subject the general problem under investigation and then asks him to reflect upon it and describe his reflections. As a research technique it is appropriate only where we require penetrating analysis of some aspects of the individual's personality or role in relation to some pattern of conditions in the organizational situation. For example, if the interviewer wished to study the role which the individual felt he was playing in relation to a primary or secondary group in an organization the problem can be stated and then as the individual attempts to describe this role to the interviewer the interviewer would do nothing but reflect the role described back to the subject in an attempt to get the subject to make even more penetrating and revealing statements about this behavior.[81]

The whole nondirective interviewing technique is based on the assumption that by reflecting the subject back to himself you give him the basis for deeper insight into his own personality.[82]

---

[81] Carl R. Rogers, "The Nondirective Method as a Technique for Social Research," The American Journal of Sociology, 50:282, July-May, 1944-45.

[82] Ibid., p. 283.

As a research instrument the undirected interview is the most unsatisfactory of all techniques. First, it is extremely expensive and requires often what seems--to be an unlimited number of interviews before insights related to the problem are apparent. Second, it takes very skilled personnel to handle these interviews effectively. Third, it takes a great deal of the respondent's time and energy and therefore is not practical except when there is some motivation in the respondent which can be used as a means of building up an identification between himself and the research project. Fourth, while it is an exciting exploratory technique, the results obtained through this method are so grossly subjective that they require refinement before they approach anything like reliability and validity.

The focused interview is a means of increasing the efficiency of our data gathering and at the same time allowing for the possibility of depth interviewing as:

a. To attempt to be as nondirective as possible in order to elicit all of the pertinent responses from the subject;

b. To be as specific as possible and get as much detail from the subject as possible;

c. To attempt to maximize the range of possible responses;

d. To attempt to get depth and personal contact, that is, to get the latent implications of the subject's responses.[83]

The focused interview is based on the assumption that we have some reliable knowledge as to what the areas of response are in the behavior we are investigating and then we design questions to bring out responses in these areas. We do not attempt to anticipate the responses which we will get completely. Care must be taken to make the questions unstructured--that is, make the questions so that they do not cue the respondent to a specific answer.[84] For example, instead of asking the question, "Do you do this or that when given things happen?" we say, "What stood out especially in your mind about this event?"

The focused interview is an attempt to keep a person's mind focused on selected aspects of behavior in which he participated but not to cue him to any particular response. It has been found that it is often difficult to get specific reactions using this method but it can be done. One way is to get the person to give a general reaction to the situation and then ask him to provide more detail or

---

[83] Robert K. Merton and Patricia L. Kendall, "The Focused Interview," The American Journal of Sociology, 51:545, July-May, 1945-46.

[84] Ibid., p. 546.

feign misunderstanding. Another means is to get the individual, when he responds with a general reaction, to break this into smaller parts. As he describes details of the event the interviewer selects these parts in which he is interested. This process may be repeated by breaking these into more detailed segments.[85]

In this type of interviewing the range of subjects is generally restricted in the planning stage. The set of questions may be modified once there is some indication in the responses of the subjects that the range is too limited. Burton suggests that the pre-planning should be so careful that alternative cues for transitions from point to point in the interview should be laid down.[86] He points out that when a response is given where the interviewer decides he wishes to get data on more depth of the subject which is made a part of the response, sympathy is often a useful technique. Also, one might ask the respondent to compare the feelings that he has with the interviewer's own feelings.[87] Certainly the focused interview is one of the most effective means that we have for exploratory studies. It is much more efficient than any of the case study techniques

---

[85] _Ibid._, p. 550.

[86] _Ibid._, p. 553.

[87] _Ibid._, p. 556.

and yet provides flexibility and depth of response.

The directed interview is a technique with very limited application since in most cases if we can use a directed interview we can just as well use a questionnaire. It is an interview situation where the interviewer feeds questions to the respondent and records his responses in discreet terms. It presupposes a rather stable theory system. It is useful when we have to ask questions against which many people fail to provide a written response because of taboo or for some other reason and when we wish to determine what type of responses we should build into a paper and pencil test, although in this case it is questionable whether it would not be better to just run a pretest and attempt to make a statistical analysis of this. In the end, the directed interview technique is largely a means of examining our methodology rather than a method of content exploration.

Finally, there is the group interview situation. Group interviews are used extensively in therapeutic work but they have not been used as a research technique. It is obvious that in this technique it is impossible to get the very deep rapport with any given number of the group since each person will be sensitive to the values and mores of the rest of the group. In essence, the group interview is a conference where the leader attempts to keep the climate as permissive as possible and then guide the group in discussing the problems common to all of the members of the group.[88]

It should be noted that in using this technique it is almost impossible to make a very rigid plan for the interview. The function of the leader is to let the members of the conference guide it themselves. In fact, it is recommended that the leader open the conference by pointing out that he could lecture on the problem to be discussed but that in all probability the combined experience of the people in the conference is so much greater than his that the best way of facing the problem is for everyone to participate in discussion. From here on the leader attempts to restrict himself to the role of coordinator.[89]

In discussing the use of this technique it is pointed out that it has been found that in order to get the discussion started the leader may have to stand and wait 25 or 30 seconds or even as long as a minute. If, after a minute, no one has volunteered to begin the discussion, the leader may make some kind of a "crack" to attempt to break

---

88 Chris Argyris and Gram Taylor, "The Member-Centered Conference and the Research Method, I," Human Organization, Vol. 9, 4:6, 1950.

89 *Loc. cit.*

the silence and then wait once more. When intensity develops again and no one still has volunteered to break the silence he can perhaps pick a member of the group who looks as if he is attempting to step forward and get him to begin the discussion. The hypothesis is that if the leader begins playing a part in the selection of the subject matter in the opening basis of the discussion, the members of the conference will look to him for this function during the rest of the discussion and he will not get the spontaneous type of free response which is one of the values of this technique.[90]

Several unique types of data can be obtained through the conference method. The relationship between the individual member of the conference and some sub-group or other member of the organization as a whole may be pointed out. Secondly, the structure and function of the primary groups in the organization may be pointed out. Third, the effects of industrial phenomena such as budget and controls, production records and the like upon the role of the individual may be pointed out. Fourth, a picture of the personality of someone not in the conference may be contributed by several participants. Fifth, information concerning the relationship of other institutions, such as

---

[90] Ibid., pp. 6-7.

the home, church and the like, to the industrial situation may be pointed out. Of course, the source of data in this conference is a verbatim record, preferably taken on tape. This is very easily done by bringing the group to suggest that minutes of the conference be kept and then encourage the suggestion that a tape recorder be used.

The first big advantage claimed for this method of interviewing is that it is economical whenever the subject matter is such that it can be used. It is mentioned here because many organization phenomena seem susceptible to this approach.

It is often possible to get more intimate response from the individual in the conference situation than in the private interview situation, for if he knows the members of the conference they may pick up where he leaves off and contribute something which he knows. He may then go on from there, so that the other members of the conference serve as a source of stimulation which no outsider could fulfill. But the major advantage of this research technique is that by having people in the conference from all levels in the organization, it is possible at one time to get a composite picture of the structure, or the function, or the process of a group. In fact, it is claimed that this technique provides a unique picture of the organization because in the conference situation people from the bottom level will stimulate people in the middle and top levels to bring out points which neither could contribute alone, so that the end result tends to be greater than the sum of its parts because of this cross stimulation between people in different levels and in different roles during the time the picture is developing in the conference. Each person who responds after this picture is becoming developed is responding from a different context. The conference leader could never have developed such a content for himself.[91]

It is interesting to note that a role-playing situation has been used to train interviewers. Groups of interviewers who were to administer a sociological questionnaire were brought into a role-playing situation of five sessions, each lasting two or three hours.[92] They learned to play the role of a sympathetic listener, the inquirer, the analyst and the director of action. It is claimed that by using this technique the number of refusals is almost eliminated. There were only three refusals in two hundred and twenty-five interviews. There is a significant

---

[91] Marguerite R. Hertz, "Rorschach: 20 Years After," Psychological Bulletin, 39:11, 1942.

[92] Lillian W. Kay and Jane H. Schick, "Role Practicing in Training Depth Interviewers," Sociometry, 8:83, 1945.

increase in the validity of the statements reviewed, also.

Typological analysis. Some of the most effective exploratory techniques are those using descriptive measurements which can be given larger meaning. The usual pattern is to make some very simple measure of behavior. The measurement is dependable because it is so simple. If a complex process is measured we are often able to make such interpretations.

Loomis tried to measure the distribution of real authority in a hierarchy. He felt this was the most true representation of the hierarchical order. He used a method devised by Chappel where authority is described in terms of the number of times person A gets person B to do something, versus the number of times person B gets person A to do something. Where there are the same number of responses in each direction the relationship is said to be two-sided or equal, and where there are significant differences in the number of responses the relationship is said to be one-sided, or a superior-subordinate relationship.[93] Loomis

---

[93] Note that using the same technique we could get one measure of the correspondence of the formal and informal organizations by merely adding a content dimension to the response pattern and checking the degree to which the combination of dominance and content correspond with the designations of delegated formal authority. Charles P. Loomis and J. Allen Beegle, "Typological Analysis of Social Systems," Sociometry, 11:159, 1948.

does not put this technique forward as a comprehensive measurement of the hierarchy but only as an operational definition of what in this study is called hierarchy. This is perfectly justifiable in the overall pattern of the typological method.

The second dimension which is hypothecated is that of the degree of solidarity which results from the conversion of interests and temperaments of subordinates and superiors, in comparison with their potential antagonistic interests. A measure is taken of the conversion of the attitudes, interests and sentiments of each of the members of the group and then a coefficient of homogeneity or solidarity is established.[94] Another important dimension is that of short primary channels of communication versus long secondary communication channels. The importance of communication, he says, is primarily that it is an indirect measure of perception. Communication is only a catalyst itself, having little or no importance as an independent organizational objective. The method is to combine measurement of length of channel, nature of channel, the number of people involved, the media of communication to be used in the channel to one common measure and develop a quotient for this.[95]

---

[94] Ibid., p. 162.

[95] Ibid., p. 164.

Rationality versus non-rational behavior is Loomis' third dimension. This dimension concerns the degree to which the actual life processes of people in a social situation are determined only on the basis of facts or whether tradition and customs and other affectual factors are important in determining the behavior of the individual. Ritual is given an important place in this concept because in administrative organization, procedures frequently take on an affective connotation of ritual. The people who are involved in these procedures come to feel that they are an end in themselves and not merely a means to an end. Another specific dimension taken into account here is the degree to which people have a "sacred attitude" toward organizational processes. That is, the degree to which they have an emotional set to question social behavior in the organization.[96]

Loomis goes on to develop what seems at first glance to be an almost unique set of dimensions for the analysis of organizational situations, using dichotomous concepts which lend themselves to scale.

Following is a representative list of the concepts which Loomis sets down in this article:

Two-way--one-way authority.

---

96 *Ibid.*, p. 168.

Voluntary--compulsory behavior.

Solidarity--antagonistic attitudes, interests.

Short communication--long communication channels.

Direct communication--indirect communication.

Sacred--secular.

Traditional--rational.

Emotional--rational.

Personalized authority--impersonalized authority.

Blanketed rights--limited authority.

Blanketed responsibility--limited responsibility.

Having set down the concepts and a means of measurement, and having actually taken measures of the two organizations under study, each of the measurements is plotted along a five-point scale with the poles of the concepts at either end of the scale. The comparison between organizations comes when all of these scales are placed one below another and a single continuous line drawn from scale to scale, representing the measurement on each scale for the organization which was studied.

Examination and criticism of the initial profile led to the development of a second list of scales. This list is felt to be so interesting that it also should be

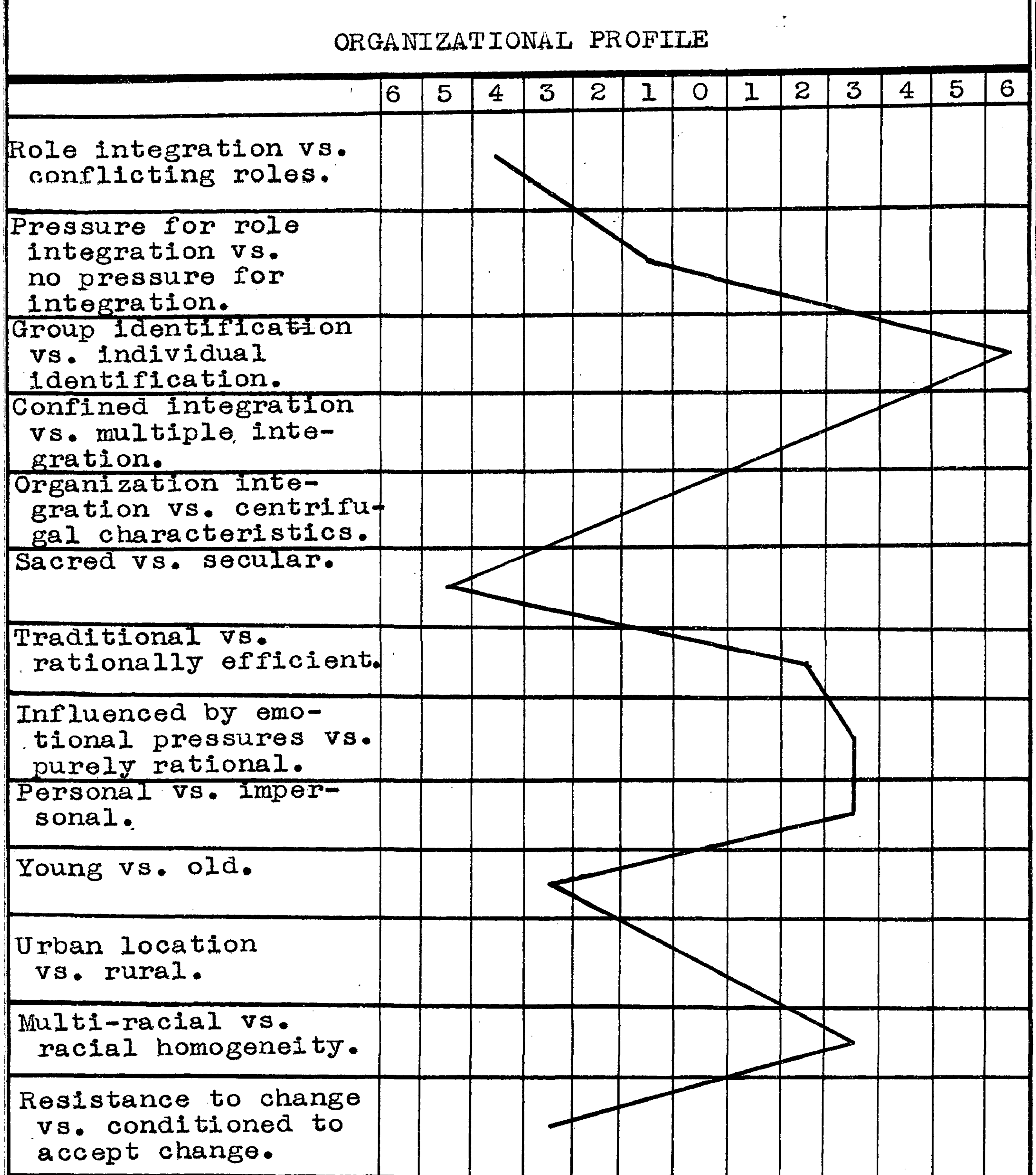
ORGANIZATIONAL PROFILE
6 5 4 3 2 1 0 1 2 3 4 5 6
Role integration vs. conflicting roles.
Pressure for role integration vs. no pressure for integration.
Group identification vs. individual identification.
Confined integration vs. multiple integration.
Organization integration vs. centrifugal characteristics.
Sacred vs. secular.
Traditional vs. rationally efficient.
Influenced by emotional pressures vs. purely rational.
Personal vs. impersonal.
Young vs. old.
Urban location vs. rural.
Multi-racial vs. racial homogeneity.
Resistance to change vs. conditioned to accept change.

FIGURE 3

reproduced here. It was as follows:

> Integration of role of individual into a system--conflicting roles of individuals in the organization.
>
> Required integration of roles in and out of the system--roles outside of the system irrelevant.
>
> All responsible for each--each responsible for all--limited individual responsibility.
>
> Members interaction confined to system--interaction distributed to many systems.
>
> Solidarity in teams--antagonism.
>
> Sacred--secular.
>
> Traditional--rationally efficient.
>
> Influenced by emotional--rational.
>
> Personal--impersonal.

One of the very interesting comparisons between these two sets of dimensions is that the role concept received a substantially more important position in the second list, after the valuation was carried out.[98]

The important characteristics of these lists for the public administration researcher is that it seems perfectly practicable to attempt to classify large

---

98 *Ibid.*, p. 178.

organizations. To be effective the concept should be as completely operational as possible and should be accompanied by a technique for the measurement of operational phenomena. Also, each dimension should be as independent of the rest as possible.

To the writer, the problem of bridging the gap between quantitative and qualitative measurement appears to be one of the more important which researchers in public administration will face. The study by Loomis illustrates how important this technique could be in the development of new dimensions of behavior in organizations and as criteria measurement. Once we have determined that these dimensions we have observed actually exist in two such completely different organizations as a Ditch-Association and an Amish family, we could set about to develop instruments for the measurement of the concise manifestations of any one of these given types of behavior, in two organizations. The first problem in each case is to determine whether or not there is actually a continuum which both of these organizations fall upon. There is no alternative to a subjective estimate as a point of beginning here.

Time sampling. Time sampling is a method of observing and categorizing behavior under the ordinary conditions of everyday life in which it occurs. Observations are made in a series of short-time periods so distributed that they provide a representative sample of the behavior under observation.[99] Time sampling is not an experimental method but a form of controlled observation where the recording of data is standardized. The segments of behavior observed are controlled through the use of sampling techniques. Time sampling is appropriate only to the description of broad behavioral situations where we can point out the pattern and general sequence of behavior but specific behavior items must be overlooked because of the nature of the procedure itself.

Time sampling resembles typological procedures in some of its important characteristics. We use the time sampling method to check classifications which we establish. It could very well be used for evaluating a typological study.

Time sampling has been used to record the frequency of behavior as well as to classify several types of behavior under broad categories and then determine the frequency within each category.[100] Time samples have varied in length from three seconds to three hours. Samples have been taken

---

[99] Ruth E. Arrington, "Time Sampling in Studies of Social Behavior: A Critical Review of Techniques and Results with Research Suggestions," Psychological Bulletin, 40:82, 1943.

[100] Ibid., p. 85.

TIME SAMPLE CHART

Organization ____________________ Unit ____________________

Recorder ____________ Date ____________ Time: From ______ To ______

	I					II					III					IV					
	A	B	C	D	E	A	B	C	D	E	A	B	C	D	E	A	B	C	D	E	Remarks
0:00																					
:30																					
1:00																					
:30																					
2:00																					
:30																					
3:00																					
:30																					
4:00																					
:30																					
5:00																					
:30																					
6:00																					
:30																					
7:00																					
:30																					
8:00																					
:30																					
9:00																					
:30																					
10:00																					
:30																					
11:00																					
:30																					
12:00																					
:30																					

FIGURE 4

for as short a period as one day and over as long a period as several months.[101] Since the sample is carefully selected the time sample has a great advantage over the autobiography or biogram technique in that it purports to actually get a representative selection of items of behavior rather than letting the individual select the behavior he will tell us about. If our problem is one where we wish to describe whole areas or sequences of behavior, such a procedure is probably much more applicable than the biogram or autobiographical method.

While the time sampling technique has been used predominantly in the description of children's behavior, the results of its limited use in the study of organizations is encouraging.[102] It has been used effectively in the observation of informal work situations where a glass screen, which could only be seen through one way, was used.

The chart on the preceding page is illustrative of one which might be used in a time sampling study. The time interval down the side of the chart depends upon the specificity of the behavior under investigation. The number of large columns across the chart would depend on the anticipated types of behavior which would occur and

---

[101] Ibid., p. 86.

[102] Ibid., p. 89.

then the number of sub-columns under the types of behavior depend upon the number of people being investigated. The capital letters stand for each person. This time chart could illustrate a hypothetical situation in a clerical pool involving five women. The major headings across the chart might stand for:

Column I -- Typing, transcribing and other use of the typewriter.

Column II -- For filing and other routine activities.

Column III -- Answering the telephone, talking on the telephone, answering and talking on inter-office communication system.

Column IV -- Represents informal interaction.

If more people are involved, different forms of lines may be used to indicate different functions. Dash lines could indicate dictation while the solid lines might indicate transcribing records and x'd lines show straight typing such as filling out forms and the like. In the last column verbal behavior might be categorized according to whether it indicates cooperation, conflict, aggression, passive acceptance and the like. Note that any other classification could be used.

The time sampling method has much to recommend it for studying organizations. It can be used to gain a broad picture of the total social situation in a work group, which would otherwise be almost unavailable. We should immediately point out that there are many differences in the records kept by two observers who work independently, but that the reliability computed on the basis of a comparison of records of the same behavior is relatively high and that validity computed on the basis of odd--even correlations of a large number of time samples is acceptable. The big disadvantage of the time sample technique is that normally we seldom take samples for more than fifteen or twenty minutes and therefore if the behavior which we are interested in investigating does not occur once in every fifteen minutes on the average during the day, our time samples will give us very little data about the total situation.[103] A second limitation on the time sample technique is that we must have behavior which can be identified through careful verbal definitions. If the behavior to be observed is so subtle that it is not easily identifiable, then we will get a great deal of disagreement between observers and so little reliability that the technique is not worth applying.[104]

The time sampling technique is another one of those methods that is a step between the pure exploratory study

---

103 *Ibid.*, p. 96.

104 *Ibid.*, p. 100.

and the perfected test. If we had the knowledge about the situation which we attempt to get through the time sample method, we would be able to make questionnaire schedules and perhaps even tests and get much more precise data. On the other hand, if we use case study or life history techniques to investigate this type of situation we risk using data which is much more biased.

Factor analysis. The only psychometric technique which the writer would class in this category of exploratory methods is factor analysis. It is, however, one of our most powerful exploratory methods for it actually opens up dimensions where none existed before. This statistical technique often baffles the layman and at the same time it impresses him since it approaches the process of pulling oneself up by one's own bootstrap, although those who are mastering the use of this technique point out that this is more apparent than real. Factor analysis is, as its name suggests, a means of analyzing tests to determine which factors any set of tests are measuring in common. It differs from the other techniques of exploration in that it requires tests. The identification of common factors is a primary objective of factor analysis. There are several published studies using factor analysis which are most interesting and which will bring out the application of the technique.

Cattell has attempted an experimental location of some dimensions of national cultures. In this study national groups were dealt with as units and the attempt was to discover the major psychological dimensions which would describe and define a national culture.[105]

One of the characteristics of factor analysis is pointed out by the problem they initially faced, which was that of treating whole populations as units. The most widely used factor analysis technique is the so-called R technique, where individuals are taken as the basic unit of the study and the factors which are reflected in the test are factors common to individuals. Here Cattell proposed to treat sixty-nine national groups as his population.

An alternative means of factor analysis is the Q technique or so-called inverted technique, where we consider the test instead of the person, and the population of units are the tests.

The third alternative is the P technique where we take the occasions on which the tests were administered as our units and we factor analyze these to get common factors. In each case you get different factors which are dimensions of different phenomena. We have dimensions of individuals,

---

[105] Cattell, _op. cit._, p. 443.

tests or occasions according to the technique of analysis. The major objection against assuming that a group can be treated as a unit and therefore that the R technique used by Cattell is justified is that ultimately groups are made up of individuals and unless we wish to assume that there is some phenomenon like a group mind this assumption is tenuous. Cattell feels, however, that we are justified in making this assumption because the group is a <u>functional</u> whole.[106]

Three broad classes of variables were to be used in making up the items to go into the test.

1. Group mentality--the acts of the group as a group, its stability and its mobility.

2. Group structure--data on the internal relationships of these national groups.

3. Population characteristics of the groups.[107]

In the end, eighty-two variables were set down from these three broad classes. Ten had to be discarded when no data were available on them leaving a total of seventy-two variables or tests on which to perform factor analysis.

Illustrative of the variables selected are: gross population, expansion of area politically controlled,

---

106 *Ibid.*, p. 444.

107 *Ibid.*, p. 445.

percentage of population protestant, percentage of governmental expenditure towards armaments, expenditures abroad per head, percentage of population illiterate, deaths from suicide, deaths from cancer per thousand population, number of Nobel prizes in science, literature, etc., severity of industrial depressions or percentage of industrial unemployment, ratio of number of tertiary to primary occupations, sugar consumption per head, percentage of population of Mongolian race, number of clashes with other countries, number of assassinations, percentage of eminent men in congress, freedom from restriction on divorce, extent of polygomy in marriage customs.[108]

It should be remembered that at the outset Cattell insisted there would be gross errors involved in the measurements of these variables but he felt that the common factors would still be indicative. In factor analysis there is not the need for obvious relationship between the test variables and some preconceived theory system so essential in many other fields. It is the configuration of variables which is of primary importance.

Without going into a description of the actual processes of the analysis of factors it can be noted that Cattell found his analysis approached several criteria

---

108 <u>Ibid.</u>, p. 445.

which give some indication of a set of meaningful factors. He located twelve meaningful factors from the seventy-two variables on sixty-nine nations. The importance of factor analysis is especially pointed up in this study. When we think of the great scholars who have attempted to account only for the fundamental dimensions of Western society or Asiatic society or some other society we gain some insight into the scope of this investigation. Such a task would be utterly impossible without the use of this powerful technique. Some of his factors were:

1. Size. There was clearly a factor of size of the country.

3. Enlightened--affluent rationalism versus narrow--poverty.

5. Emancipated urban rationalism versus unsophisticated stability.

8. Bourgeoisie philistinism versus reckless Bohemianism.

10. Fastidiousness versus forcefulness.[109]

Cattell says time and again that these are very, very tentative and tenuous common factors but nonetheless they represent a beginning in what ultimately could be an empirical science of politics. If one can accept the

---

109 *Ibid.*, pp. 458-68.

present limitations in content, the conclusions here are of great importance. Even if one rejects the assumption of using groups in place of individuals it must be admitted that as soon as we can work out the problem, this technique has important implications in the future study of administration and political science.

Another study by Cattell is directed at laying out the dimensions of personality. Here the pattern was essentially the same as above. A large number of so-called tests were developed, each containing two or three items (some say that the tests should contain at least 10 items) on a great number of hypotheticated dimensions of personality. These were then analyzed using the above-mentioned R technique. This study points out one of the reasons why factor analysis is not quite able to pull itself up by its own bootstraps. We actually hypotheticate items which will go into our tests, and therefore our results.

Cattell points out that while it is often denied that we can study the mental interior of the individual or his private world, he found nineteen dimensions of this private world which may not be extremely reliable and which may have little or no validity, but which at least represent an attempt which, on the face of it, appears worthwhile. An illustrative list of factors include:

3. Interested and understanding nature.

6. Hard-headed rationalism and intellectual leadership.

7. Independent, adventurous self-sufficiency.

8. Will control and character stability.

12. Masculinity versus feminity.

14. Hysteroid aggressive.

17. Conventional unimaginative annoyance at superiority.

In concluding the discussion of factor analysis as an exploratory technique it should be noted that those who have done the most with this statistical technique are often very critical of its use in exploratory studies. We should probably point out that at this time the results of factor analytic technique have only a little more weight than the results of case studies or any of the other qualitative techniques. It appears, however, that as this statistical method becomes perfected it will gain increasing importance as an exploratory technique because of the tremendous power resulting from the first experimental method for determining categories with known interrelationships.

<u>Sociometric techniques</u>. Like the psychometric methods, few sociometric techniques are appropriate in the exploratory phase of science. It is characteristic of

sociometric techniques that they presuppose a good deal of knowledge about the situation under investigation. Certainly Moreno's techniques are of little use here. In one study reported by Schall an exploratory experimental procedure was imposed upon sociometric technique. Here a sociogram was made of a primary group; the indigenous leader and the isolates were picked out. Then a rumor was started with the indigenous leader and with the isolates to see if there would be any differential in the way the rumor was translated to the other people in the group. The hypothesis was that there would be more of a motive for the isolates to sharpen the rumor than there would be for the leader.

In a study of a whole community a sampling technique was used to develop the status structure. It has generally been assumed that the only way to determine the class structure in a community was to follow the Yankee City model and make a complete case study of every individual in the city. Here, in a village consisting of fifty-five family units, with a population of about two hundred and thirty-five individuals, of whom approximately one-half lived in the village and one-half in the hinterland around it, an attempt was made to build a picture of the class structure on the basis of judges.

The judges were selected so that they would be representative of all of the anticipated major groupings in the community. Characteristics considered were: prestige rank of the judges, ethnic affiliation, age, sex, occupation, participation in organized life of the community, years lived in the community, residence in the village, or in open country.[110] The judges were carefully instructed by the investigator to arrange people in order according to "their popularity" and "their rating in society."[111] The attempt was to have each of the judges rate every one of the individuals in the community. Actually, ninety percent of the people were rated by at least four of the judges and this was considered a sufficient approximation of the goal. It was felt that the rating would be useless without at least four judgments, so only four hundred and eighteen people, representing ninety percent of the fifty-five families, were used in the development of the class structure. It was felt that these four hundred and eighteen people were representative of nearly all of the families in the community and that most of the members of each family would have the same economic status. Therefore, the sample was sufficient.

It was found that in ninety-five percent of the

---

110 Harold F. Kaufman, "Defining Prestige Rank in a Rural Community," Sociometry, 8:200, 1945.

111 Loc. cit.

ratings on these four hundred and eighteen people the judges did not vary significantly from each other. An interesting conclusion was that the percentage distributions of persons into classes formed an almost perfect normal distribution, running from two percent in the extreme classes and twenty-five percent in the middle.[112] All over intercorrelation between the judges rating was .88.

This study is reported here because it is felt that in many psychometric studies we should have a definite concept of the status structure in a large organization if we are going to make interpretations which have real power. This technique would be even more effective in an organizational situation because there would be more probability of judges knowing most of the people who would be rated. The background data for the selection of judges could be obtained directly from a statistical analysis of personnel records, where these were available, and what appears to be a very reliable and perhaps quite valid picture of the status system in an organization could be developed efficiently and extremely economically.

<u>Projective technique</u>. Most experimentalists would not classify projective techniques as quantitative research

---

112 Ibid., p. 202.

methods. They are classified as such here, because in comparison with the case study technique, for instance, they tend to be much more toward the quantitative end of the continuum. As Sargent has said, "The variety and richness of material which the projective methods provide is at once the delight of the clinician and the despair of the experimentalist."[113]

The first assumption upon which projective techniques are designed is that all behavior is expressive of the individual's personality.[114] Working on this assumption, those who designed the projective tests attempted to get a representative sampling of the behavior of the individual reflected by presenting a set of stimuli which embody a representative sample of the individual's behavior. A second assumption underlying the projective technique is that the individual is unable to provide the investigator with much of the description of his personality verbally unless he is aided by some graphic or verbal device. This is largely because of the subconscious source of this

---

[113] Helen Sargent, "Projective Methods: Their Origins, Theory, and Application in Personality Research," <u>Psychological Bulletin</u>, 42:275, 1945.

[114] Annelese Korner, "Theoretical Considerations Concerning the Scope and Limitations of Projective Techniques," <u>Journal of Abnormal and Social Psychology</u>, 45:619, 1950.

information. The basic design of the projective technique is to present the individual with a series of ambiguous stimuli which have been found to represent a sample of any personality in our culture.[115] These are both founded upon the third assumption that none of the behavior evoked by the test is a chance event.[116] For example, if the Thematic Apperception test shows that an individual has an aspiration to be the governor of his state it must be assumed that this aspiration, while it may not be a significant part of the individual's personality, is representative of a constellation of aspirations and not just a capricious aspiration of the moment.

The oldest materials for projective tests are the ink blots.[117] There are some who hold that the ink blots are still the best because they are the most versatile of all the media of projective tests. A second type of material for projective tests is the ambiguous picture, such as that utilized in Thematic Apperception Tests. It is said to be approximately as popular as the Rorschach because it presents a total stimulus to the individual. Other media of the projective technique are: art, often in the form of

---

115 Ibid., p. 620.

116 Ibid., p. 621.

117 Sargent, op. cit., p. 267.

finger painting; clay modeling; drama; completion tests, sometimes called the word-association tests; story telling; drama ranging from puppet shows through psycho-drama; and the use of various types of nonsense syllables.[118] Only the first two types of tests will be discussed here, for the remainder lack the standardization which will give them much utility in public administration research for the present.

Claims upon the uses to which the projective test may be put range from those who feel that it has almost no scientific utility to those who feel that it could well supplant all other types of scientific investigation. Those who are acquainted with the projective technique and have made a careful survey of it seem to come to the conclusion that it has great value as a diagnostic instrument in clinical work, some value as a therapeutic tool in clinical work, and at least limited value as a research device.[119]

As a research device Parsons feels the projective techniques can be used to:

1. Initiate new elements into the theory system.
2. Reformulate elements of the theory system.

---

118 *Ibid.*, p. 268. Also, Talcott Parsons, *Structure of Social Action* (New York: McGraw-Hill Book Co., 1937), p. 175 ff.

119 *Sargent, op. cit.*, pp. 270-1.

3. Deflect unhealthy concepts from being entered into the theory system.[120]

The use of projective techniques in public administration is restricted largely to the analysis of groups. Parsons holds that the projective techniques are useful in analyzing:

1. The general theme of a group (that is, the general objective around which it is organized).

2. The dominant tone of a group (passive, cooperative, aggressive, etc.).

3. The means of operation of a group (full participation, autocratic control).

4. The environmental effect upon the person.

5. The person's effect upon the environment (that is, the group).[121]

An experiment was run where a sociologist used an intensive interview and a psychologist used the Thematic Appreception Test in the analysis of the same group and it was found that the Thematic Apperception Test was much more useful as an exploratory device, for the interview only provided a picture of what existed while the T. A. T. provided this and also some indication of the origin and

---

[120] Parsons, _op_. _cit_., p. 176.

[121] _Ibid_., p. 179.

therefore the reason why the group had come about. On the other hand, the T. A. T. had a big disadvantage in that it only told you about the individual in relation to the total situation and it did not give you a picture of the relationship between individual members of the group. For example, where the interview showed that a person typically rebelled on the job and had trouble getting along with the supervisor, the T. A. T. showed that he had an inherent parental aggression and therefore inherent resistance to authority, but it did not show the direction of this resistance in the supervisorial situation. This probably implies that in public administration research we should attempt to supplement the one with the other whenever possible.

The Rorschach has been administered to many, many different types of individuals; superior individuals, people with low intelligence, average intelligence, average children, pre-school children, college students, depressed adults, stable and unstable individuals, etc. It has also been administered to almost every known type of abnormal personality. While it is a long way from being an acceptable research instrument in the final stages, Parsons holds that it is a useful exploratory instrument.[122] The importance of the Rorschach having been administered to so many

---

[122] Hertz, _op. cit._, p. 549.

different types of individuals is that it is standardized for all of these types of individuals: that is, we know how a given type of individual will normally respond and therefore we are able to work backwards and classify individuals we are dealing with. Indiscriminate application of the Rorschach in its present state is hazardous, but where we stick to groups where standardized patterns are known, it is undoubtedly a useful research technique. It is certain that the use of the Rorschach in public administration research would be restrictive but it also might be invaluable. The explanation of antagonism and cooperation between administrators, especially in the higher levels in administration, might be one of the areas where this would be useful. When combined with the interview procedure it seems that this projective technique might also be useful in the selection of people who will be compatible with a given work group. It might also be useful in the analysis of matching personal and formal organizational goals to achieve optimum motivation.

This concludes the discussion of exploratory techniques. The extensive list of techniques which has been discussed in this chapter should be sufficient evidence to convince any layman that the exploratory phase of research is of great importance. In the writer's opinion, it is certainly of equal importance with either of the other two stages of research and it is certainly the most exciting and at the same time, the one which involves the most risk. It is the phase which allows the greatest play of the creative ability of the researcher, and at the same time the phase which is the most unrewarding in scientific results. It is a phase of research which has been largely overlooked by the true experimentalist.

III. IDENTIFICATION TECHNIQUES

A second step in the exploratory phase is the closer identification of variables we have suggested. It is not enough to hypothecate variables. We must test them to see that they represent real and meaningful dimensions of behavior. We must fit them into our theory system and make them consistent with the rest of the concepts and propositions in our system. This may involve only a redefinition of the new concept or it may involve a complete realignment of the whole theory system, for as pointed out in Chapter III, this is precisely the point at which science becomes self-corrective. Regardless of how the investigator "feels" about a new hypothesis, he must make a careful examination of the manner in which it fits with his theory system. The moment for a wallop of paint and a splash of turpentine is past by the time we reach the second stage of a scientific method. Having filled in the broad outlines of our painting, this is the time to step back and make a general evaluation of the degree to which we have actually represented the behavior we are observing. There is no place for inspiration and emotion in this phase. Success here depends upon cold objectivity and the mastery of the mechanics of knowledge-making.

It is probable that since we used our theory system as the basis for the orientation of our exploratory techniques, there is some integration of the new concept with the theory system. The development of reliable knowledge requires that we know what the degree of integration is.

This phase of scientific research might well be called the selection stage for it involves the choice of concepts which provide maximum analytical power. The fundamental problem which all science faces is the development of concepts which must accurately represent the interaction of variables in the situation under study. It makes no difference whether one is speaking of the physical or the social sciences. In no science is it possible to represent all of the aspects of the whole situation. Reliable knowledge would be denied us with our present techniques if we had to represent every aspect of the whole situation to gain control. It is the universal practice of scientists to locate those variables which will account for the behavior which we wish to understand to a given degree of predictability--say, 80% or 95%. The other 20% or 5% may involve hundreds of variables but scientists overlook these because for practical purposes of control 80% or 90% predictability is sufficient.

That is just as true in the physical sciences as it is in the social sciences, and it is pointed up by the fact that in building a bridge, for example, engineers always allow a 15 to 20 safety factor. That is, they anticipate the load and the stresses from weather which the bridge will have to bear and then they build the bridge 15 to 20 times as strong as this estimation. Part of this is, of course, because they cannot anticipate weather, but the major reason is that they cannot completely anticipate the action of the materials which are put into the bridge. By building it at a safety factor of 15 to 20, the engineer reduces the possibility that some combination of variables which he does not know about, lying in the 5% to 20% unknown area, will react to cause the bridge to fall. Thousands of bridges are built every year, all over the world using this principle and almost none of them fall.

The insurance companies are so satisfied with the reliability of scientific knowledge that we have in building bridges that they insure bridges for hundreds of thousands of dollars. But once in a while a bridge does collapse through that one-in-a-thousand combination of unknown variables which is powerful enough to overcome even the greatest safety factor. When this happens, further time and research efforts are expended in the exploration of the unknown area until variables which appear to account for the failure of the bridge are located. These are then integrated with the theory system.

In the social sciences we are still attempting to develop theory systems which give us even thirty to fifty percent predictability and we are therefore much more open to realignment of our whole theory system. One way of saying this is to say that we are not yet certain of even the most important dimensions of behavior. Almost every year exploratory studies turn up new concepts which appear to have as much or more interpretative power than the concepts we have tentatively placed in our theory system. The problem is to evaluate the alternative concepts attempting to fit them into a complete social science theory system. It is this problem that we are facing in the section on identification techniques.

After we have tentatively settled upon a concept as the most appropriate for measuring a given aspect of behavior, our problem is to find out just exactly what behavior this variable actually represents. This is the problem of validation. Theoretically, with a completely valid concept we would know exactly what behavior the concepts represented and exactly what behavior they did not represent and the relationship between the behavior represented and the rest of the behavior in the total situation. To do this we must have concepts which represent all of the pertinent aspects of each of the parts of the study situation. In the end, the determination of which aspects of the parts of the situation must be represented by a concept to account for all of the variance and the total concept of what the situation involves become the primary problem of scientific endeavor.

In the development of a theory system for public administration it is apparent that the first great problem is the laying down of boundaries of administration, so that we can differentiate—it from other processes in our culture. When talking in abstracts, it is easy to get the feeling that defining the universe is rather simple, but as soon as we begin talking in terms of a specific universe it becomes obvious that this is almost a never-ending process, for as soon as we limit the area and begin investigating it we gain knowledge which helps us to make a redefinition of our boundaries.

Secondly, administration, as many students of the subject point out, is a total life process. In fact, it is one of the several total life processes which account for the total cultural processes in our society. We could not, over a period of thousands of years, using hundreds of thousands of scientists and our present scientific tools, create a verbal representation of administration which mirrored every social detail of this cultural process. The question then is, of course, what aspects of administration must we build into a theory system to gain substantial understanding.

Social scientists have worked out some procedures which, though awkward at the moment, allow us to locate boundaries and the identification of independent internal dimensions. One of the means of selecting new concepts is to examine them and find out what assumptions are implicit in them about the nature of the political process. For example, it is held that the basis of motivation is the matching of individual with organizational goals. We might examine this statement. The statement implies that the main determining factor in the expenditure of individual energy is the individual goal. It also implies that organizations are essentially groups and that individual goals and group goals are comparable. It is further implicit in this statement that the primary unit of an organization is not a small work group or even the work team but rather the individual. Now each of these assumptions could be checked against the common set of assumptions underlying the rest of our theory system.

Another means of evaluating a proposed concept for the theory system is to check what it measures, and especially what it measures in relation to the rest of the concepts in our theory system. A frequently mentioned concept of late is that of identification with the organization by each and the total ego involvement of the individual with the organization is said to be made up of his identification with his own work unit and then with the division of the organization of which his work unit is a part, then with the whole company, corporation or agency of which that division is a part. One of the subtle problems that we meet is whether or not identification with one's primary work group is naturally similar to identification with one's company or corporation. Can we actually say that the two processes are essentially the same, only different sizes of groups are involved, or is there a qualitative difference between the ego involvement with the smaller group and that with the larger? In short, do we actually have the same dimension of the organization or do we have two dimensions which, though closely related, are independent; independent in the sense that they account for different types of variance? One of the means of determining this would be to measure the relationship of identification with the primary work group with the individual production and measure the relationship between corporation identification with the same production. Through statistical analysis we might then be able to state whether or not the two types of identification were scalable on the same scale. That is, whether they actually measured the same dimension or whether the two dimensions were slightly different.

To have scientific utility, our concepts must measure dimensions of the same universe; that is, our concepts in a public administration theory system would all have to measure the cultural phenomena which we had delimited as administration. It is impossible to determine this merely by observation. If our knowledge is to be reliable, we must have some basis for knowing that a measure of identification with the work group is actually a measure of administration rather than a measure of some other cultural process such as the family. Note that while the first means of evaluating a concept is a logical process, this one is largely a statistical process.

A third means of evaluating an exploratory concept is an analysis of the degree to which it embodies all of the pertinent details of the part of the total situation it is supposed to be representing. Assuming that we have successfully set down the boundaries of the universe administration, and then that we have made our sum of most of the dimensions of this universe of variables, the question then is, whether or not a given variable represents as many of the pertinent details as possible of the segment of the total universe it represents; or, another way of stating this would be how representative of the given area of the total situation is our concept? Can we find another concept which will do the job better? For example, many theorists in public administration have defined communication in terms of the operation of handing down formal directives, participation in formal conferences and the like, complete disregarding the actual interpersonal interaction which takes place during the coffee hour, during the lunch hour or over the desk with the person who works next to you during working time.

Having delimited the area which we will call administration, and then having broken this area down into leadership, communication, formal division of labor, internal controls, coordination, motivation and so forth, we are very likely to find that after we have defined each of these concepts that there will be a great deal of variance unaccounted for if we accept the formal definition of communication as the referent for this concept. In this case we must reexamine the concept during the early exploratory studies in terms of the above two evaluations and in terms of the individual insight of the investigator, and add to it all the variance included in informal communication. Subsequent investigation may indicate that informal and formal communication are actually scalable on the same scale; that is, these two types of communication represent only different hues of the same color. But we have to know this with some confidence.

There is but a single problem--the matching of our concepts with a defined segment of the total phenomena under study and at the same time integrating our concepts with the rest of the theory system in terms of assumptions and independent dimensions. It is an arbitrary decision, but in this discussion we are calling parts of this process by different names for ease of discussion. The first stages of this process are called exploration and the later stages where very accurate measures are developed and where specific relationships are "nailed down," the measurement phase. It should be noted that during the selection stage one of the things we do is to determine whether or not the factors that we have hypothecated should be measured and therefore handled in the measurement stage. It must be noted, and this is extremely important, that these processes are not hard and fast procedures of a standardized nature. If they were, social science would be much more science and much less art than it is.

The first process, that of examining the assumptions involved in a concept and comparing them with our assumptions about the universe, is a purely logical process and we have no assurance that logical processes will inevitably provide us with objective answers.

In the next procedures, those of selecting a concept, we are able to use some mechanical aids, but these mechanical aids are not at all dependable. For example, just one of the problems in using the statistical technique is the grave question of whether mathematical assumptions underlying statistical techniques can be applied to a social situation such as administration where the group dimension is of equal importance with the individual dimension. One of the really important problems in any successful social research is working out the compromise between the use of rigorous methodology and producing results. In almost every social investigation we could spend the whole budget on perfecting one or two techniques for measuring the phenomena in which we are interested. Almost by definition, if we actually produce results we have more than likely made many methodological compromises. This is not a damning criticism of social science but rather a proclamation that social science is now established upon such a sound foundation that it can stand to frankly admit its limitations.

The discussion immediately to follow is largely introductory, for the statistical devices play the most important roles in the measurement phase. There are some experimental procedures which have been developed which are applicable to problems of identification. Polls and surveys may be used to determine the extent of a given variable in the administrative universe, for instance. Sociometric techniques may be used to get some idea on the various types of groups which exist. These and other techniques of more precise exploration will be discussed here.

Polls and surveys. Polls and surveys are fast maturing and have now grown so that they are well recognized as a commercial instrument.[123] Although limited to almost a single research function, it will probably be necessary to use polls and surveys in defining the contours between bench marks in administrative theory. The substantive limitation on polls and surveys is that they provide us with data only on affective aspects of personality.[124] This is not a reason for denying the utility of the poll or survey but only a drastic restriction upon it.

Harper cites the study of LaPiere where a mail questionnaire was sent to a sample of hotel and motel

---

[123] Calvin F. Schmid, "The Measurement of Public Opinion," Sociology and Social Research, 34:84, November-December, 1949.

[124] Robert A. Harper, "The Present Status of Questionnaire-Derived Opinion Data," Social Forces, 25:294, October-May, 1946-47.

operators asking them their opinions toward Chinese. Those who expressed a positive feeling toward lodging a Chinese in their establishment were then confronted with a Chinese student who was traveling with the investigator, to see whether they would actually accept Chinese persons. The correlation between the response to the questionnaire and the response to the actual Chinese individual was very, very low, and it was held that this was evidence that the attitude survey did not represent potential behavior but only an affective response.[125]

However, in another study it was found that the correlation between responses on an attitude study concerning beer preferences and actual selection of brands of beer was .76 which is substantial by any standards. One explanation which is offered for this disparity is that the attitude survey can present a situation which may measure actual potential behavior responses but the stimulus offered the subjects may be so limited that it does not correspond sufficiently with the total social situation and therefore presents a low correlation with it. Harper's conclusion is that while the question is not settled one way or the other at the moment, it is probably safest to hold that we are measuring mostly affective responses in using

---

[125] _Ibid._, p. 294.

questionnaire-derived data.[126]

Of equal importance to the political scientist is the substantive limitation of the methodological problems in using the questionnaire. It is held that questionnaires are appropriately used after we have some knowledge about the situation under study but before we really can be confident as to what the true dimensions involved are. The use of closed end questions in surveys limits the variety of responses which the subject can give and thus hides some of the information that we are searching for.[127] Of greater methodological importance is the fact that it is virtually impossible to measure the validity of attitude data.[128] Since this is an identification technique rather than a validation technique, we can tolerate more doubt about validity here than we ever could in the third stage of research but there is a question whether the use of polling is desirable when we could just as well use tests with low validity and at least with a test we have the potentiality of increasing validity where with polls we have to perfect a whole technique before this is possible.

---

126 Ibid., p. 295.

127 Ibid., p. 294.

128 Douglas W. Bray, "The Prediction of Behavior from the Attitude Scale," Journal of Abnormal and Social Psychology, 45:82-84, 1950.

There is now general agreement as to the major steps in the opinion-measurement process. The first step in opinion research is the formulation of the problem.[129] This step is broken down by Blankenship into four phases; the first is situational analysis. It involves setting down all of the reliable knowledge and hypotheses concerning the situation which is going to be investigated.[130] Next an informal investigation is made at the site, or at least at representative sites, using undirected interviews with people and talking to people who have a broad knowledge about the situation. The major function of this phase is to get the perceptions of the people in the sample so that one has some idea of the way the questions should be phrased and the type of questions which will be useful.[131]

With this background, one sets down the specific objective of the investigation with enough precision to indicate the actual steps which will be needed. Following this, plans should be made for sample designs, questionnaire development, pre-tests, actual administration of the questionnaire and analysis. As a check upon possible bias,

---

[129] Schmid, <u>op. cit.</u>, p. 84.

[130] Albert B. Blankenship, <u>Consumer and Opinion Research</u> (New York: Harper & Bros., 1943), p. 32.

[131] <u>Ibid.</u>, p. 34.

the results which are anticipated should be set down.[132]

The statistical designs are treated elsewhere. The assembly of the questionnaire is a primary step. It involves the selection of the type of questions, multiple choice, true-false, free answer, etc., that is to be used.[133] The opening questions should be simple ones which will develop rapport. Any personal questions should be placed well into the body of the questionnaire, preferably near the end.[134] While it is difficult to determine what the impact of the sequence of questions will be upon the subject, this can be examined rather precisely in the analysis of the pre-test.

The pre-test of the questionnaire is essential. An effective way is to use several alternate forms of questionnaire in the pre-test so that different sequences of questions and different questions can be compared as to the data they provide.[135] Questionnaires should probably be of different lengths and in some cases may even contain different types of questions. In every case, any flaws

---

[132] Ibid., pp. 35-38.

[133] Mildred Partes, Surveys, Polls, and Samples: Practical Procedures (New York: Harper & Bros., 1950), p. 188.

[134] Blankenship, op. cit., p. 77.

[135] Ibid., p. 83.

which show up in the pre-test should be eliminated and pre-testing should be continued until all flaws are eliminated.[136]

The last step is the least complex but the most productive for it is here that we actually collect the data. We have to be concerned with training the interviewer, maintaining controls on the samples and making sure that the questionnaire is administered within the required time span. The analysis of the questionnaire data concludes this step.

Lest anyone feel that the polling technique is nearly perfected, Blankenship lists the following common failures:

1. Questions are biased.
2. Respondents are asked to reply to questions where they have no basis for responses.
3. Pre-testing is inefficient.
4. Surveys do not measure intensity of reaction.
5. Sampling errors are tolerated.
6. Interviewers are incompetent.
7. Sampling design is based upon wrong characteristics.

---

136 Ibid., p. 91.

8. Inappropriate statistical analysis.[137]

There are three basic types of social surveys used today. The first, the personal contact interview; second, the mail questionnaire, and third, the phone contact. These will be discussed in this order.

Probably the most popular technique of polling and certainly the most effective from the research point of view is the personal contact interview. Partes notes that there are five common interview patterns. (1) The interview where we use a schedule of closed end questions; (2) the free-story technique, where the interviewer states the information which he wants and then just listens; (3) the focused interview; (4) the unaided recall, where the individual is asked to remember a specific point, and (5) the aided recall, where the individual is asked to recognize or identify something.[138]

The big limitation on the personal interview contact, from the researcher's point of view, is that it introduces bias. The amount of bias, says Hyman, depends upon (1) the type of respondent, (2) the type of information gathered, (3) the types of interviewers used, and (4) the mode of

---

137 _Ibid._, pp. 215-222.

138 Partes, _op. cit._, pp. 74-75.

questioning.[139] Studies have shown specific sources of error. One study showed that there were significant differences in response when two different methods of recording data were used. In another study it was found that the interviewer, through non-verbal cues, biases the respondent. In one case, completely different answers to the same set of questions were obtained by two interviewers from the same subject, in a two hour period.[140] Other studies have shown that the ethnic background of the individual as perceived by the respondent is important, especially if there are any questions with an ethnic content in the schedule. Further, it was found that the sex of the interviewer, the economic class, the educational background and the political leanings of the interviewers all biased the responses which the subjects give.[141] Another study showed that experienced interviewers tend to get a stereotype of the way individuals respond and soon after the interview opens he will have classed the person into one or another of these stereotypes and their perception of his responses

---

[139] Herbert Hyman, "Problems in the Collection of Opinion-Research Data," American Journal of Sociology, 55: 363, January, 1950.

[140] Ibid., p. 365.

[141] Ibid., p. 366.

is significantly biased because of this.[142]

An interesting new area of attitude study involves the measurement of the intensity of the attitude as well as its direction. Cantril has a double column of questions on each page of a questionnaire. The first column contains the question proper and the second column asks the respondent to check one of 5 categories reporting how strongly or how much concern he feels about this particular point.[143] One of the important implications of this point for a study of public administration is that it gives us a much higher degree of predictability, for when intensity is taken into account we know something about the probability of the person acting upon a given attitude.

The uses of survey and poll techniques are fairly obvious as a tool for the analysis of the environment of administration. Fundamentally, all organizations must achieve goals set by their clientele. If we are to completely understand the organization we must understand the process through which the clientele determines organizational goals. The attitude survey has been used many, many times to study the affective responses people in

---

[142] Ibid., p. 367.

[143] Hadley Cantril, "The Intensity of an Attitude," Journal of Abnormal and Social Psychology, 41:130, 1946.

organizations have. The work of Leighton indicates the value of attitude scales developed for given organizations to take the temperature of these social organizations periodically.

<u>Sociometry</u>. This technique for the description of groups is of considerable importance in public administration research since groups are one of the important dimensions in administration. The first sociometric technique was the sociogram, initiated by Moreno. Its most admirable characteristic as a research technique is its simplicity. Sociograms are constructed out of inter-personal preferences as listed on sheets of paper. The scheme briefly is to ask people in a given group to list those who are their friends or those whom they would like to most work with or talk with, etc. in order of priority and it is usually specified that 1, 2, 3, or 5 names be listed. Then using a series of circles to represent the people making the choices and arrows to represent the choices they made, a diagram which shows the people who are chosen is prepared. The utility of the sociogram is that it points out almost immediately reciprocation in feeling between people in the organization and in this way it identifies the cliques and secondary groups, the most popular people in the group and the people who are isolationists. It is to the credit of Dr. Moreno

that a whole field of research techniques has been erected upon the basic pattern of his sociogram. Those which appear to be applicable to public administration will be discussed below.

A minor variation on the sociogram technique is to gather supplementary data about persons involved in the group situation and attempt to give the sociogram more interpretive power by extending its dimension. In a study reported by Bonney, church affiliation and family size of the individuals was recorded on the slip where the preferences were listed and these were used to segregate the cells on the sociogram in an attempt to determine whether there is some factor in family size which naturally produces preferences. The same thing was attempted for church membership.[144]

While Bonney's study did not provide really important supplementary interpretation, there are many studies where economic status has been determined and it has been found that this does have a real bearing upon the preferences which develop. It is interesting to note that this bio-data is meaningful when used to supplement the sociogram while

---

[144] Merle Bonney, "A Study of Friendship Choices in College in Relation to Church Affiliation, In-church Preferences, Family Size, and Length of Enrollment in College," Journal of Social Psychology, 29:161, 1949.

there has been little success in attempting to correlate this type of data with psychometric techniques. Guilford and Comrey report a study where data were not effective in predicting administrative ability,[145] and Hadley reports that bio-data was not satisfactory in predicting success in a radio training school.[146] It appears to be the consensus of opinion of the psychometricians that bio-data is of little use in predicting in this area because of the precision of psychometric techniques. The studies of Bonney and others seem to point to the value of using this technique in sociometric studies.

A variation on the sociogram technique was reported by Jacobs. Instead of submitting a blank piece of paper to the subjects, they were given a paper with a place for their names in the upper left hand corner, and the date and code number at the right, and then four columns. The first column contained the names of all the people in the immediate work group, the second column they could check whether they were friendly with the person and the middle

---

[145] J. P. Guilford and Andrew L. Comrey, "Predictions of Proficiency of Administrative Personnel from Personal-History Data," Educational and Psychological Measurement, 8:295, 1948.

[146] John M. Hadley, "The Relation of Personal Data to Achievement in Radio Training School," Psychological Bulletin, 41:62, 1944.

columns provided a place for a check to register indifference, and the last column provided a place to register an "unfriendly" attitude. Then there was a final column where they could list their reasons for acceptance, indifference or rejection.[147]

It was found in the study of a clerical poll that a great number of girls checking the indifferent column said they did not know the person. This has a similarity to the forced choice technique where every girl is supposed to make some response to every other girl. The large number of checks in the indifferent column were followed by the comment that "I don't know." When the data on the first test was supplemented with a straightforward sociogram, we get both the structure of the group and the affective feelings of each of the girls toward the other members of the group.[148] The importance of such an instrument for research is that we should expect to find certain affective patterns and certain structural patterns typical of effective organization for particular types of functions.[149]

The target diagram is an important type of sociogram.

---

147 John H. Jacobs, "The Application of Sociometry to Industry," Sociometry, 8:186, 1945.

148 Ibid., p. 197.

149 Ibid., p. 188.

Here we use the customary approach of classifying people on a basis of some sort of preference but in addition we get background information about the individual who is doing the listing. In presenting our data by the use of circles and arrows, we use a large circle, hence the name, target. All of the cells are placed inside of this large circle and then the large circle may be divided by a diameter running vertically on the paper so that all the male cells are on the left side and the female on the right. Then these halves of the circle can be again subdivided by radii which show ethnic background, religious background, educational background, etc. If in addition to this we superimpose horizontal lines which show economic class of the individuals, the target diagram simultaneously presents graphically the structure of sub-groups in the larger group, the status of sub-groups, the basis of sub-groups, the lines of informal authority and the isolationists.

It is often held that the sociometric technique have such a high subjective content that we can hardly give them credence as objective techniques. In a very interesting study factor analysis techniques were used to check upon the factors measured by sociograms. The specific problem investigated here was to determine whether so-called buddy-ratings actually measured leadership or popularity, authority or merely sociability and the ability to get

TARGET DIAGRAM

OF FIRST AND SECOND CHOICES

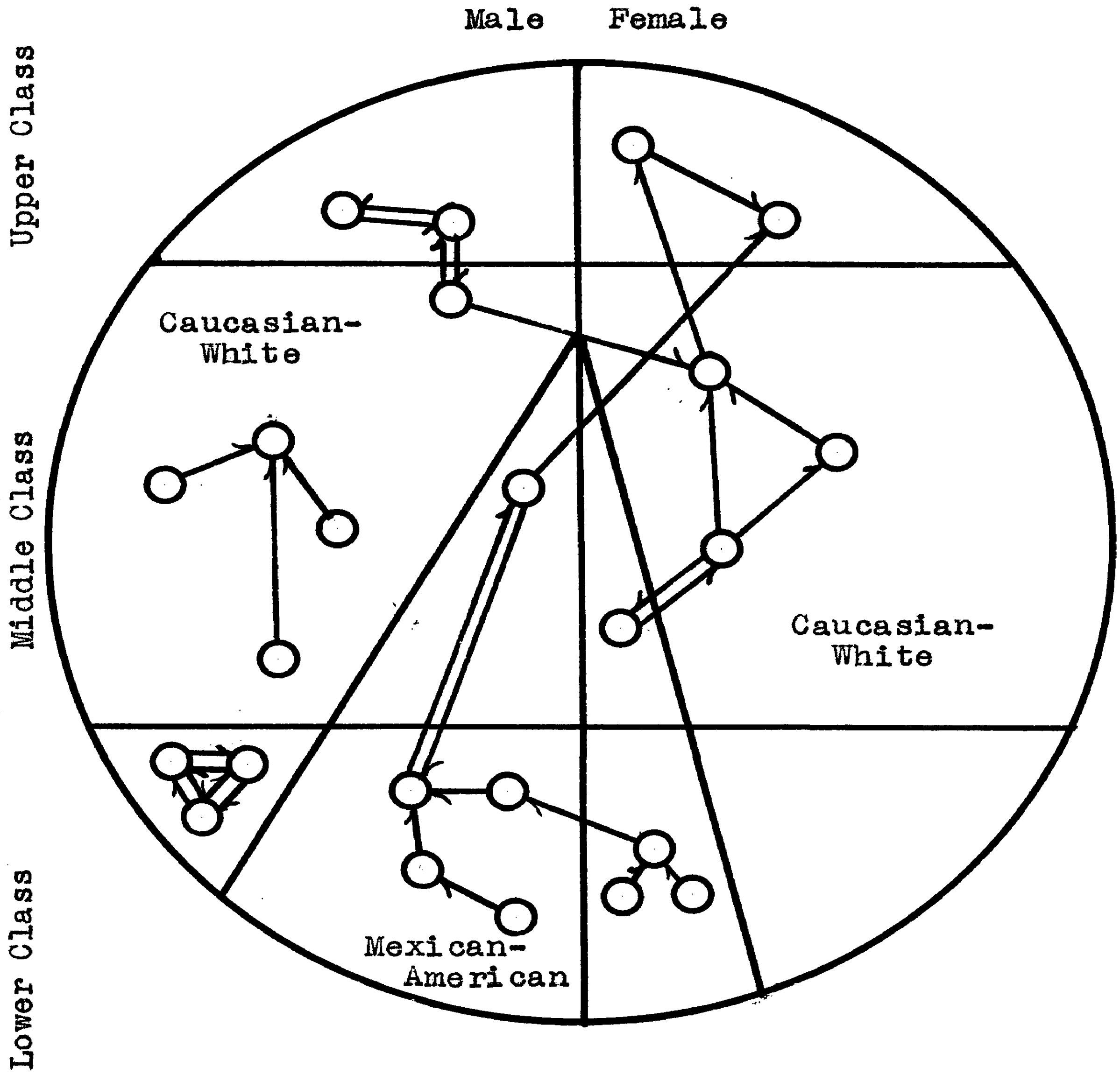

FIGURE 5

along. The groups studied were officer candidates at a signal corps school and the study was continued over a period of six months. The design of this experiment was to take eleven measures of behavior which might be considered leadership behavior. The eleven measures taken were:

1. Anonymous nominations by students for their own section alone.

2. Anonymous nominations by students for the whole class.

3. Average leadership rating, by students for the other members of their own section.

4. Average leadership ratings by junior tactical officers for the whole class.

5. Leadership ratings by the senior tactical officers for the whole class.

6. Anonymous nominations by junior tactical officers for each section.

7. Anonymous nominations by tactical officers for the whole class.

8. Officer efficiency report by tactical officers. (This is a formal performance rating).

9. Average leadership rating by academic instructors for the whole class.

10. Average academic grades given by all academic instructors.

11. Standard intelligence test scores.[150]

The following factors were located:

First, a factor called "academic standard." It had high loading on variables 9, 10, and 11, which is very natural.

Second, leadership factor had a high loading on variables 1, 2, and 3 and on 4, 5, 6, 7, and 8 after the officers had been in contact with the class for four months.

Factor 3 was called "tactical standing," which had high loadings on 6 and 7 and significant loadings on 4, 5, and 8 also.[151] It should be noted that the conclusions obtained here are jeopardized by certain aspects of the factor analysis design. First, there are probably too few variables for the number of factors which were found. Actually, a fourth, although indeterminate, factor was projected and this would mean there is hardly more than two variables per factor and the safe minimum number of variables should be three per factor. Also, the number of people involved in the signal corps school was none too high. One of the interesting techniques they developed was a slightly different rating score. Where a number of

---

150 Robert J. Wherry and Douglas H. Fryer, "Buddy Ratings: Popularity Contests, Leadership Criterion?" Sociometry, 12:180-182, 1949.

151 Ibid., pp. 184-185.

people were familiar with an individual, they were all asked to rate him either good or poor and the number of poor ratings were subtracted from the number of good ratings and this was divided by the total number of ratings made of the individual. This score was found to be an extremely stable means of establishing a rank order when compared with all of the other eleven measures that they took.[152]

The conclusion of this study is that the factor analysis of the eleven ratings indicates that if there is a popularity factor it is essentially the same as the leadership factor anyway, for these eleven measures reflected only one social dimension. The tactical standing factor is a technical proficiency factor and the academic standing, of course, is an intelligence type of ability factor. This is certainly a very interesting attempt to validate the sociometric technique and until other variables are developed to supplement the eleven used here, it is rather convincing proof that the charge that the sociogram measures sociability instead of informal power, is not well founded. It appears that the stars in a sociogram are the leaders and also the most popular.

Stoghill has developed a sociometric technique based on the presupposition that we can get a more stable and

---

152 Ibid., p. 180.

meaningful measurement of group structure and membership by measuring the amount of time people spend working with others in the work situation because people desire to spend their time with those who enhance their feeling of personal worth and give them a feeling of accomplishment and satisfaction and of belongingness as well as assisting them in the execution of their responsibilities.[153] Therefore, if we ask people to pick out those people whom they spend most of their working time with we will get a more reliable measure of the group aspects of the situation than we would merely by asking for friendship preferences or some other type of preferences.[154]

The method is to ask the respondent to select a typical distribution of his time from a typical day in that week. He is then asked to list 5 names in descending order on the basis of the amount of time spent and describe briefly the business transacted during each contact. It was found that when the individual was asked to make two lists, one of contacts with subordinates and the other of contacts with superiors, reliability of the list was increased. Sociometric measures of leadership are correlated

---

[153] Ralph M. Stoghill, "The Sociometry of Working Relationship in Formal Organization," Sociometry, 12:277, 1949.

[154] *Loc. cit.*

with a number of other measures and these provide rather interesting insight into the characteristics of leaders in the population study.

Some of the results of the study are given to show the interpretative value of the technique. (In the following list a "positive" means that there were a substantial positive correlation between the leadership status of the person and the operation measured, while "negative" means there was a minus correlation and "indeterminate" means there was an insignificant correlation).

Level in organization, highly positive; military rank, substantially positive; methods with subordinates, positive; responsibility, indeterminate; authority, indeterminate; delegation, indeterminate; time spent in inspection, negative; time spent in research, substantially negative; time spent in planning, positive; time spent in evaluation, indeterminate; time spent in supervision, indeterminate; time spent with superiors, indeterminate; time spent with subordinates, indeterminate; time spent in conference, positive; time spent in answering mail, positive; time spent in writing reports, substantially negative. The author's comment is:

> These findings indicate that those who occupy high level positions in the organizational structure may generally be expected to have some status as leaders. However, the correlation coefficients reported above are not so high as to suggest that

> leadership status is entirely a matter of location in the organization scale.[155]

To the writer the only possible conclusion we could make here is that Stoghill's assumption is unfounded, for the majority of research in this area controverts the findings of this study. The other explanation is that Stoghill has a valid measure but it is not a measure of leadership as he believes. The highly positive correlation on military rank and level in the organization are two measures of essentially the same thing, while time spent in conferences and in writing reports are typical activities of those in the upper levels of the organization. This leads the writer to submit that Stoghill actually was measuring formal status while he holds that he was measuring leadership and the two are evidently not related, if we can put any confidence in the correlations that Stoghill derives. If this is a measure of formal status it is all the more important, for the only measure of formal status that we have at the moment is pay title or job title and position on the organization chart, and these are only crude measures at best. Research on formal versus informal organization will require a reliable measure of formal organization and this may be it.

---

155 *Ibid.*, p. 281.

Stoghill has another very interesting technique. He imposes a traditional sociogram on an adapted formal organization chart so that lines of authority are shown in one color and choice arrows are in another, and this makes it immediately apparent, graphically, where the loci of actual power are and where those with formal responsibility do not have the real power to execute their duties.[156]

The writer sees important research possibilities in this pattern of Stoghill. Using the near-sociometric technique (merely a standard sociogram approach where we have the person list the reason for their preference behind the name of each person preferred) it would seem that we could locate a number of reasons for indigenous leadership and failure of formal appointed leaders, by combining the explanations of those who choose indigenous leaders into a sort of thumb-nail profile of the pertinent characteristics of the indigenous leaders in the work situation. Another variation of this technique might show us some interesting things about patterns of coordination and control. The patterns of first choices and second and third choices on sociograms are often significantly different. It would seem to be a tenable hypothesis that the pattern of first and second choices indicates leadership in the work

---

156 *Ibid.*, p. 282.

situation while the pattern on third, fourth and fifth choices might well indicate coordinative relationships. It is typical, for instance, that on the third, fourth and fifth choice chart there will be a great deal more horizontal selection than on the first and second choice chart. This is especially true as one descends in the status of individuals. Using a near-sociometric analysis, it seems that we might gain some very interesting insights here.

In a sociometric study reported by an anthropologist, a series of questions were given and a series of preferences were then received. In the first section the respondent was asked to submit a name after each of the following characteristics from his college class: most likable, easiest to know, has few intimate friends, best dressed, best groomed, most unapproachable, somewhat snobbish, outstanding campus leader, somewhat crude, unpolished. The second section told the individual that four people would be sent to an inter-college conference to represent their school there, and their school would therefore be judged by the representatives they sent. They were asked to list the four people they would nominate.[157] Obviously, some of the characteristics listed above are complimentary

---

[157] Orzen Smucker, "Prestige Status Stratification on a College Campus," Applied Anthropology, 3:22, 1944.

or positive and some are deprecatory or negative, and it was arbitrarily decided to give one plus point for each time a person was selected for a positive characteristic and one negative for each time he was selected on a negative characteristic.[158] Total prestige score was calculated for each student, and it was found that out of several hundred students a great number got zero, meaning they were not selected one way or the other for any of these characteristics or functions. These people were said to have medium prestige, and then there were some with significantly low scores and many with significantly high, and these were said to have, respectively, low and high prestige in the student body. The importance of this technique for public administration research is that it gives us still another measure of inter-personal relationships in the work situation and therefore a chance of discovering a new dimension or further identifying one which is already present.

Lundberg and Beazley provide us with a formula for calculating an index of preference. For example, it was found that 37 drama majors out of a population of 253 drama students, gave 43 of their 108 choices to members of the drama department. To get the index of preference we divide the 43 choices within the group, by the 65 choices

---

[158] *Ibid.*, p. 23.

remaining of the 108 which were outside the group. Then we divide this coefficient by a number which represents the probability of 37 drama students out of 253 making 43 choices within their department.[159] It was found that this index was a very stable measure. In terms of public administration research it appears to the writer to be an interesting means of determining work group identification with an agency. This is a factor which is coming more and more to be recognized as having great importance but as yet we have no real measure of it.

A very interesting sociometric technique is reported in a study of rural sociology. Here a sort of stratified sample of judges or raters was made in a whole community. Background data were collected which indicated all of the major groups in the community and then individuals were chosen so that members of all these groups would be represented. Thirteen judges were selected, and they were each given a complete list of all the adults in the community and asked to rate them on the basis of their prestige in the community. Reliability was computed by measuring the agreement of judges who rated the same people, and it was found to be substantial, although there were significant

---

[159] George A. Lundberg and Virginia Beazley, "Consciousness of Kind in a College Population," Sociometry, 11:61, 1948.

differences between the degree to which judges could make fine discrimination.[160]

They found that in the end they were able to make a rather detailed description of the status system in this community on the basis of these prestige ratings by the stratified sample of judges.[161] Where we do not need absolute precision, the efficiency and economy of this technique as compared with the complete census taken in the Yankee City studies is impressive and certainly appears to be a worthwhile technique.

Kaufman feels this scheme could be used well by administrators and others to study whole communities in terms of the predisposition to act and the type of appeal that will move to action for these often differ for persons in these different groups.[162] The writer certainly agrees with this. Steward reports a study which accomplished essentially the same thing in a community with a population of six thousand, where it would have been virtually impossible to have raters stratify everyone in the community, and so a sample was taken. He reports an apparently valid

---

[160] Harold F. Kaufman, "Members of a Rural Community as Judges of Prestige Rank," Sociometry, 9:72, 1946.

[161] Ibid., p. 81.

[162] Ibid., p. 83.

identification of the status system here, using the same technique that Kaufman used only he obtained a sample of the status system.[163]

Where might these techniques be used in public administration research? (1) As measures of morale by determining the homogeneity of individual goals and work group goals. (2) As means of identifying tangible factors of cooperation. (3) Supervisory selection. (4) Recording the individual's progress in adjusting to his job. (5) Assigning people the work group. (6) Locating social factors which affect production. (7) The improvement of management rating devices.[164]

As Deri points out, sociometry and the survey and poll techniques are only three of the many tools of the sociologist.[165] There are undoubtedly techniques for investigating individual differences which should be classified here and are not, and are discussed in the next section, but this is for merely a matter of convenience. It is the writer's impression that of all the techniques

---

[163] Frank A. Steward, "Sociometric Testing at the Adult Level," Sociometry, 9:147-148, 1946.

[164] Jacobs, op. cit., p. 197.

[165] Susan Deri, Dorothy Dinnerstein, John Harding and Albert C. Pepitone, "Techniques for the Diagnosis and Measurement of Intergroup Attitudes and Behavior," Psychological Bulletin, 45:248, May, 1948.

discussed here, the techniques in this section will provide the most useful immediate results because at this time we have background enough to use most of them while we are not yet sophisticated enough to use the psychometric techniques extensively.

CHAPTER VI

MEASUREMENT PROCEDURES

Three major phases of the scientific method have been stipulated: the location of meaningful dimensions of behavior, the definition of these dimensions, and the measurement of behavior in terms of these dimensions to identify relationships between a particular dimension and the rest of the variables in the theory system. The techniques applicable to the first two phases have been discussed. These methods are necessarily subjective in nature. Some social scientists deny these techniques any value.

I. MEASUREMENT OF BEHAVIOR

The point of view taken here is that these two elementary functions of science are distinct enough to be independent. In the activities which serve to initiate the inquiry we are concerned with the location of terms which will allow us to describe new areas of behavior or we are looking for terms which will give us new interpretative power in understanding behavior. Aspects of behavior which are not known are often the dimensions which will account for the errors in our attempts at prediction in the measurement phase. There is no method which will, with mechanical certainty, point up these dimensions. In the end all our present methods depend ultimately upon the imagination and creative prowess of the investigator. Preliminary attempts at definition must be subjective to some degree. This is not an arbitrary limit upon the scientific method but results from the fact that measurement must be based upon knowledge.

Effective measurement is closely related to the conclusions derived in the first two phases in three ways. First, we must locate a unit of measurement which has enough universality to be used as a means of measuring most of the behavior in our area of study. The simplest reason for a common unit of substantial scope is to reduce the number of dimensions that one must understand to comprehend the area of behavior. In dealing with a phenomenon of the complexity of human behavior we must find ways of representing it in a manageable number of constructs if man is to work with the whole at one time because of the psychological limitation upon our ability to hold a large number of ideas in mind at once. The more important reason for a common unit of measure is to allow the comparison of one item of behavior with another and thereby facilitate the location of a relationship between them. One of the fundamental goals of science is to provide a theory system composed of concepts and postulates, which embodies a model establishing the relationships between the elements of the theory system. Real power is gained from effective units. Studies of monotony, for instance, showed little relation between progress in school and boredom but with the more penetrating and universal measure of mental ability made by the Intelligence Test an important relationship was found between these two factors.

Second, measurements which give us interpretative insight depend upon knowledge of the relationships which exist in a situation. One of the reasons we need prior knowledge of the relationships between experimental variables is that this gives some indication of where to look for other relationships. Known relationships are essential in setting up experimental controls for further investigation. If we know the relationship of heredity to personality, for example, we will know whether to include it or leave it out of a study of administrative role. When the phenomenon under study is anything as intricate as the role of an administrator there are literally hundreds of variables which can be suggested which must be investigated and the importance of our findings will be affected a great deal according to whether or not we are able to disregard a constellation of variables as large as those associated with heredity.

A third way in which the measurement of behavior is dependent upon preliminary conclusions is that the measurement of new areas of behavior requires subtle techniques. Often effective instruments are possible only when we know a good deal about the activity we are attempting to measure. This aspect of the measurement process points up the spiral nature of social measurement. On the basis of some fairly reliable knowledge we design the best instrument we can and fit it into an experimental design. This gives us a more precise measure of the behavior we are studying, but on the basis of this new measure we are able to modify the instrument and our next measure is even more precise. This refined measure may be used in other experimental designs to maintain tighter experimental controls. This is one example of the cumulative nature of scientific knowledge mentioned in Chapter III.

The creation of reliable knowledge is a process of successive approximation where each refinement is based upon the step just taken. The importance of this principle is that no technique is independent of the whole sequence of exploration, definition or measurement. Thus, from the time a factor is first projected until the time instruments are developed for its accurate measurement a single instrument may be used, but both the instrument and the concept undergo drastic change during the progression from identification to measurement. The fundamental differences in the methods and the rationale used to execute the exploratory phase and the measurement function of research are so sharp that the layman often perceives exploratory techniques and measurement techniques as two different parts of scientific method. Probably one of the reasons for this is the heavy reliance upon statistics in the latter activity and the importance of logic and other verbal mechanisms in the former. Interviewing, sociograms, and questionnaires are methods that most people can understand with the aid of cursory explanation. Statistical procedures and experimental designs are beyond the comprehension of most who do not have some grounding in mathematics. The aim of this chapter is to outline some measurement procedures and to try to show how these are related to the end product of scientific research, reliable conclusions.

It must be pointed out that making social measurements is quite different from making measurements in the physical sciences. There the methods of measurement have long been agreed upon and the units are described in detail. Factors related to an accurate measure of length, such as temperature, pressure, stress and the like are well known and need not be considered in the use of most instruments because the instrument itself is designed so that it takes these variables into account. In the social sciences we know very little of the processes that we are measuring and we can never be sure that we are not actually measuring the equivalent of some diagonal on the side of a cube instead of its edge. Actually the physical scientist doing basic research faces many of the same problems that the social scientist faces, but the problems all have the basic difference that social phenomena are more dynamic than physical phenomena.

Measurement necessary to understand behavior. It is accepted that measurement is necessary if we are to understand behavior. But by measurement most people mean estimation. They anticipate that the process of social measurement should be similar in complexity to taking a gross measurement of length with a yardstick. Laymen quite naturally expect that if an Honesty Test is administered to a person or a group of persons one should have a measure of honesty. Several characteristics of all social phenomena make necessary a number of statistical controls before it is possible to determine what our measurement is and if it is actually a measurement of honesty.

First, the only means of measuring behavior must be made either by observing what an individual actually does or asking him what he does or would do in a particular situation. Since observation is almost out of the question, we customarily use instruments containing a brief series of questions held to be representative of all the individual's behavior in an area. _If_ the individual gives us sincere answers, _if_ our questions are understood by him to mean the same thing as they meant to us and _if_ the test situation is such that the individual feels about the way he usually does we have some opportunity of getting information which will give us predictive power. Second, _if_ we have a pretty good idea of what the behavior we are measuring is like, _if_ we have a type of behavior which can be described by the respondent, _if_ we have some experience with constructing questions in this area, _then_ we have some chance of developing a questionnaire which provides us with a really representative sample of the individual's total behavior in this area and thus we have some opportunity for predictive power. Finally, _if_ we have a situation for which statistical techniques are available, _if_ we have a stable theory system, and _if_ we have great fortune in construction of our assumptions, _then_ we have fulfilled the conditions of measurement to some degree.

Lack of stable unit of measure. In taking social measurements many things which are obvious in physical measurement are not yet controlled. If one takes a yardstick and sets out to measure the size of a room there is no question as to where the boundary of the room is, but if one attempts to measure the perimeter of a person's honesty there is little agreement upon what one should be measuring. It is the lack of such simple things as a stable unit of measure that introduce additional problems for the social scientist. While temperature might produce certain inconsequential changes in room size the increase of emotional tension may make our yardstick for the measurement of honesty look like it is made of rubber. This sort of thing leads many people to feel that we should have as many measures of honesty as there are classifications of emotional tension. The scientist knows that for predictive power we must be economical in the proliferation of measuring units. For thousands of years men had one set of techniques for measuring and explaining the action of ice, another for water and finally a third for the description of the action of steam. With the discovery that two dimensions, pressure and temperature, were completely sufficient for the explanation of liquids in any state, we eliminated several basic dimensions and made possible operations which took advantage of the new continuum like refrigeration. Because we increased our interpretive power to understand the forms of liquids by discovery of more fundamental dimensions of their behavior we were able to improve the health of some sixty percent of the people in this country by modification of eating habits.

While it would be useful to have some understanding of seventy different forms of honesty, it would be infinitely more important if we were to understand honesty, integrity, trust, faith, confidence and sixty-five other such variables in terms of two or three fundamental dimensions such as ego and role or the like. Since we are never sure where the behavior types that we are investigating end and since we are never quite sure that two activities that appear unrelated are not actually related, our measurements contain a set of procedures which function as a continual check that indicates whether or not our measurement is actually representative of the true situation.

Lack of definite classifications for measurement. Social measurement is further complicated by the fact that few of our instruments tell us for sure what we are measuring. In the physical sciences we have such an inclusive classification of characteristics that a person who is measuring a circular piece of rubber with a hole in the center has no trouble determining whether it is a tire, an inner tube, a rubber washer or a rubber doughnut. Further, if a person knows what make of tire he is measuring he can tell how much nylon cord is in it and how thick the bead wire is and many other things.

In social science we are fortunate when we know enough about a behavior pattern to be able to identify it with confidence and we almost never know enough about it to be able to infer all of its characteristics merely because we are able to classify some of them. Thus, if we knew as much about people as we do about tires, we should be able to infer the personality of an individual who is male, white, married, 45 years old and born in Mexico.

Therefore, we must include some scheme for insuring that we are actually measuring what we think our instrument is measuring. Criteria or scaling procedures are often used for this purpose. Since we are not dealing with manufactured phenomena we usually want to find some way of insuring that our generalizations once determined apply to all of the natural behavior similar to that which we studied in our investigation.

Sampling statistics are used to check upon generality. In our "mass production" society we are used to perceiving standardized products as being exactly the same, whether they are or not. For some reason we are cued to perceive each item of behavior as different from every other item of behavior. Certainly in many cases there are behavior patterns no different from one another than two Chevrolets of the same year and model. There is no significant functional difference between the two Chevrolets. We have learned to disregard the rest of the differences which exist. In our society most people have learned to believe that the great differences between the behavior of individual precludes the finding of any reliable similarities. The social statisticians have come forward with objective means of determining whether or not a set of differences are significant or not and therefore whether they allow us to maintain that there are fundamental similarities between two sets of behavior.

It is for these and perhaps many other reasons that the layman usually feels that when the social scientist speaks of "measurement" he actually means something else, something that no layman can understand. He may even feel that there is or should be some simpler way of measuring. The fact that social phenomena are largely subtle and the fact that they are so complex makes measurement difficult. This, coupled with the immature stage of our theory systems, makes precise measurement impossible and reliable measurement a complex operation. As our theory systems are filled in and as they become more settled we shall be able to control a great deal of variation by assumption which we now must deal with through experimental controls.

Testing theory systems. In the process of taking measurements we are also testing the theory system. From the research point of view this is, in fact, the importance of measurement. While the administrator may wish to know just what the attitudes of his personnel are toward him, the investigator wishes to know what possible patterns of subordinate--superior relationships exist. When we assume a theory structure and then set out to measure a specific variable assuming that it operates in this setting we establish a link between reality and our theory system. It is only through controlled measurement that we can determine such a connection. When we know that a given variable, honesty, for instance, exists, we do not have to assume its existence and we can then make our next measurement in the light of the theory system which has one less assumption. Not only this, we are able to lay down the relationship between a second variable, leadership, for instance, with some detail.

The advantage which we gain as variables are established is cumulative and illustrated by the identification of the primary mental abilities in psychology. As verbal and number ability became well established, the variance that remained in any given test, such as the clerical aptitude tests, could be more easily explored for we knew with confidence that it was not composed of either of these two and probably not of any variable in any way closely related to these.

One variable hypothecated was memory. This was tested and it was found that while it accounted for a small amount of variance, there was still an important segment remaining. Mechanical ability was hypothecated and in conjunction with the other three virtually all of the variance measured by common clerical aptitude tests was explained. It can be easily seen that the location of a factor which we can maintain is of fundamental importance in the analysis of behavior. The only way in which to say with certainty that honesty, as defined in an honesty test, is really related to successful supervision is to make precise measurements of honesty and supervision and see if a relationship appears.

When the layman hears a psychometrician quietly state that certainty is conditional on measurement he hardly sees the implications. He must see them and understand their importance in comparison with the methods of the typical political scientist if he is to see the real difference between science and history or philosophizing or just plain wishful thinking. Gifted political theorists may, without the aid of a method, sift and turn great quantities of detail and hypothecate a variable or a relationship. No man has the mental capacity to establish the existence of this factor or relationship in nature, especially if it is of some generality. Nature is a functioning whole. Measurement devices provide man with a means of observing the whole at once or in a series and singling out some variable or variables for observation while they are interacting in their dynamic setting. In large measure it is the experimental design, and especially the measurement phase, which enables us to do this. Only if one is so naive that he has no realization that the human mind has these limitations can he easily deny the necessity of measurement as the final phase of science.

This fundamental point must be made. Measurement is not a mechanical process. The use of statistical procedures involves a set of assumptions. They are generally assumptions about the degree to which the mathematical presuppositions underlying our statistical procedures are fulfilled by the phenomena under measurement. For example, building upon the theory of probability, many statistical procedures assume that the variable under investigation is so distributed in the population under investigation that if we actually had all of the measurements on this variable we would find that it was distributed in some systematic way, usually according to normality. We can never know whether this assumption is correct until all of the measurements are actually precisely made. This is the type of thing which leads the sophisticated social scientist to hold that every investigation involves a "risk." This concept of risk then comes to be a measure of the criterion of sound experimental work. Any small scale study with a large degree of risk is mediocre while any large study with no risk is probably mediocre too, since it probably involves some large assumptions.

We are now ready to discuss the actual techniques of measurement. The purely statistical techniques will be taken up first since they form the setting for the discussion of experimental designs.

II. STATISTICAL PROCEDURES: DESCRIPTIVE

Statistical procedures are divided into two types: descriptive and sampling. Descriptive statistics are those which describe any continuous variance.[1] In the process of deducing generalizations about nature from observation we must determine observations which represent all of the manifestations which we can imagine. This means that we seldom deal with a single number. If we are measuring the intelligence of all supervisors in government agencies we will have thousands of individual intelligence scores. Our problem is to represent these by some symbol which we can comprehend and so that we can compare. Descriptive statistics enable us to represent complex sets of measurements with a simple statistic and

---

[1] John Gray Peatman, Descriptive and Sampling Statistics (New York: Harper & Bros., 1947), p. 127.

thus describe the whole of a set of measurements in understandable terms.

Sampling statistics are used to examine the relationship between sets of data represented by descriptive statistics. It is this set of techniques that allows us to establish relationship and take the measurements discussed in the introduction to this chapter. The time is approaching when most social scientists will have to at least understand what these measures mean if they are to continue to read the literature. The non-experimental social scientist customarily establishes a relationship by deciding that it exists and once he has made the decision he lets the matter go unless he suddenly becomes convinced that the relationship actually does not exist at all. The experimentalist on the other hand feels that there is no relationship established until he is able to demonstrate it through the use of one or another sampling statistical procedure and then the relationship is only contingent.

Having applied his statistical technique the researcher specifies that there is x probability that a is related to b such that under conditions m, n, o, p, and q, a will precede the event b. At this point we do not know whether it is the set of conditions m, n, o, p, and q, which cause b or whether it is a or whether there is some third as yet unknown cause or whether this is a case of multiple causation. Establishing the nature of the relationship requires a whole new experiment and hundreds of thousands of dollars. This type of progress seems tedious at best to the non-experimentalist. The simple truth is that if we are to utilize the findings of experimentalists we shall have to have some idea of what constitutes sound experimental method so that we can make at least cursory evaluations of the results we are attempting to apply. For many years the person concerned with the application of social science results will be confronted with denial of the validity on the part of those who should be practicing the new theory. If the administrator denies the reasonability of the new concept, political scientists must be able to explain the support behind the theory.

<u>Descriptive statistics</u>. One of the continuing problems of administration is the representation of large aggregates. There is actually only one way to do this. Like parts must be added together to make subclasses and then these must be added together to make larger classes until we end up with three or four different primary classes which represent a great number of small classes and hundreds of individual elements. This is essentially what happens when a governmental agency uses a budget. Salaries for all clerk-typists are grouped together under

one heading, then the clerk-typist category is made a part of a general clerical heading and this finally becomes one of the headings under "personnel costs." No one denies that there are great differences between a clerk-typist who does a great deal of filing and one who spends her time transcribing dictation but these differences must be overlooked if we are ever to get a picture of the total operation. Descriptive statistics give us a most precise method for making these combinations.

The critical question always is what differences can be overlooked and which differences must be reflected in the final set of categories. In a budget document we have no way of knowing this (although we are fairly sure that the categories we now have are inadequate). There are descriptive statistical techniques which can be used to represent any aggregate of things. These are appropriately called measures of central tendency for they determine the points toward which the individual elements in a distribution would tend if they were all to be brought to the center of the distribution. The startling thing about measures of central tendency is that they are a means of representing a whole distribution, one involving hundreds or even thousands of measures.

The description of a given distribution is not really meaningful unless we know how far it varies on either side of its measure of central tendency so a measure of this is provided in the measure of dispersion or variability. If we know that almost all of the measures in a distribution are grouped close to the central point we can feel that we know what the distribution is really like. On the other hand, if the elements of the distribution are greatly dispersed, the measure of central tendency is not as satisfactory a means of representing the whole distribution and we know this.

Finally, we want to know how dependable our measures are. The several coefficients of reliability tell us this. It should be noted once more that these simple descriptive statistical functions achieve their major experimental importance as entries in the more powerful techniques to be outlined below. A frequency distribution is such a simple matter that it is difficult to explain its importance to the intelligent layman for he understands it at a glance. The important thing is that frequency distributions are the basis of the calculation of the mean, correlation coefficients, common factors and many other powerful statistical measures.

Frequency distribution. The frequency distribution is perhaps the simplest of all statistical techniques. It consists simply of a series of intervals with an equal number of steps to each interval and a tabulation of the number of measures falling in each of the intervals. This presentation tells us some very simple things such as the range of variation of the factor we are studying and some indication of the nature of the distribution--that is, are most of the measures in the center with a few at each end or is there about equal density of measures over all the intervals of the range? The frequency distribution represents the beginning of the process of generalization for here we say that for all practical purposes all of the measures that fall in a given interval are similar. That is, the differences between them are less important than the differences between them and those in other intervals, and, of course, this is true, if we consider the quantitative differences.

This is illustrative of the concrete nature of the assumptions that we make in the derivation of statistical generalizations ~~and~~ it illustrates the reason for the superiority of statistically derived generalizations over any developed purely through logical methods, at least as concerns their precision and interpretative power.

Mode. The simplest measure of central tendency is the mode. It is by definition the measure that appears most frequently in all the measures made. Large numbers of

FREQUENCY DISTRIBUTION CHART

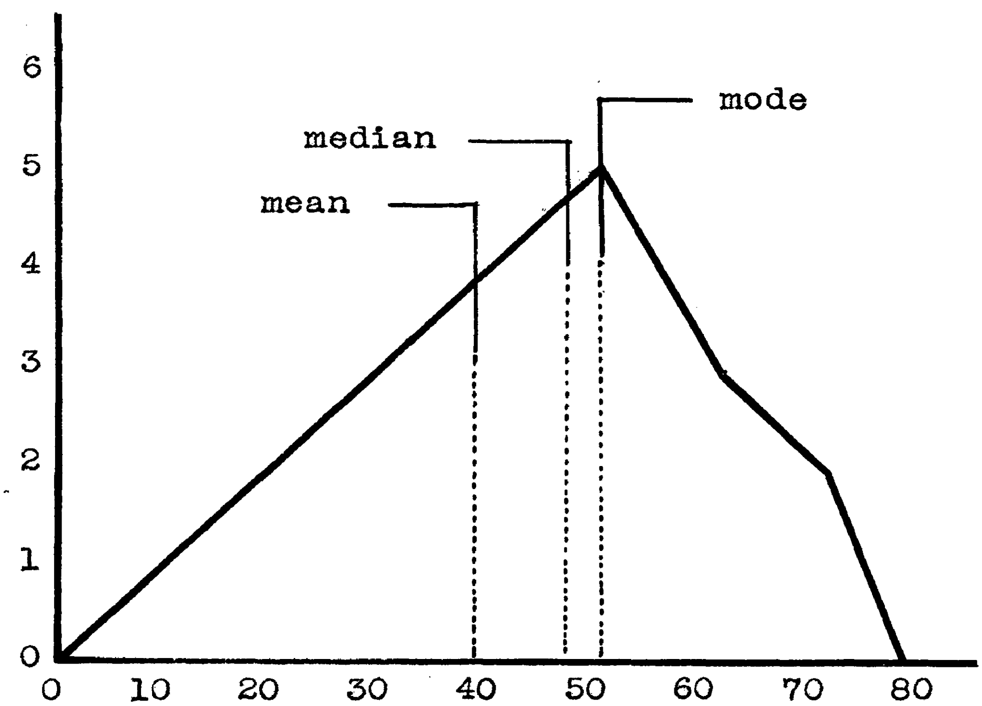

FIGURE 6

investigations have shown that many types of variables tend to operate such that most measures fall in the center of the distribution. This is very reasonable when one stops to consider that most people are about average in a majority of attributes. There are always a few some distance above most and there are a few substantially below the majority. The measurements of intelligence, for example, show that we may expect any group of people to fall near 100 since it is likely that most people are average and this is the average score. The mode is not a very satisfactory measure of central tendency since it is so unstable. Guilford suggests that the mode is useful when we do not have time to calculate the more stable measures of central tendency and when a rough estimate of the most typical case will do.[2]

<u>Median</u>. The median is the measure which represents the center point in the distribution in that it is the half-way point. The median is calculated so that it is the middle point in a distribution. This characteristic of the median gives it a special function. A distribution is said to be skewed when there is a noticeable concentration

---

[2] J. P. Guilford, <u>Fundamental Statistics in Psychology and Education</u> (New York: McGraw-Hill Book Co., 1950), p. 79.

of scores away from the center. When this happens the mode and the mean are poor representatives of the whole distribution. The median is not a very good means of representation, but it is much more representative than either of the others. When the distribution is skewed, when we do not have the time to calculate the mean, or when there is doubt about the equality of the unit of measurement, Guilford suggests the use of the median.[3]

Mean. The mean is second only to the standard deviation in utility. Essentially a simple average, the mean is best described as the center of gravity of a distribution. Taking into account irregularities in the form of the distribution, the mean gives us a measure which is most representative of any distribution which approaches normality. That is, the mean is the best single measure that we have for the description of any large set of measures where the measures reflect variance which is normally distributed. A simple approximation of normality is symmetricity about the mean. One of the most important characteristics of the mean is that it is computed from squared deviations from the center of the distribution and this gives it a particular sensitivity to clusters of

---

[3] *Ibid.*, p. 79.

scores in the distribution. These clusters often represent real dimensions within the variance under investigation and the mean is therefore more responsive to interrelationships than the more gross measures of central tendency. This leads to the mean having more stability as the number of measures is increased. It is this same sensitivity to clusters which makes the mean of little use on distributions which are not symmetrical, for a cluster near the edge of the distribution which skews it will pull the mean so far away from the center of the distribution that it is no longer representative of the whole.

The importance of the mean can hardly be exaggerated as a tool for the exploration of natural phenomena. When we study human behavior, and it matters not which one of the institutions of the culture we investigate, we find that it is typical that natural factors, such as height, weight, intelligence, leadership ability, dominance tendency, neuroticism and psychotic tendencies are distributed among the population from complete absence to absolute saturation. If we are to find out anything about the nature of such traits as these we must have some way of indicating that the general tendency of American males is to grow about five feet ten inches tall without stumbling all over ourselves. The mean is a tool which allows us to make the best possible representation of the different heights that we actually measure in the populace. The layman often feels that the mean is an artificial method which produces something which does not exist. It is certainly true that there are few people exactly five feet ten inches high in our culture in comparison with the whole remainder of the population. But if we are not to become swamped with detail in comparing height with other variables we must have some simple method of representing the quantitative extent of variance of the American male on the variable height without bringing up each individual measurement of height.

Statisticians have surveyed the field of problems and methods and they have determined that in spite of any technique the layman might conceive, the mean is mathematically optimal as a symbol for representing a whole set of measurements. There is a tendency to deny the efficacy of the mean by a person who happens to be aware of the range of variance and feels that every bit of detail that he perceives must be reflected in any statistic which he will accept. The mean is designed to serve a specified set of functions and it is a mathematical certainty that it will satisfy these within the limitations demonstrated by the standard error of the mean. If the layman cannot see the sense in a mean he should attempt to increase his grounding in mathematical functions.

Scientists have been confronted by the complexity of nature and have turned away with a frank admission of defeat in an attempt to get a single symbol to represent nature perfectly. Their modus operandi is to define carefully the characteristics which are pertinent to their problems and proceed to work with only those characteristics. The layman has to come to see that what the scientist sees--that some of the variance he disregards is relative but he must stick with his method if he is ever to get at that variance.

Since the mean is such a fine method of representing a whole distribution it is very important in higher statistical operations where we wish to determine relationship by comparing a single measure. The mean is often coupled with some measure of variation to give a more complete description of a distribution. Thus, we often say that the mean height of American males is seventy inches and the standard deviation of the distribution is three inches. This tells us that the average American male is seventy inches high and that almost none of them are less than sixty-one inches high or more than seventy-nine inches high. We shall have to discuss the properties of measures of variation to explain this inference.

Range of distribution. The simplest measure of variation is the range of a distribution. It tells us only what the smallest and the largest measures are and therefore the limits of variation. If this is all that we need to know, the range is sufficient unless there are some extreme scores on one or the other end of the distribution. Extreme scores have the effect of stretching the range and therefore giving us a mistaken impression of the extent of variation. When we know or suspect that there is a stretching effect in the range we can use what is known as the interquartile range. This is a range measurement from the point separating the top twenty-five percent of the scores from the rest and the point separating the bottom seventy-five percent from the rest of the distribution. This has the effect of eliminating the more unstable extreme scores and providing a more dependable measure. The interquintile range, the interdecile range or any other span might be used for this purpose.

Standard deviation. The most important measure of variation is the standard deviation. This is probably the most widely used tool in statistics and one of the most important. The standard deviation has several uses. When used as a descriptive tool the standard deviation is a measure of variation from the mean. The standard deviation is computed on the same principle as is the mean and so has the same qualities of stability and representativeness. Mathematically, it is the second moment, which means that it is even more stable than the mean. In connection with the normal or bell shaped curve the standard deviation becomes very descriptive. It is a means of determining how much of the distribution lies between any two points.

Again, the layman might counter that these so-called descriptive statistics do not describe the characteristics that he is interested in at all. The writer has heard administrators deny the utility of statistics because what one needed was knowledge of the nature and character of factors and all statistics tells one is that the mean is something and the sigma is something else. These administrators feel that a standard deviation is not a meaningful measure. What these administrators actually want to know is the relationship <u>between</u> the factor they are interested in and some others. Descriptive statistics simply do not establish this type of meaning. A person who has some difficulty understanding the mean or the standard deviation often comes to feel that it must be a powerful technique if he has such difficulty understanding it. The point is that the measures of central tendency and variation are very simple statistical tools which tell us only about the characteristics of a single distribution. The number of things that any one can infer from a single set of observations is limited. The answers that most laymen want from statistics are possible only from the most complex statistical manipulations.

One of the things that informed laymen want to know about a given set of data is whether or not it represents a true relationship. If we make a survey and find that sixty-seven percent of the salesmen in a department are honest, how sure can we be that they are actually honest? Researchers have found out that controlling the conditions of measurement in any situation requires a sophisticated theory system. This is often not available. The statisticians met this problem with what they called a coefficient of reliability.

Reliability coefficient. The coefficient of reliability is based not upon the subjective dimensions of any situation but upon the more universal principle of probability. The statistician denies the necessity for considering variation due to external factors any more than we must put up with undependable measures if someone gives us a rubber foot rule. The solution of the statistician was to find a means of measuring and thus controlling the "stretch" in our measuring instrument. The reliability coefficient is based upon the principle that even using a rule made of rubber thousands of individual measurements of the same stick would provide us with a distribution of measures whose mean tended to approach the true length of the stick. Note that it makes no difference whether one person makes thousands of measurements of a single stick or whether thousands of persons each make a single measure of similar sticks.

The reliability coefficient is designed to tell us what proportion of these measurements represent true measurements and what portion of the measures are the result of error in the measuring instrument, and thus is an undependable part of the measurement. If we think of any measure as being made up of a true component and an error component it is simple to see how the reliability coefficient is nothing more than the proportion of these two components. It is essential that the reliability coefficient be considered in terms of the conditions which actually determine its effectiveness. <u>Any reliability coefficient is meaningful only for a particular instrument in a specific situation</u>. No instrument is completely free from situational determinants. Many tests approach independence from situational determinants however. The Intelligence Tests have been standardized to the point where they are quite reliable in thousands of test situations, though we know that their reliability is lower on persons with limited education and on people nearing and past middle age. Tests that measure less important universals than learning ability are certain to have lower reliability coefficients over the whole population on this principle.

From the administration of any test, then, we produce only an obtained score. The problem is to estimate how near our obtained score is to the true measurement. This requires that we make some estimation of the relationship between the true score and the obtained score. Guilford points out that we can conceive of this relationship easily if we picture a scatter diagram with plots of the true and estimated score for each measure. The standard error of estimate will then be a pair of lines parallel to the regression line.[4] This gives us a very precise description of the standard error of estimate but we have no means of actually computing the true score so we must turn to some other method of deriving the reliability coefficient.

There is general agreement that there are three classes of techniques for obtaining the standard error of the estimate of obtained scores: (1) internal-consistency reliability; (2) alternate-forms reliability or comparable-forms reliability; and (3) retest-reliability or test-

---

[4] *Ibid.*, p. 478.

retest reliability. Actually the first may be thought of as a test of consistency, the second as a coefficient of equivalence and the third as a coefficient of stability.[5] There is no best way of estimating the reliability of a measure. One or more of these methods may be more effective in a given situation however. The one to be selected will depend upon the unique contribution to total reliability which this technique makes.

Even though we locate a fundamental, somewhat universal factor, there are bound to be variations in the score that a particular individual will make because of the relationship between certain external factors and the test ability. Metabolic equilibrium at the time of the test, the incidence of some domestic maladjustment and the like will have some perceptible impact upon the test-retest coefficient for any single individual. Anyone can see that fatigue, improvement in skill in dealing with a new question form, and increase or decrease in motivation and intensity of application would have a real impact upon the internal consistency form of reliability coefficient. The alternative forms measure of reliability is bound to be influenced by the differences in content both because the individual may have chance blind spots in his subject

---

5 *Ibid*., p. 481.

matter construct and because of differences in perception based upon differences in experience. It is probably safe to hold that each of these types of reliability is unique or has a unique component so that a complete measure of reliability would be approximated only if we made these three types of measures.

Actually, there probably is some communality between these measures of reliability and we have an estimate of the multiple estimate of reliability if we take only one of them. Practical considerations being equal we should take the measure which provides us with the unique component which seems most important in the subject matter problems we have to meet. The detailed considerations involved in the selection of a particular method are expertly discussed in some detail in Guilford.[6]

Validity. A second major problem in the use of statistical techniques is that of making sure that our measurements mean what we think they mean. This is a part of the general problem of establishing the epistemic correlation and it is the statistical problem of validity. This is one of the most complex problems of modern science and certainly one that is far from solution. One often hears

[6] *Ibid.*, pp. 481-510.

someone deny the utility of a test because they have heard its validity is low. This is simply not a tenable position. Validity has a very complex relation to the functional effectiveness of a particular test in a certain situation if the test is working satisfactorily. It may have almost no impact upon utility. This is because there are several kinds of validity according to the function or purpose we have in mind.

Classification of validity. The classification of validity here is between practical validity and true validity. A measuring device has practical validity if it provides us with measurements which allow us to control or modify or otherwise work with some social process. Our test for honesty discussed above would have practical validity if it allowed us to select honest clerks or if it aided us in defining honest behavior. This is no indication that the Honesty Test measures some primary natural phenomenon which we identify as an honesty trait. The determination of whether or not a trait which corresponds with the folk idea "honesty" exists is a problem in establishing an epistemic correlation and functional control is no criteria here.

In establishing practical validity we may follow several courses. We may correlate test scores with success on a job after several years have passed. A Safe Driver's Test was validated in this manner by checking to see how many drivers who had accidents were spotted by the test and how many good drivers could have been predicted by the test. Where a test has such limited application it is common to make a cursory reproduction of the job situation in the test and hope that the critical elements have been included. Thus a performance test for bombardiers involved a set of knobs and levers, a test for pilots--a set of levers and pedals, and a test for gunners--a replica of a small machine gun. Here there is little effort to lay down the basic traits of a good gunner or a good pilot or a good driver. A whole population of drivers or pilots is tested (all of the drivers for a single milk distribution firm, for example) and then records are kept of their accidents. A good score in the test is determined by the score pattern of those drivers who had a satisfactory record actually driving a truck on the city streets. There are many predictors which would be used. Usually the effort is to select those predictors which give the maximum predictive efficiency for the least cost. Studies of this type are not designed to locate causal relationships.

<u>Criterion</u> <u>measure</u>. Often some check is made upon the object of the measurement instrument through a second

and independent, though associated, measure called a criterion measure. Here, though we may be very concerned with predictive efficiency, we are also attempting to develop some indication of a theory-reality relationship. The general pattern for doing this is to select some easily observable behavior and get a gross measure of it through some means such as the rating process. There are many who hold that the criterion being an objective measure the test which is validated against it must be a near reproduction of reality if the correlation between the test and the criterion is high. Actually, in any situation there are many criteria that could be used. In most cases any behavior measure which is available is used and this is then defended. If there are several criteria measures to choose from, the selection must surely be made in the light of the theory system so that we should expect to find at least some correlation between a test constructed from the theory system and a criterion selected in the light of it. Where there is only one measure available, we have almost no assurance that the criterion measure is representative of what we think we are looking for.

The criterion technique is certainly important but this writer feels that its importance should be established in perspective instead of on the basis of its popularity. Establishing an epistemic correlation involves these three steps, as has been pointed out. First, we must have a sufficiently detailed theory system constructed in as close congruence to reality as is possible from an arm chair. Secondly, we must develop measurement tools which will represent the nature and relationships of our variables as they appear in the social mileau. Third, we must determine what our instruments are measuring. If they are not measuring what we think they are measuring, we must replace the corresponding variable in our theory system with the one implicit in the measuring tool. This will lead to realignments in the theory system and redesigning the measurement tools and a closer approximation of a reliable representation of reality.

The principal fault of the criterion technique research tool for establishing an epistemic correlation is that we cannot determine just what parts of the theory system the criteria are reflecting. We can take pains to select a criterion which represents part of our theory system but there is no known means of determining just what parts of the theory system it represents except by using the correlation between the test and the criterion measure and this, of course, is circular. It is like saying that Mr. A is a good policeman because he receives a high service rating and explaining that we will correlate our "Good Policeman Test" with this service rating. Now our "Good Policeman Test" consists of having the man shoot on the pistol range. Policemen in a city take this test and we find that our test correlates perfectly with the criterion and we then say that we have a true test for locating good policemen. To throw out what we now think characteristic of a good policeman on the basis of this correlation would be dangerous for this reason. We have no way of knowing that the people who made out the service ratings did not do it with scores on the pistol range before them. If this is true, all that we have proved is that our test is as good a measure of marksmanship as is the one used by those making out service ratings.

Factorial validity. The only conception of validity which corresponds to the logical problem of the epistemic correlation is factorial validity. This concept is very carefully developed by Guilford.[7]

Factorial validity is actually the correlation of a given measurement instrument with a common factor. This is a mathematical process allowing almost no opportunity for bias. This same procedure is used to establish the validity of a test with a criterion measure. The difference between these two methods is in the difference between

[7] *Ibid.*, pp. 513-522.

criterion measures and common factors. We have demonstrated that the validity coefficient resulting from the correlation of a test with a criterion is not as meaningful epistemologically as a substantial correlation coefficient might imply. Common factors, on the other hand, approach operational entities. That is, common factors are determined in such a way that they often correspond very closely with natural phenomenon. Being defined from a theory system they also fit it when the process of definition is completed. They are, in short, the best media that we now know for establishing a reality oriented theory system--one where all of the major propositions can be said, with some confidence, to represent behavior sequences.

The fundamental reason for the interpretative power of common factors is the mathematical complexity of the statistical basis of the factor analysis approach. The simple descriptive statistical techniques outlined above are, as was pointed out, often rejected by laymen because they represent only such limited aspects of nature. The experienced administrator feels that he knows a good deal about the underlying nature of the social processes he deals with. When someone attempts to demonstrate to him that one of his stereotypes does not correspond to the mean derived from objective analysis of a specific variable he often takes a "yes, . . . but" view and disregards the statistical conclusions because he feels that there are so many important variables not accounted for in the mean that this measure has little meaning. He is perfectly correct in this view.

Many social scientists do not wish to admit it but we are able to represent nature's complexities only to the extent that our experimental methods approach a degree of complexity similar to nature. This is an easily understood principle. In the days when most people owned a Model A Ford there were only four or five primary relationships in the equilibrium in a motor and a mechanic could put them in balance by checking three or four key points in the motor. Modern motors are built to obtain maximum power and a minimum of vibration. This means that they involve an equilibrium between some twenty odd factors. The mechanic now must use a machine which measures most of these factors when he is examining a motor.

Early sociologists were concerned simply with predicting whether or not a family would stay together. They utilized relatively simple insights. Now, when we are attempting to eliminate some of the vibrations which threaten family life, we have to rely upon whole batteries of complex questionnaires.

Factor analysis represents an attempt on the part of statisticians to provide us with a technique complex enough to handle data on social processes and derive operational terms which truly reflect the subtleties of our data. It is for the lack of such techniques that the case study approach has had to do as a means of scientific investigation. The human mind is not a reliable machine for the analysis of data where we need generalizations which approach the complexity of nature. It is to the everlasting credit of the social anthropologists, for example, that they provided us with a body of useful theory without these devices.

Factor analysis is like any other scientific procedure--it rests upon a set of assumptions. These assumptions are essentially mathematical guesses as to the nature of behavioral dimensions. They imply, for instance, that nature is not total confusion, but that there are relatively independent social forces which act in conjunction to produce the behavior which we observe. It assumes that we can construct instruments for the measurement of various aspects of behavior which have equal intervals between their units--in the same way that the interval between one inch and two inches is the same as that between two and three inches on a yardstick. We have no way of knowing now that behavior actually does not require a "rubbery" ruler. If we can make these, and other assumptions, we can apply the factor analysis method to social data and get dependable operational concepts. We can do this because factor analysis is a means of forming concepts which fit the implications of our data.

This brief review of descriptive statistics has served to point up some of the problems in the use of statistics as well as to point out the function of these statistical techniques. There has been no attempt here to describe the application of these methods. There are many good books on the problem. The only question here was how these techniques function as tools of the researcher. In the following discussion of sampling statistics space prohibits a discussion which does not assume some familiarity with higher statistical methods. This is unfortunate for as was noted above, these higher methods are just the ones which satisfy the demands of the layman for complexity.

III. STATISTICAL PROCEDURES: SAMPLING STATISTICS

The formal measurement techniques. Sedecor holds that the major function of the statistician, at present, is to measure the fallibility of the conclusions of other social scientists.[8] His concern is not with the meaning of experimental results nor with the applicability of these

---

[8] George W. Sedecor, "On a Unique Feature of Statistics," _Journal of the American Statistical Association_, 44:2, 1949.

results. He indicates the reliability and dependability of experimental results. This is a function quite different from the early concern of statisticians with problems of representing aggregates of information. Implicit in this interest in the reliability of conclusions is the concept of prediction, for prediction is the tool that is used to measure reliability.

The old dichotomy "reliability--validity" is helpful in explaining the function of prediction as a measure of dependable knowledge. In this oversimplified dichotomy, reliability indicates the extent to which our experiment picked out generalities which have applicability in more than one time slice. Validity refers to the wider meanings that our conclusions hold because of their integration with the elements of our theory system. These two qualities are obtained through almost independent processes so that we do not have one because we have the other.

> In social science, successful prediction may sometimes be made in the absence of adequate explanation. However, if explanation is adequate, prediction is always possible. Efforts to predict often stimulate efforts to explain.[9]

The characteristics of validity are such that if we have complete validity we have complete reliability,

---

[9] Paul Horst, et al, The Prediction of Personal Adjustment (New York: Social Science Research Council, 1941), Bulletin No. 48, p. 3.

according to Horst.[10] The inverse of this relationship is not true. In the statistician's eyes validity is not an immediate goal of measurement except as it is a means of increasing reliability and reducing error. The statistician's efforts to increase validity does lead to increasingly effective experimental conclusions. The layman often reacts against these because he perceives them as "cold-blooded" and feels that the statistician lacks moral responsibility. This reaction is not justified by the results of the statistician's efforts.

This discussion is based upon the proposition that sophisticated measurement is carried out through a sequence of statistical activities. There are many different sequences of statistical procedures available, any of which amount to substantially the same thing. This is because there have to be a number of ways of carrying out each of the steps in the measurement sequence so that for almost any study situation there will be a procedure to fit its peculiarities. The first step in any measurement process involves definition of the characteristic to be measured and the subjects to be included in the sample. The techniques used to identify our subjects are called sampling techniques and they enable us to define the population

---

[10] Ibid., p. 3.

which we shall measure.

Another set of techniques are available to help us define the thing that we are going to measure. These will be called techniques for locating dimensions. Then there are a set of techniques for determining the relationships between variables. These are very important techniques in all social science investigations. Finally, there are techniques with which to test hypotheses. These might be classed as the most powerful techniques of the social scientist for they lead to knowledge.

If one disregards all of the details connected with these procedures for scientific measurement it amounts to a very simple sequence. We determine just what we are going to measure, we decide just what group of persons shall be measured on this variable, we determine the general nature of the relationships between our variable and other pertinent variables and then we formulate and test hypotheses based upon this information. The result obtained from testing the hypotheses we call knowledge, and we know about how much we can depend upon it.

<u>Sampling procedures</u>. One of the most important, if not the most important, of statistical devices is the sampling procedure. The sampling design performs at least three functions. First, it tells exactly where to make

our measurements. The sampling design helps determine exactly what measurements to make in two ways: It determines the scope of our measurement through the definition of the population which will be under study, and it determines the precise members of that population who shall be included in the study sample through the sampling unit.

The second broad function of sampling design is that it provides assurance that we have dealt with the most varieties of the behavior we are studying. Unless we have some assurance that we have represented total variance in the measurements of our population, we have no way of setting down conclusions which have any universality.

The third important function of the sampling design is that it makes experimentation feasible in many areas where the necessity for studying each and every member of the population would make research uneconomical. The great importance of the sampling design in experimentation springs from the fact that it is the manifestation of probability theory in action, as the basis of the measurement process.

There are several basic concepts of sampling design which ought to be stated to identify the terms of reference. They contribute little directly to the understanding of sampling problems, however. One of the fundamental ideas in statistics is the difference between the concept "population" and "statistic." A population is the universe made up of all of those things which have the characteristic which we specify any member of the population should have. This is, of course, no definition, but it helps us set down the idea of population.

By setting down the characteristics of the thing or process that we are interested in studying we arbitrarily include everything which has these characteristics in our population. It is a purely functional concept.[11] The group that actually is measured is called the sample. The whole purpose of the sampling design, of course, is to obtain a set of measurements. The measurements which we actually obtain are on the sample and they are called statistics.[12] The measurements which we are attempting to estimate through the use of our statistical devices are those on the whole universe or the total population which the sample represents. These are called parameters to distinguish them from the actual measures the sample provides.[13]

The primary objective of the statistician is to gain as reliable an estimate of the parameter as possible. The

---

[11] Allen L. Edwards, Experimental Design in Psychological Research (New York: Rinehart & Co., Inc., 1950), p. 15.

[12] Guilford, op. cit., p. 176.

[13] Loc. cit.

general procedure by means of which he does this is sometimes termed the "control of sampling biases."[14] A statistic which provides an almost perfect estimate of the population parameter is said to be a completely unbiased estimate.

The first phase in sampling design is definition of the population to be studied.[15] The definition of the universe to be studied should be in terms of the geographical area over which the population units spread, the relevant characteristics of the behavior involved, so that it can be identified easily by the investigators, and the probable form the behavior will take in social settings at the time of measurement.[16] As Stephan points out, the more precise the specification of the characteristics of population to be studied, the more control of sampling bias we shall have, for this will reduce to a minimum the measurement of variance which actually does not fall within our problem area.[17]

---

[14] Edwards, op. cit., p. 18.

[15] Calvin F. Schmid, "The Measurement of Public Opinion," Sociology and Social Research, 34:85, November-December, 1949.

[16] Mildred Partes, Surveys, Polls and Samples: Practical Procedures (New York: Harper & Bros., 1950), p. 116.

[17] Frederick F. Stephan, "Sampling," The American Journal of Sociology, 55:373, January, 1950.

We must point out that even the sampling procedure is based upon a broad supposition. That is,

> . . . with respect to the material sampled, the assumption is that there is a large "universe" of uniform conditions, in that throughout the universe individual items vary among themselves in response to the same causes with about the same variability.[18]

Ezekiel is pointing out that in the light of the fact that everything is partially related to everything else, we have to make the assumption that the behavior which we are observing can be delimited with such precision that our measurements actually reflect only the variance of the specific variable under study, and no other.

Thus, if it is specified that the population includes all Negroes, does this mean that we include a person from a Negro family who has "passed" and has the attitudes and behavior of the middle class caucasian white? When we specify that the population includes all persons who are Catholics, does this mean those persons who attend the Catholic church on Sunday or persons who were baptized in the Catholic church or persons who call themselves Catholics?

It is obvious that the development of a really rigorous sampling design is just another one of the things

---

[18] Mordacai Ezekiel, Methods of Correlation Analysis (New York: John Wiley & Sons, Inc., 1941), p. 15.

that is dependent upon the exploratory studies and therefore should be one of the things which is included among the purposes of the exploratory study. Even such clear-cut limits as geographical boundaries become very difficult at times because the purpose of geographical boundaries are to cut off a homogeneous segment of population. In the urban situation this is virtually impossible. The practice is to specify the population in terms of the characteristics of subjects. The most satisfactory definitions of sample population are those which include a maximum number of characteristics of the people we wish to include in our population.

The second major step in developing a sample design is the identification of sampling units that are to be used. That is, shall we use a family, an individual, a work group, a secondary group, etc.[19] There is a slight ambiguity between this point and the last, since if we define our population in terms of the characteristics of people it is implicit that the sampling unit will be the individual. In administrative sampling there may well be times when our universe will be made up of groups, especially primary groups such as the work team, and in these situations we shall perhaps find it best to lay down the

[19] Partes, *op. cit.*, p. 117.

characteristics of our population in terms of the nature of the work team.

In random-stratified sampling method, the most common sampling unit is geographic, such as a block in the city, home on a block, township, county, state. There is no single sampling unit that could be used satisfactorily in every situation.[20] The major criterion for any effective sampling unit is that it be readily identifiable and as distinct from other phenomena in the social context as is possible. If there is any ambiguity in the identification of the sampling unit, it will very possibly introduce bias. Here again is a subject appropriate for investigation in the exploratory phase of research.

Hammond points out that there are actually two dimensions to the definition of a sampling unit. First, we must get conclusions representative of all the population units, and second, we must make sure the subject matter items on which we make population measures are representative of the whole subject matter being studied.[21]

---

[20] Frederick F. Stephan, "History of the Uses of Modern Sampling Procedures," Journal of the American Statistical Association, 43:23, 1948.

[21] Herbert H. Meyer, "Factors Related to Success in the Human Relations Aspects of Work Group Leadership," Psychological Monographs, General and Applied, 65:530, No. 320, 1951.

Thus, a psychological research team studying work satisfaction in a giant factory, was very careful to set up the sampling unit, which in this case was the individual, and then to select a group of individuals which was practically a complete representation of the whole work force in that factory. But they made almost no attempt to define work satisfaction and the attitude connected with it and so their attitude items put forth as measures of work satisfaction only represented a scattered set of dimensions of work satisfaction and could not at all be said to represent the whole variable.

This is a problem where the statistical techniques for locating dimensions should be applied. Some work could be done on this problem in the exploratory phase also by using alternative question forms.

Type of sample. The next major phase of the sampling design and one of the most crucial, is the determination of the basis upon which the actual unit to be sampled will be selected.[22] It is here that the most violent bias can enter. The selection of sets of sampling units which are not truly representative of the universe prohibits reliable estimation of population parameters.

---

[22] Partes, op. cit., p. 122.

Guilford notes that there are four important means of selecting the sample: random basis, random-stratified, purposive, and incidental.[23] A random sample is merely one where every unit in the population has an equal chance of being chosen.[24] This was the basis upon which selective service numbers were chosen in the last war. The principle is that in this scheme sampling controls are set down such that the laws of chance are allowed to operate with complete freedom. Purely on the basis of probability, if a substantial number of sample measures are taken, measures representative of all the variations of our variables will appear in the sample.

One of the techniques used to produce randomization is to procure a list containing the name of every individual in the population and assign a number to each one; through a table of random numbers select those individuals on which measurements will be taken. The numbers in the table of random numbers have been arranged scientifically in a sequence on some kind of a lottery procedure. If we are using the principle of the operation of the laws of chance, a sample which represents a substantial percentage of the total population must be taken. If the universe is

---

[23] Guilford, op. cit., p. 177.

[24] Loc. cit.

extremely large the use of a random sample becomes an expensive procedure. It is often difficult to get measurements on every individual selected through a table of random numbers because of illness, death, refusal, etc.

As pointed out in the last chapter this is the big disadvantage of the mail questionnaire, where the people who are most interested and therefore represent only one part of the population are most likely to answer.

Guilford notes that "A collosal example of biased samples is that of the Literary Digest Public Opinion Poll during the 1936 Presidential Campaign."[25] Here several million post cards which should have been a sample of adequate number were circulated, but the post cards were sent out from mailing lists prepared from telephone directories and automobile registration lists. In assuming that these lists represented the total population, the Literary Digest made a grievous mistake for actually an important segment of the population did not have telephones and did not have automobiles and therefore did not appear in the sample population. To combat this type of systematic error, stratification is becoming more and more important where we are estimating parameters of large populations.

---

[25] Ibid., p. 178.

The principle of stratification is that through previous knowledge the major groups in the total population of sample units are delimited and then the attempt is to get random samples within each of these groups. This technique helps overcome the type of bias which ruined the Literary Digest poll. Common bases for stratifying samples are male--female, Democrat--Republican--Independent, professional--business--office worker, semi-skilled--laborer, and Protestant--Catholic--Jewish--no religious preference. Even finer divisions of the total populations can be obtained by using one or more of these continua simultaneously so that we have one of our sub-groups male--professional--Republican--Catholic, and then the rest of these combinations are used to lay down the rest of the stratification. The major advantage of the random-stratified sample is that we have a whole series of random samples within each of the stratified groups, so that we have more independent groups where the laws of chance have been allowed to operate freely and therefore more probability that there will be more true estimates of the population parameter. This advantage is simply a mathematical function of the design.

Guilford says that a purposive sample is one which is arbitrarily selected because there is evidence indicating that the sample is representative of the total population.[26] Experience has shown that in public opinion polling, for instance, there are certain states and regions that come much closer to representing national opinion at any given time than many, many other states and regions. These regions are so-called indices or barometers. That is, of course, a very convenient and economical procedure but it must always depend upon reliable prior information. It also runs the risk that conditions may change without our realizing it and the state or region would lose its representativeness without our knowing it.

Incidental samples are these taken because they are available. These are often called accidental samples, but the name accidental is not too appropriate because there is nothing accidental about their selection.[27] Thus, a great majority of the psychological studies made so far in this country have had students in beginning psychology classes as subjects, merely because they were available. The great disadvantage in an incidental sample is that it is probably not representative of any major universe.

In scientific research there are really only two alternative sampling bases; the complete random sample or the random-stratified sample. In general, the random-stratified

---

[26] Ibid., p. 180.

[27] Loc. cit.

principle is desirable if we have prior information for laying down dependable strata because of the advantage pointed out above that it overcomes the "one-shot" nature of complete random design. If we lack prior knowledge to set up a dependable stratification but have a complete list of all of the members of our population the random sample design may be used with equal satisfaction. If we lack the advantage of either of these situations then we must make a compromise. If the population is rather small and we do not have a "touchy" subject to investigate so that there will be a great number of refusals, perhaps a pure random sample is the best design. If we have a large population and cannot possibly take a sample of great size from it, we can make at least a guess as to reliable strata and use the second scheme.

> There is no "best" method of sampling that can be followed blindly in all instances. The most effective sampling methods are those that are selected specifically to test the situations in which they are used.[28]

There are a number of specific problems which arise in the use of any selection principle. All of those who have used the pure random principle have faced the problem of bias which results from their inability to get measurements on a number of units which were selected totally at

---

[28] Stephan, "Sampling," *op. cit.*, p. 372.

random.[29] Samples are always being biased because the most interested, the most accessible, the more intelligent, the more understanding or the more sensitive have a more sympathetic set toward the subject matter of the experiment and therefore allow a greater number of measurements. A good publicity program, sponsorship by an organization of which most of the people in the population are members, the use of strict reporting forms for all the people who are making the measurements, patiently repeated attempts to get measurements and other devices have been found effective in almost every case.[30]

Stephan suggests that if we are engaged in research which is in any of the final stages of scientific methods, we should conduct background studies which include a check on the randomized nature of our sample.

Size of sample. The fourth major phase of the sampling design is the determination of the size of the sample needed. The size of the sample which is required is determined by two factors: (1) the accuracy which is demanded and, (2) the degree to which responses to questions

---

[29] Ibid., p. 373.

[30] Dorothy Dickins, "The Problems of Sampling in Connection With the Study of Family Economics," Social Forces, 25:328-9, October-May, 1946-47.

form a 50-50 split.[31] Harris notes that another factor in any determination of optimum size of the sample is the problem of practicability.[32] Each of these factors is important in determining the dependability of our results but a consideration of equal importance is the relationship between the size of the sample and the significance of our conclusion. All of the tables that are used in connection with the t test and the F test and the chi-square test have as one of their dimensions either the number of subjects involved or the degrees of freedom involved and in every case these tables are such that the larger the sample the smaller need be the observed difference to be significant. Actually, if we have determined the complete statistical design for our study, and know what tests of significance we shall be using, it is a simple matter to estimate the size of the sample which we need merely by substituting in a formula and solving for N.

As we have mentioned before, all of the statistical devices are based ultimately upon the laws of probability. If these laws are to operate freely, the first necessity is

---

[31] Albert B. Blankenship, Consumer and Opinion Research (New York: Harper and Bros., Pub., 1943), p. 113.

[32] Marilyn Harris, "On the Determination of Sample Sizes in Designing Experiments," Journal of the American Statistical Association, 43:397, 1948.

that the sample which we select is as representative as possible of the population we are studying. One of the important criteria of reliable knowledge is that our conclusion rests upon statistics which are the best possible estimate of population parameter. Although a sampling design is never sufficient for the derivation of reliable knowledge, it is certainly one of the necessary conditions of all social science investigations.

IV. LOCATION OF DIMENSIONS

The second step in the measurement process is to determine dimensions on which to take measurements. If there were a comprehensive set of primary variables in social science and a standard set of measurement instruments to go with them, this step should not be necessary. As it is, one must identify very carefully the dimensions that he is interested in whenever he investigates a new area of behavior. One reason for concern with these elaborate procedures for defining the dimension which is to be measured is that this facilitates fitting the new dimension into existing theory systems and thereby allows inference immediately on the basis of the relationships between the dimension and the rest of the variables in the theory system. It must be granted that these inferences are not of reliable precision until the hypothesis testing stage, the last stage to be discussed below. A second reason for this interest in setting down the dimension of new areas of behavior in great detail is that it aids in the derivation of dimensions which have universality. For example, if one were interested in the attitude of supervisors he should get dimensions on those attitudes which are common to supervisors in an iron foundry, a government agency, and perhaps even supervisors in Australia or England. This is much more difficult than merely finding dimensions which will describe the behavior of supervisors in a given factory situation or in a given industry.

In concerning ourselves with the problem of meaningful dimensions, we are, in fact, concerning ourselves with the question of the scientific model discussed as one of the philosophical problems which any good science takes into account. For any given segment of behavior on which we gather data, there are an infinite number of possible interpretations, descriptions and explanations. The problem of the social scientist is to find a set of dimensions which can be used to give optimum interpretative or predictive power for the maximum expanse of variance. Once he arrives at this set of dimensions he has defined the theory system model. The scientific function of these techniques for locating dimensions is to enable us to establish a model which has maximum universality but also maximum interpretative power in particular situations; genuine simplicity but also enough complexity to reflect all the subtleties in the behavior under study; and sufficient predictive power without requiring elaborate and expensive measurement tools. The desirable scientific model strikes a balance between a number of contradictory conditions, each of which is necessary but none of which is sufficient alone to give us both interpretative and predictive power.

At the core of this problem one finds the epistemic correlation. We can go at the development of the model in two ways. First, we can set up a theory or even a theory system and then identify data which substantiates it; second, we can collect data on behavior which we are interested in understanding, and then attempt to derive a theory system from these data. The statistical procedures for locating dimensions that will be discussed here are all applications of the second approach because any statistical procedure rests upon a set of mathematical assumptions. There are a number of techniques to be discussed below because we have had to invent a variety of techniques so that we have a set of mathematical assumptions which approximate a great variety of social situations. The first type of procedure has been termed scaling methods and the second procedure to be discussed here is factor analysis.

<u>Scaling procedures</u>. The broad assumption underlying scaling techniques is that

> . . . there exists a set of latent classes, such that the manifest relationship between any two or more items of a test can be accounted for by the existence of these basic classes and by these alone.[33]

It is probably impossible to develop a set of classes which would account for every bit of the variance involved in our data and at the same time satisfy all the myriad criteria which we have discussed above. It is frequently possible to find a set of classes which will approximate this goal. That is all one expects any of these procedures to do. If one gets a match such that the mathematical assumptions are even approximately fulfilled he considers the situation satisfactory for scaling. If, then, the scales accounted for approximately all the variance in the data, we consider these satisfactory. Guilford's discussion is authoritative and we use this as the basis for the description of elementary scaling procedures.

The simplest method of scaling is the technique of paired comparisons. It involves the simple comparison of each item to be scaled with every other item to be scaled. One of the means of doing this is to take a slip of paper

---

[33] Samuel A. Stouffer, <u>et al</u>, <u>Measurement and Prediction</u>, <u>Vol</u>. <u>4</u> (Princeton, New Jersey: Princeton University Press, 1950), p. 6.

and place each item with one other item on a single slip and then compare the items on each slip. Computation of scale values involves determination of the proportion of favorable judgments a single item receives and then the comparison of this proportion with the proportion of favorable judgments that each of the other items receive. Since the method of paired comparisons is based on the law of comparative judgment, we then must test and see whether our obtained scale values conform with this assumption.[34]

Where there are a great number of items to be scaled, it is often a proper procedure to select a few representative items from the total number of items to be scaled and scale these through the method of paired comparisons and then scale the remaining items in groups about these "key" items.[35] The range of applicability of the paired comparisons technique is perhaps the greatest of all the scaling methods. This is because of the very limited number of mathematical assumptions underlying this approach.

A second simple means of scaling is the method of rank order. It is especially useful where there are a relatively large number of items to be scaled. This method

---

[34] J. P. Guilford, Psychometric Methods (New York: MacMillan and Co., 1936), pp. 220-230.

[35] Ibid., pp. 235-6.

involves the same principle as the method of paired comparisons but here a series of judges are asked to place items in rank order on the basis of the degree to which they present or lack a characteristic. Scale value is computed by averaging the rank given the item by the different judges. Guilford reports five different methods of deriving rank order scale values.[36] Each of them is particularly appropriate to some type of situation.

The rank order method suffers from the proneness of judges to want to classify all of the items in the middle of the range. This results in the ranks at the extremes having considerably larger intervals than the center ranks. Guilford feels that the technique of rank order is probably, in the over all, more satisfactory than the method of paired comparisons because of its great economy in time, without a corresponding loss in dependability.[37]

The third common method of scaling is the use of rating-scale. This familiar device involves classification of behavioral items on the basis of some stimuli presented along a continuum. In the graphic rating scale a description of the variants under study is given and the judge merely marks the point along the scale at which he feels

---

36 *Ibid*., pp. 246-254.

37 *Ibid*., p. 260.

the specific item should fall and then this is given a numerical value by simple measurement. On the standard scale, the guidepost may be given in terms of actual samples, for instance, of handwriting, or a description of behavior, and the judge classifies the item according to this. In the numerical rating scale, descriptive phrases or sentences may be used and they are fixed to a letter or number and the judge records the number which he feels best represents the classification of the given item.[38] In some cases two or more of these scale types may be combined into one. A big problem in the use of rating scales is the prior determination of the unit.

The great limitation on the rating method is that the judges are even more prone to the error of central tendency. In fact, this error is so widely recognized that it has been given the colloquial name "the halo effect." A second source of error is called the logical error and this involves the tendency of raters to rate items which they see as similar at approximately the same value. Thus, if they see honesty and trustworthiness as being related and they rate a person high on honesty they are very likely to rate him high on trustworthiness also. Guilford points out that this error represents a presupposition or set

---

[38] Ibid., p. 263.

in the mind of the rater.[39]

In contrast with the paired comparison method or the rank order method, rating method requires much less time. It is more interesting to the judges and therefore may have more validity, and it can be used with psychologically naive judges. On the other hand, rating certainly lacks the validity to be used as a final measurement of any behavior and it appears that the reliability of rating a scale is less than that obtained from either of the other two methods.[40]

These are the simple methods of deriving scale values. They are methods of making measurements where we do not have tests or other devices where a known dimension and reliable units are available. Obviously they are not methods for picking dimensions out of thin air. If we do not have some prior knowledge that the continuum exists, we must hypothesize the existence of a continuum before we can ever derive scale values through the use of any of these three techniques. Their great disadvantage, and it is a serious disadvantage, is that we never have any way of checking whether or not our scale actually represents a single dimension. A scale, for instance, might represent

---

39 Ibid., p. 275.

40 Ibid., pp. 280, 281.

the operation of two or three or more variables which only happen to be operating simultaneously in a particular situation when we make a study. Certainly we could not classify these scaling techniques as exploratory measurement devices but they do not have the precision of many of the other statistical techniques that we will discuss.

During World War II a method of scaling called the scalogram technique was developed. This method is built on the assumption that for any given segment of social behavior we can find a single set of variables and lay down a single set of relationships between these variables for the analysis and interpretation of this body of variance.[41] The importance of this technique over the simpler scaling procedures mentioned above is that it allows us to determine with some confidence dimensions and relationships which are actually represented in the data which we gather from studying the behavior in which we are interested. A practical advantage, and one which is at the same time a disadvantage, is that this technique provides us with meaningful interpretative hypotheses while not fitting all of the data into a complex, logical construct.[42] This is the advantage of the factor analysis technique.

---

41 Stouffer, *et al*, *op. cit.*, p. 6.

42 *Ibid.*, p. 7.

The scalogram technique is based upon the principle that a person who is very honest is also quite honest and of average honesty and somewhat honest, because these are merely the behavior parts of which complete honesty is made up. The real importance of this technique springs from the reverse of this fact, that is, that the behavior of the somewhat honest person is actually related to the behavior of the completely honest person and that these two types of behavior are not, as they might well appear to an observer, different types of behavior.

In using the scalogram technique what we are actually doing is first locating hypothecated dimensions which are logically independent of each other and then secondly, we find point where these dimensions tie in to actual behavior so that we have a means of interpreting behavior mathematically in terms of independent measures which we have established as the first step in the process. The scalogram technique is especially appropriate as a means of determining dimensions which fit the different variances represented in data.

Stouffer notes that while the scalogram technique has gone a long way toward helping us solve the first problem, it has not helped much toward the second one.[43]

---

[43] *Ibid.*, p. 47.

The mechanics of scaling with a scalogram board involves determining whether or not the data describing many different types of specific behavior patterns can be fitted along a single continuum. It can be seen from this that as soon as we have determined which items of data fit along out continuum, we have established the dimension or line which we are looking for and we have established some markers along this line. The scalogram technique does not, however, tell us what guideposts are missing from our line. We must derive these logically. In the factor analysis technique we do get some indication of the blind spots on our continuum.

As we pointed out above, the great advantage of the scalogram technique is that we can verify the existence of a dimension which we have hypothecated in the questions used to collect our data. Stouffer himself, however, denies that the scalogram technique represents any lasting solution to the problem of the epistemic correlation, although he does not use this term. He feels strongly that the scalogram technique will give us predictive dimensions but he points out that predictive dimensions are often indeterminate because they are not unique.[44] The great problem of definition is to locate dimensions which are

---

[44] _Ibid._, p. 56.

unique both in the sense that they allow us to represent in one dimension a great amount of detail in the data but which also fit our other variables. Stouffer makes this differentiation by holding that there is external or predictive validity and also internal or tautological validity.[45]

An example is the case of using age as one of the measurements of individuals. If we ask a man his age and the response which we get checks with his birthdate, we have established external validity, but internal validity involves the uniqueness of age itself as a measurement in terms of its relation to personality, stability, emotional maturity, physical condition and the like. That is, how is age related to the personality of the individual, his emotional characteristics, and his physical characteristics. What we get down to, in terms of the scalogram technique, is whether or not the guideposts which we use to establish our connection between the discovered dimension and actual behavior are the best ones to represent this behavior and Stouffer is quick to admit that we have no way at present of determining this statistically.[46]

The basic mechanism of the scalogram technique is

---

45 Ibid., p. 27.

46 Ibid., p. 59.

this, according to Stouffer:

> We shall call a set of items of common content a scale, if a person with a higher rank than any other person is just as high or higher on every item than the other person.[47]

This eliminates the problem of ranking every item in relation to every other item. Secondly, if one response category is higher than another, then all the people with responses on the higher category must have higher scale rank than those in the lower category. These two principles are the concepts upon which the scalogram analysis rests. The scalogram board, which is used to manipulate the responses of all of the people to a given item, is simply a means of determining which items are consistent with these two principles. In the beginning of the analysis we do not know which items are going to show the pattern, but by putting all of our items into the scalogram board and then manipulating the order of the items on the board, a scale is derived.[48]

It is obvious even to the layman that the scalogram board performs a function which is very little different from the researcher simply sitting down and attempting to determine what items in a questionnaire are related to

---

47 *Ibid.*, p. 62.

48 *Ibid.*, p. 63.

each other and what the order of this relationship is. The big advantage of the scalogram board is that it provides us with a mechanical instead of a logical means for deriving these relationships, therefore largely eliminates the bias and the indetermination which enters through logical processes. Using the coefficient of reproductability we have a statistical check upon the scalability of any given item. This is a check that we have upon any logical procedure.

It is very difficult to evaluate the scalogram technique in relation to other statistical techniques at this time because it is so new. Most reputable statisticians are taking the position that it looks like a promising technique for the location of dimensions which have interpretative power, but that we shall have to wait until it has been used in a great many different situations before we can make any final judgment upon its effectiveness as a measurement tool. Certainly the problem of determining whether or not the items which we use to establish the scale are the most meaningful items that we could possibly have selected is one which will inevitably leave the conclusion derived using this technique open to criticism. But this is a criticism which can be leveled at almost any statistical technique and that certainly should not be considered as damning weakness of the scalogram approach. The very fact that reputable statisticians are withholding their judgment certainly indicates that this technique is important enough to be used if it appears applicable.

Factor analysis. To the writer, the most powerful statistical technique that is available to us in this stage of development of a science of administration is the factor analytic method. The factor analysis technique is a means of deriving the minimum number of concepts or dimensions which will completely describe the given set of data. This is a technique based upon the mathematical manipulation of individual differences.

> When a particular domain is investigated by means of individual differences, one can proceed in one of two ways. One can invent a hypothesis regarding the processes that underlie the individual differences, and one can then set up a factorial experiment, or one can utilize a more direct laboratory experiment, to test hypotheses . . . .[49]

In the first case we start with a hypothesis that determines the nature of the measurements that we will make, then enter into factor analysis. In the second, we start with no hypothesis but proceed instead with a set of measurements or indices that cover the domain. Actually, all factor analysis utilizes both of these approaches,

---

[49] L. L. Thurstone, Multiple-Factor Analysis (Chicago: University of Chicago Press, 1945), pp. 55-6.

because we cannot by definition merely test an uncharted domain, we have to have some idea of what we are going after when we make up the tests which go into our factor matrix. To the extent, however, that we are able to proceed with an indefinite set of measurements and derive basic dimensions from these, factor analysis is a means of lifting ourselves by our own bootstraps. It is this characteristic of factor analysis that probably gives it such broad general interest as a scientific method. The second characteristic of factor analysis which gives it such importance is that the common factors which we derive are at least in a limited sense, unique factors. They are unique to the extent that the rules used in carrying out the factor analysis computation produce concepts which are consistent with the rest of our theory system and which represent all of the vagarities of our data. Factor analysis often provides us with concepts which are more meaningful and have more interpretative powers than concepts developed through any other method.

Factor analysis is essentially a procedure by means of which we analyze a table of intercorrelations between a battery of tests to see which of these relationships represent the impact of "common" factors; that is, which of the correlation coefficients in a table of intercorrelations between 10 or 15 or 20 tests represent the effect of the operation of several powerful variables. The first step in this procedure leads to the development of what is known as simple structure. Simple structure, however, is not a unique solution in any sense, so the next step in the extraction of common factors is called a rotation, and within the limit of mathematical foundations of the factor analysis technique, through this sequence of computations we get a set of common factors which are unique. There is a good deal of discussion at the moment as to whether the mathematical foundations of factor analysis are broad enough and whether all the pertinent criteria for unique solution are included. But certainly there is at least a degree of uniqueness in the common factor pattern that results from a factor analysis.

One aspect of factor analysis which makes it such a powerful tool is that it can be used to cut through individual differences which appear between subjects on a battery of tests to locate the common configuration or pattern of ability which accounts for these differences. Without aid it is difficult for the mind to find the generalizations inherent in a single test battery. It is clearly beyond the power of man to look through the individual differences represented by a battery of tests and find the common factors inherent in them.

From a very practical point of view one of the major advantages of factor analysis is that we do not have to have a sample which conforms to the rigorous conditions just discussed. Our only concern is that the sample of subjects tested represents as wide a variation as exists in the universe. While it takes at least some care in the selection of the sample to assure maximum variation, the need for a great effort necessary to secure a truly random stratified sample is eliminated. Thurstone suggests that some care to assure randomization in the selection of the sample will probably produce the greatest variance.[50]

The great disadvantage of the factor analysis technique is that it does not provide us with logically equivalent concepts. The extraction of common factors and the rotation to simple structure is a complete mathematical process which provides us with mathematical entities. We do not, as in the case of the scalogram method, have any guides which establish a relationship between our common factors and the actual behavior represented in the data from which our common factors are derived. This has no limiting effect upon the common factors as devices for the development of predictive instruments. It does, however, pose a problem in fitting our new dimensions into the existing theory system, since our theory systems are, at

---

[50] Ibid., p. 324.

present at least, developed logically and these common factors are mathematical entities. Thurstone takes the point of view that we do not need interpretative concepts as long as we have complete predictive control.[51]

The writer's feeling is that while this would be true if we had an ideal science where we actually achieved complete predictive control, in the next million or so years before we achieve this state it would be nice to be able to use some of the results obtained through this powerful technique and to do this we shall have to fit them into our theory system or else develop new theory systems through this technique. The task of developing an experimentally based theory system through the factor analysis technique is prohibitive in practice by the amount of time and money required, so the only alternative is to attempt to fit these mathematical concepts into our theory systems as best we can.

Homogeneous tests. Another means of locating basic dimensions is the so-called technique of homogeneous tests. Most instruments which we call tests are actually collections of items which have some predictive value. There is no single dimension of behavior. The homogeneous or

---

[51] Ibid., p. 332.

univocal test is one whose items all measure the same element of behavior. Loevinger sets down two types of homogeneous tests. The first is the cumulative type where a subject scores correctly on all of the items which describe and then negatively on the rest of the items in the test which do not apply to him. This is the general principle upon which many ability tests are designed, of course. The differential homogeneous test is one where a person scores correctly for a number of answers and then negatively for a number of answers and then positive again; or he may score alternately positive and negative several times. Many attitude scales form a pattern of this type.[52]

On the cumulative homogeneous tests the only differences between people result from the length of their series of correct answers. In the differential homogeneous tests the person is differentiated both on the number of pluses which he has and the location of sequences of correct answers in the order of test items.

Loevinger feels that the technique of homogeneous tests is more appropriate in many experimental situations than either factor analysis or scaling procedure. Her reasons are:

---

[52] Jane Loevinger, "The Technique of Homogeneous Tests Compared with Some Aspects of Scale Analysis and Factor Analysis," Psychological Bulletin, 45:509-10, 1948.

1. The technic of homogeneous tests rests on pure, more plausible testable assumptions.
2. For a given number of items the method of homogeneous tests involves appreciably less work.
3. Psychologists can understand everything about the "theory" of homogeneous tests, if the methodology is worth the term "theory" with no more mathematics than high school algebra . . .
4. The technic of homogeneous tests leads directly to development of a test of predictable homogeneity. More often than not, factor analysis seems to lead not to tests but to hypotheses which it is hoped will lead later investigators to construct tests.[53]

To the writer, what Loevinger is actually saying is that in some situations the technique of homogeneous tests seems more appropriate than either factor analysis or scalograming. This does not decrease the importance of either of these other two techniques. The technique of homogeneous tests rests upon a great deal more knowledge than a factor analysis battery requires and therefore since it is more economical in terms of time it might well be used after factor analytic techniques had laid out some of the dependable dimensions around which test items might be developed. As such, the technique of homogeneous test construction is one of the most refined statistical techniques that we now have if it will do what Loevinger claims is possible.

We now have three sets of techniques for locating

---

[53] *Ibid.*, p. 525.

fundamental dimensions and measuring them with some precision. The simple scaling procedures and the more complex scalogram technique can be used very effectively in situations which are almost completely unstructured and where we are attempting for the first time to find some indication as to the effective dimensions of social behavior. It has been used especially in the latter exploration of the more complex verbal behavior pattern, such as attitudes, opinions and the like. It is a very logical and productive step to proceed from case studies to the construction of attitude items and then the manipulation of these items in a scalogram analysis or some other type of scaling procedure, in an attempt to determine whether items actually fall along some hypothecated dimension. To any one familiar with scaling procedures it is obvious that one of the disadvantages of this technique is that it cannot be used when we are dealing with a myriad of variances. In the really complex and therefore almost by definition most important areas of behavior, we are dealing with such great complexity of possible dimension that factor analysis is the only practical means that we have, for an almost unlimited number of test batteries can be thrown into a factor matrix. In fact, generally, the larger the number of tests in the matrix, the more sound is our factor analysis. Once we have established, with some confidence, the meaningful dimensions of behavior, homogeneous tests or further factor analysis may be used to find the configurations of many specific social situations in terms of our hypothecated dimension. Before we can ever have a science of administration or a science of politics, all of these powerful statistical techniques will have to be applied extensively in research. Our present verbal mechanisms of definitions and classification must inevitably give way to these more precise mathematical phenomena.

Measures of relationship. If we are to have a theory system made up of interrelated concepts or dimensions, we need, besides the statistical techniques for locating dimensions and defining concepts, a set of techniques which will enable us to set down the relationship between these concepts. For this purpose the statisticians have provided us with correlational techniques. These are statistical procedures by means of which we establish relationships between intelligence and ability to write. If we are to construct an adequate theory system we must know the interrelationship between each of the concepts in our system. Many people are familiar with the Pearson product-moment of correlation which is so widely used, but there are five or six other important means of establishing a relationship.

The basic procedure for all the correlational techniques is the examination of a considerable number of pairs of measurements of behavior. We know that on any given characteristic of human beings there will be variance whether it be the height of the American male, the average national birth rate of the United Nations member nations, or in the number of boxes handled by a group of car-loaders. This variance can be easily understood, however, if we know for instance, that all short men have small feet and all tall men have large feet; that is, even though there is a variance in the height of men and in the size of men's feet, if these two characteristics vary concomitantly, we can gain predictive power about the height of a man by knowing his shoe size.

All correlational processes can be represented by thinking of a grid on which we characterize the height of a population of men on the vertical dimension and their shoe size on the horizontal dimension and then take a series of measurements on a population sample of American men and place a tally mark in a single square in the grid, which is located on the horizontal axis by shoe size and on the vertical axis by height dimension. If there is no relationship between the size of men's feet and their height in inches, there will be tally marks scattered randomly in all the boxes in our grid. However, if there is a relationship between foot size and height, there will be a pattern in the grid boxes in which tally marks are concentrated. If the tally marks are restricted to the diagonal grid boxes running from a lower to an upper corner of the grid, we say there is a 1 to 1 or perfect correlationship. If the pattern of tally marks forms a curve, we say there is a curvilinear relationship. If the pattern of tally marks forms an S or a U or a J, we say that it is a complex curvilinear relationship and a sequence of procedures for handling this situation called "curve-fitting" procedures will be discussed below.

The most general index of relationship is called the correlation ratio. It rests upon the most limited number of mathematical assumptions and can be used where there are curvilinear relationships and in cases where data tabulated on the scatter diagram does not form a normal distribution. This makes it especially important as a device for measuring relationships in studies where we have no way of knowing that any correlation exists. One peculiar thing about the correlation ratio is that it produces two correlation coefficients and two regression lines. We get a correlation of X with Y and another correlation for Y with X. These two <u>eta</u> coefficients are not necessarily the same in value, although they may be. The big drawback with <u>eta</u> is that it is extremely laborious to compute.

The chief advantage and the chief use of the <u>eta</u>

coefficient is the determination of the relationship between two variables when the regression is either known or suspected to be non-linear. *Eta* literally shows the goodness of fit of the data to the regression lines determined by the means of the columns and rows. This is both the strength and the weakness of *eta* coefficient. Because the *eta* coefficient allows the regression curve to follow whatever pattern is established by the means of the columns and rows, it is much more subject to sampling errors than are the more refined measures of relationship which depend upon the curve-fitting principle. All in all, Guilford feels that the best approach in a research study is to attempt to develop some mathematical function and fit the curve to it rather than to use the *eta* coefficient because of the many sources of error which it is subject to.[54]

A second widely used method of correlation is Spearman's rank-difference technique. *Rho* is practically the numerical equivalent of Pearson's r. It is actually a product-moment type of coefficient, worked from the curve-fitting principle, but instead of being computed directly on the product-moment principle as is the Pearson r, it is computed in terms of the differences in rank between two

---

[54] Guilford, *Fundamental Statistics in Psychology and Education*, *op. cit.*, p. 323.

measurements of a variable. The process of computation is very lengthy and it gets prohibitive as the number of items to be correlated exceeds 25 or 30. The samples which can be handled by means of the <u>rho</u> are so small that it results in little more than a means of estimating whether a correlation exists from the small sample.

The most widely used of all correlation methods is the Pearson product-moment method. The coefficient of correlation derived from this method, r, tells us both the degree of relationship and the direction of relationship; that is, it tells us the degree to which X accounts for the variance in Y and tells us whether a large amount of X produces a large amount of Y or a small amount of Y. The Pearson product-moment coefficient of correlation is a powerful research tool because it does tell us both of these things. However, this correlation coefficient is based upon three mathematical assumptions which are a good deal more demanding than those underlying either of the correlational techniques just discussed above. It demands: (1) that the measurement for each of the variables being related form a distribution which approaches normality; (2) that the standard deviations of the dispersions of columns and rows in the correlation table be approximately equal--that is called the assumption of homoscedasticity; and (3) it assumes that the variables vary continuously between the largest and smallest values represented in the distribution.

The effect of all of these assumptions is to demand that we have a distribution of individual measurements which can be represented by a straight line. This does not mean that all of the measurements have to fall in a perfectly straight line in the scatter diagram. It does mean that we can represent the axis about which the distribution would balance by a straight line. These assumptions limit the use of Pearson's product-moment correlational technique, for we seldom find two distributions which conform perfectly to these mathematical assumptions and we almost never find distributions which conform to all three of them satisfactorily.

There are three measures which we can use in evaluating the Pearson r. The first of these is called the standard error of r. This is a measure of the degree to which our linear or curvilinear functions, representing measurements of the whole universe. In this sense it is not an estimation of the dependability of our measure of relationship but rather a measurement of the possibility that the relationship reflected by collected data presents a true relationship existent in reality.[55]

---

[55] Ezekiel, _op. cit._, p. 131.

A second often used measure for evaluating the Pearson r is called the coefficient of determination and it is nothing more than r squared. This coefficient shows what proportion of the variant in the dependent variables is accounted for by the variance of the independent variables. This is obviously an extremely important measure in any research where we are attempting to account for the behavior of some social variable. A third means of evaluating the Pearson r is the coefficient of regression, usually called the regression coefficient, which tells us the slope of the regression line and therefore permits us to make specific predictions. Fourth, there is a chi-square test of linearity where we take <u>eta</u> and the Pearson r and put them in a formula which also involves N and k and determine whether there is a significant difference between <u>eta</u> and r and therefore whether the curvature present in the regression line is only a chance phenomenon or whether it represents a genuine attribute of the relationship.

<u>Special techniques for correlation</u>. The multiple and partial correlation techniques are among the more useful special correlation techniques. The multiple correlation method is a means of determining what factors are responsible for a given variance and even more than this, what part each plays in producing the total variance. The

multiple regression equation tells us what variables are involved and the proportion of the total variance accounted for by each of these variables and the proportion of total variance unaccounted for by the variables used in our multiple correlation. Thus, if we wish to know how important verbal ability, mathematical ability, mechanical ability and persuasiveness are, in administration, we can correlate these with some effective criterion of administration and determine the percentage of unexplained variance that we still had to account for, say, through personality measurement. The coefficient of multiple correlation, R, is a proven means of establishing the relationship between more than two variables. It must be noted, however, that if the number of variables exceeds five or six, the calculation of the multiple regression equation becomes a rather involved procedure.

Partial correlation is a technique which can be used to determine the relationship between two variables when there are a number of interrelated variables on which we have correlation coefficients. If we have some idea that a good leader has high verbal ability, a good deal of persuasive ability, a particular attitude structure and a characteristic I. Q., the multiple correlation coefficient will tell us whether or not these four variables account for successful leadership in a given situation, but it does not indicate the relationship between leadership and any of the other variables. Partial correlation is used to hold verbal ability and attitude variables constant while we measure the relationship between persuasiveness and intelligence. And, of course, we could get a partial correlation coefficient for the relationship between any of these two pairs of these variables. Actually, the partial correlation technique need not concern us much here, because it has been almost completely supplanted by the analysis of variance methods which provide fuller control and do not leave us dependent upon Pearson r's and all the errors which they are subject to.

A second group of special correlation techniques has been developed for situations where we do not have both of the variables continuously distributed over the range. Thus, if we are correlating the yes--no responses to a question in a test, with the individual's total score on a test, we will be correlating one variable which can never be anything but dichotomous with another variable which can range from zero up to some limit. The correlation technique that we might use here is called the biserial correlation method. Biserial r is based upon the assumption that the people who responded either yes or no would have responded along a continuum of quantitatively larger categories if they had been provided the opportunity and therefore what actually has happened is that a whole continuum has been divided into two sections where variation would be continuous if we had not arbitrarily chosen to dichotomize. r_b was designed so that it is meaningful as an estimate of Pearson r. It is based on the assumption of linear aggression and homoscedasticity. As with Pearson r, biserial is based on the additional assumption of normality, as we just pointed out. Since biserial r is nothing more than a means of estimating Pearson r, one should always calculate Pearson r if he has the opportunity.

In a situation where we have a genuine dichotomy such as black and white, or male and female, it is impossible to assume that there would be a normal distribution. A method called point biserial correlation has been developed. The great advantage of the point biserial correlation coefficient is that it is not based upon the mathematical assumption of continuous distribution and can be used in many varied situations. Being a product-moment coefficient, it does rest upon the assumption of linear regression, however. What we probably should do if in doubt is compute r_{bi} rather than r_b. Its great disadvantage is that although it is a product-moment value, it is not numerically comparable with Pearson r and therefore it is difficult to interpret. Tetrachoric correlation method is a product-moment device when both of the variables are dichotomous and it rests upon the same assumptions as does the biserial method and has the same disadvantage. The tetrachoric correlation is not too widely used but when it is used it is often extremely valuable because it is either the only method that we have of determining relationship or it is a very quick and efficient method. If we use graphic devices for estimating r_t it is a very efficient and economical method of estimating relationship.

Finally, when both of the variables are dichotomous the phi coefficient can be used. The big disadvantage of phi is that it does not have stable limits as do the rest of the correlation coefficients.

In summing up this discussion of statistical techniques for measuring relationship between variables we should point out several fundamental characteristics of these statistical coefficients. The first and by far the most important point that should be made is that all correlation coefficients are relative measures. None of them are absolute measures in any sense. Many laymen feel that a correlation measure lower than .4 must be disregarded. Actually a correlation measure which does not differ significantly from zero is of equal importance or even greater importance than many of the measures between .6 and .9 for they tell us that we can, with some confidence, assume no relationship. Also, a low correlation in an area where we are just beginning to take measurements may not indicate that there is a weak relationship but only indicate that our measurement devices need refinement or that we take more variables into account or some other such factor. A coefficient is always relative to the nature of the population which we have measured and to the nature of our measurement devices. It is a matter of simple logic that if our measurement devices are highly refined and well understood, our correlations are going to be larger for that reason alone.

A second fundamental point that we should make is that the correlation coefficient which is obtained is partially dependent upon the range of variability of the factor being measured. Factors which vary within a narrow range are almost certain to produce relatively higher correlation coefficients than those which operate over a wide area. This has powerful implication for the study of administration, where many of our factors have the greatest variance. We shall not find high correlation coefficients here until we have virtually perfected our measurement devices. It will be ridiculous to deny relationship between leadership ability or administrative ability and effective organizational processes, for instance, merely because we get low correlation between measures of these processes.

Finally, the heterogeneity of the population sample will influence the size of our correlation coefficient. This is obviously partially a function again of complexity and since administrative situations are of relatively great complexity in comparison with many psychological and sociological phenomena, we shall find low correlation coefficients for this reason.

In summary, we must say that the correlation techniques are powerful statistical devices but they are open to criticism from many avenues and we must recognize that this criticism is bound to come and be prepared to discount much of it.

V. TESTING HYPOTHESES

We have now reached the final stage of scientific experimentation. It should be noted that by the time we have reached this stage in the scientific process almost all of the work has been completed. The testing of a single hypothesis is a rather mechanical process because almost every step in the procedure has been worked out in previous phases of experimentation.

Theory of errors. Implicit in the discussions throughout this paper have been the theory of errors approach. The scientific method, from one point of view, is a means of avoiding many of the errors which the layman is subject to when establishing a conclusion. In the second chapter of this paper it was pointed out that common sense knowledge, the result of the layman's intuitive processes of developing knowledge, can never be substituted for reliable knowledge. At that time common sense knowledge was rejected because we said it lacked certain scientific characteristics. Now we wish to point out that the more fundamental reasons for rejecting common sense knowledge is that it is not developed in such a way that we have any assurance that the errors have been controlled. The usual means of establishing a common sense conclusion is for a person to use his intuition and "make up his mind" and then "stick by" his conclusions, "come hell or high water," as the saying goes. This process is open to every conceivable type of error.

The common sense knowledge of the academician has involved a method which eliminates at least some types of error. The academician takes great pains to define his problem and gather "facts." If the problem is rigorously defined we eliminate the possibility of error from an ambiguous question. If the "facts" which the academician gathers happen to be the appropriate ones, we eliminate errors due to incomplete evidence.

Some of the early social scientists controlled an additional source of error by setting down alternative hypotheses which covered all of the possible outcomes of the experiment and thus controlled personal biases in the perception of what data should be collected. This accompanied the attempt to use instruments which provided quantitative rather than qualitative results since using qualitative data never provides any way of telling the extent to which bias is responsible for conclusions. There is no doubt that some investigators can survey qualitative data and derive completely objective conclusions, but there is also little doubt but what most investigators looking at the same data would derive conclusions which were fifty or sixty or seventy-five percent the result of preconceptions on their part.

Today's social scientist may take all of these several sources of error into account and attempt to control them and, in addition, he may make preliminary exploratory studies or pilot studies of various sorts where he determines at least to some degree the likely pattern of data which he will find and uses this as a guide to the design of his major experiment.

The social scientist engaged in fundamental research is almost the only researcher who can afford to use an experimental design of sufficient complexity to control all important sources of error. The development of reliable knowledge could well be defined as the development of knowledge which was subject to a minimum amount of error. By definition this would involve a method which had a maximum of power to control error variances. At the heart of the concept of the experimental design is the idea that each of these experimental procedures, while integrated with the whole design, plays some part in controlling some type of error and then pairs and triads of these experimental procedures, when operating in conjunction, eliminate still other sources of error. This is the reason for the continued emphasis upon the need for pre-planning which takes account of all three phases. If we do not do this we leave cracks in the walls of our design unchinked and though our individual techniques are the best, gross error may be allowed to enter unnoticed.

In addition to controlling the source of error discussed above, the modern social scientist has a sampling design which defines carefully the population he is testing. He has laid down tests of significance for his conclusions and he has made enough exploratory studies so that he has some idea of the degree to which he approximates the mathematical assumptions underlying all of his major statistical procedure. One of the objects of this discussion is to point out, from this new perspective, the function which the exploratory studies perform. If we are to use statistical procedures which are based upon an assumption of continuous variation or normal distribution, we have to have some idea of whether our data are going to conform to this pattern. If the data do not conform to this pattern the experimental design has to include other statistical procedures. But if we include other statistical procedures, then the data which we shall collect will be somewhat different in form if not somewhat different in kind. It all points up the fact that scientific method is not one method at all but a number of methods which must be used in sequence, each phase of the method being just as important as any following phase because the theory system which we develop rests upon each one of the steps.

There are actually two broad types of tests which we make to test the hypothesis. The first is to test the dependability of our estimate of parameters. This falls under the topic of sampling design which has been discussed below. The second broad area of tests are called tests of significance; using the theory of errors approach, what we are basically doing is testing whether the pattern in the data which we collected is significantly different from the pattern of data hypothecated in our hypothesis and therefore whether or not we have substantiated or reputed our hypothesis. Here we are actually determining whether or not we can hold that the epistemic correlation is established. The mechanics of executing this rather simple function become extremely complex because there are many different classes of situations demanding different types of testing procedures.

Test of significance. The null hypothesis is perhaps the most widely used of all devices for tests of significance. As Mood points out, to the layman, the null hypothesis seems like an awkward way of going about something positive. Actually, there is good reason for this. Having attempted to cope with the many sources of error in our experimental design, the null hypothesis becomes a sort of all-over insurance policy which gives us a check on the larger aspects of the experiment as a whole. The null hypothesis has been developed to meet what Fisher calls the type I and the type II errors.[56] These two types of error result ultimately from the fact that all statistics, as we know them now, are based upon probability. Consequently, any statistical proof which we may put forward is a proof not of any absolute, but an indication that we have located some pattern. (Actually, we have no scientific means of absolute proof as yet.) All statistical proof being based upon probability, there is always the possibility then that

---

[56] R. A. Fisher, Design of Experiment (London: Oliver and Boyd, 1937), p. 19.

even though there are only five chances in a hundred that we have an outside case in every hundred experiments we actually will claim to have substantiated a hypothesis when we have one of the outside cases.[57]

The test of significance is tied in with the theory of probability in this way: having carried out our exploratory studies and established enough experimental control to provide some assurance that our data are truly related to the hypothesis under investigation, we set up the formal hypothesis to be tested in final form and collect our data, and then through descriptive statistics we determine whether or not the pattern in our data conforms to the pattern which we hypothecate.

The statistical device which we use to do this is the formal theory of errors, as propounded by Fisher and others. Using the theory of errors we set up our test in the negative form of the null hypothesis, where we say that if the difference between the pattern which shows up in our data and the pattern which we hypothecated is not significant, we will then reject the null hypothesis that our formal hypothesis is false, with a given probability, say five in one hundred. That is, if the difference is not significant (there is bound to be some difference) we say

---

[57] Edwards, *op. cit.*, p. 28.

that there are 95 chances in 100 that the formal hypothesis we are testing has been substantiated by our data. If the difference is substantial, then we say we accept the null hypothesis that our formal hypothesis is false. Again, we know what the probability is--that our formal hypothesis is false perhaps 95 in 100.

What we are actually doing is evaluating our experiment completely in terms of the probability theory, on the basis of the chance that the difference between or the correspondence between our pattern of data and the pattern in our formal hypothesis is significant. When we do this we are subject to two types of statistical error. The type I error is that we will reject our formal hypothesis as false when it is actually true. The type II error is that we shall accept our formal hypothesis as true when it is actually false. In the terms of a single experiment, of course, this rationale is not especially meaningful. In terms of thousands of experiments it becomes very meaningful and has real implications for any one of these experiments. The problem narrows itself down to this: what probability ratio shall we establish for accepting and rejecting a hypothesis so that we will be subject to a minimum of these two types of error?

Obviously, if we accept a fifty-fifty probability as our standard, we shall be rejecting a great many true hypotheses as being false. However, if we insist upon the one percent level of significance (99 chances in 100) we shall be accepting a number of hypotheses as true which are actually false. It should be noted that there is a strong tendency for any given experimenter to set up the one percent level of significance as his cut off point for he sees it as giving him the greatest possibility of accepting a true hypothesis, but we can point out in the light of this rationale that he is actually increasing the probability that he will select a false hypothesis as true and he is increasing the probability that he will be subject to a type I error. Edwards seems to speak for the majority of the mathematically inclined statisticians when he holds that it is the five percent level of significance which gives us the minimum possibility of either type I or type II error.[58]

Many laymen reject the theory of errors as a technique for establishing "proof" of a hypothesis, on the basis that it does not give us any absolute criterion of the truth value of hypothesis. The simple fact is we have no other way of establishing the truth value of a hypothesis than this, and so we must make it do. Fisher says that what the theory of errors actually allows us to do is to

---

[58] *Ibid.*, p. 29.

derive the best compromise that we can within the limitations of the scheme.[59] With the rationale of the null hypothesis set down, we now turn to specific tests of significance.

Fisher notes that in general the tests of significance are based upon the principle that if we get enough variables involved in our experimental situation, that total numbers of possible configurations of variables are so great that the probability of a chance pattern which corresponds to our hypothecated pattern is very small.[60] This is one of the most fundamental concepts of statistics.

The tests of significance to be discussed below are actually statistical means of determining, then, first, the degree to which our data imply a configuration, and second, the probability that this configuration is meaningful on the basis of the possibility that it could have occurred by chance. To determine this probability, we have to have designed our experiment so that it conforms to the mathematical assumptions underlying the statistical techniques and so that the data are in a form such that they can be manipulated through one or more of these statistical procedures.

---

59 Fisher, _op. cit._, p. 40.

60 _Ibid._, p. 60.

Because it is impossible to collect all data in a form suitable for the same test of significance, a number of statistical tests have been developed and we must, in designing each experiment, survey these different techniques and determine which of them will best fit the nature of the data we anticipate on the basis of our exploratory study. Fisher holds that an experiment which does not contain a test of significance for the selection of our hypothesis, is nothing more than an experience.[61]

The chi square. One of the most common tests of the null hypothesis is the chi square technique. It is especially adapted to situations where we hypothecate a particular distribution in the data and then wish to test whether or not the distribution which we actually get is significantly different from this.

Thus, if we hypothecate that there is no variance in the grade which oral interviewers give job applicants which results from the ethnic background of the applicant, you might expect that if all other variance was controlled, the same proportion of job applicants from all the major ethnic groups would be hired. If we used the experimental control of matching, for instance, and developed a large sample of

---

61 *Ibid*., p. 39.

people who had been interviewed for jobs having the same intelligence rating, same personality grading, etc., and who differed only on their ethnic background, we could then use the chi square test to determine whether or not the number of people of each ethnic background indicates a significant difference from chance. Note: the sample must be fairly large. In all tests indicated here it is assumed sampling problems are adequately handled as discussed above.

Chi square tests can be used to determine significant differences from the expected distribution. We frequently hypothecate that the distribution will be random, or normal, or that the distribution will be in some other form, and then use the chi square test to determine whether or not our data conform to this type of distribution. In an almost classic article, Lewis and Burke have laid down the conditions under which chi square can be meaningfully used. Being a rather rigorous statistical technique, the chi square test rests upon some rather detailed assumptions, but the two assumptions most frequently violated are: that the separate events measured are not really statistically independent and that the smallest cell entry is often not large enough.[62] They point out that Yates has developed a

---

[62] Don Lewis and B. J. Burke, "The Use and the Misuse of the Chi-Square Tests," Psychological Bulletin, 46:443, 1949.

correction to try to meet this problem of small cell entries, but the only real solution is to make sure that the minimum N in every cell is above 5 and preferably above 10.[63]

One of the big advantages of the chi square method is that it allows for continuous variation in the degree of significance and tells us whether the significance of our measurement is in the ratio of once in 20 times or once in a hundred times or once in a thousand times. Another advantage of the chi square test is that since it takes into account the degrees of freedom involved, that is, the number of comparisons which can be made between the variables which are involved (thereby, of course, allowing the use of more than two variables) it can be used in a great many more situations than can the student's test.

The t test. The t test is defined by Guilford as the ratio of a deviation to a standard error.[64]

Here, of course, we are dealing with a different means of testing the significance of a set of data for instead of hypothecating the given pattern and then making a comparison, we merely check the amount of deviation which

---

[63] *Ibid.*, p. 463.

[64] Guilford, *Fundamental Statistics in Psychology and Education*, *op. cit.*, p. 208.

we find existent in the data against the standard error inherent in the data to see the dependability of the deviation which we located. It is worth noting once more that here the deviation which we found may be absolutely accurate; since we have no absolute measure, we must resort to a probability test. As such, t is a means of estimating the probability that our sample represents the larger parameter. Obviously, if we had measured the whole population on any measurement scale, we would have absolute proof that the pattern which is reflected in the data, exists. We would still have no absolute proof of the meaning of the data but we would have some proof that the pattern observable in the data does exist.

As Guilford points out, "We cannot prove the truth of the null hypothesis; we can only demonstrate its improbability."[65] The practical problem is that we never can take a complete measurement of all of the members of our population and therefore we have to estimate the larger parameter using sampling devices. The function of t is to help us determine the reliability of our estimate. Regardless of what the t value is, any investigator can always maintain that the rejection of a null hypothesis proves that in his sample his formal hypothesis holds. But this, of course,

---

[65] Ibid., p. 216.

is not what most of the investigators are concerned with. To the layman there appears to be an almost unlimited number of ways in using t as a test of significance. A few of these will be discussed here.

Perhaps the simplest of all uses of the t test is in the difference between two means. Here, after finding two means, we compute the standard error of these two means from their own standard deviation and then divide the actual difference between the two means by the standard error of the difference of these means and we call this the t ratio. Using the t table and the number of degrees of freedom involved we determine whether or not our null hypothesis is improbable. Edwards gives a formula for the use of t in testing whether or not the correlation coefficient is significantly larger than zero. In the light of the nature of the correlation coefficient, as pointed out above, this of course, becomes a very important test when we are dealing with small or near zero correlation coefficients.[66]

Guilford gives a formula for the use of a t test to determine what the difference is between proportions in significance.[67]

---

66 Edwards, op. cit., p. 124.

67 Guilford, Fundamental Statistics in Psychology and Education, op. cit., p. 209.

One of the disadvantages in the use of the t test formerly was that an experimenter often went to unnecessary expense to collect correct data upon a very large sample to help assure that his t would be significant. In sequential analysis the experimenter has a device which allows a more or less running check upon the level of significance he has obtained and having pre-set the level of significance he desires, sampling is automatically stopped at this point. This is a highly efficient procedure.

<u>The F test</u>. The F test is very similar to the t test except that it is more restricted. The basic mechanism of this test is to determine whether or not two samples which were drawn from the same population are representative of two different populations. It is similar to the t test because we use the ratio of variances which the data from each sample reflects to make our test. t and F are actually related so that when two groups of paired observations are involved, F is roughly equivalent to t squared. The practical application of F results from the fact that the t test is not too satisfactory when we are dealing with extremely small samples because the standard error becomes rather unstable. The F test, which deals with the ratio between variances rather than the ratio of some descriptive statistic to standard error, is much more stable at this

point and so we commonly use the F test when our samples are very small, say below 30. Although different in form, the F test gives us essentially the same conclusions the t test provides.

The analysis of variance. Our final statistical test of the null hypothesis also is based upon manipulation of variance and is called analysis of variance. Again, analysis of variance is basically a technique for determining whether or not observed differences are significant. More specifically, analysis of variance is a means of dealing with differences between more than just two sets of measurements. For instance, if we have five sets of measurements, we have ten possible comparisons; six sets means fifteen possible comparisons, ten sets, forty-five and so forth. Of course, we could handle this situation by pairing off two sets at a time and working a t test or an F test, whichever was appropriate. But analysis of variance increases our efficiency in this situation because it embodies the F-ratio test in this design. What we actually do is to test the significance of the difference between the within and the between group variance.

Guilford notes that analysis of variance is based upon four assumptions. These are:

1. The contribution to variance in the total sample must be additive. Naturally if the units

involved in a group measurement are not similar we cannot make any of these comparisons.

2. The observations within sets must be mutually independent. This is merely to say that our observations have to be uncorrelated for we are basing our analysis upon a probability theory again.

3. The variances within experimentally homogeneous sets must be approximately equal. This is only to say that the within-group variances must be about equal.

4. The variations within experimentally homogeneous sets should be normally distributed.[68]

In evaluation of the various tests of significance the analysis of variance test is one of the most useful and meaningful in a study of administrative problems because of the requirement that we take into account the number of variables in the same moment. It is impossible, of course, to make any all-over evaluations in these various techniques for, as was pointed out initially, they are all satisfactory for the purposes for which they are designed and any satisfactory experimentation must inevitably make use of the appropriate one.

---

[68] *Ibid.*, p. 247-8.

CHAPTER VII

AN EXPERIMENTAL PROCEDURE FOR RESEARCH IN POLITICAL SCIENCE

The purpose of this chapter is to bring together many of the things that need to be kept in mind in the development of experimental conclusions. It is now obvious to the reader that the scientific method is not a series of steps which need only to be followed to turn out scientific truths. The scientific method is only a guard against subjective perceptions, certain types of false logic, the type I and type II statistical error and any other errors for which control procedures are included in the experimental design. Many laymen hold the belief that if one doggedly follows a prescribed sequence of steps, dependable conclusions are inevitable. This is not true. Any investigation must begin with background study and the next practical step is some sort of exploratory work. Measurement may be possible after this.

The important process in a study is not that these three phases follow in sequence. It is virtually impossible to carry out a study in any other way. The power of the scientific method is in the experimental controls built into the measurement phase through interrelating each step in the first two phases with the last step of the scientific method--testing the hypotheses. Laymen may not realize that there are several kinds of controls needed and that the selection of the most effective controls and the integration of the controls selected is an art. The creation of knowledge which has cold precision requires the warmth and skill and native ability of an artist or an artisan.

This procedure reflects the principle of repetition as the basis for the creation of reliable knowledge. We are in complete agreement with the person who suggested that scientific conclusions are only "educated guesses." What this wit neglected to add is that if one makes a series of "educated guesses," each of them more reliable than the preceding one, our final "educated guess" is an approximation of truth. If our scientific investigations are successful, each guess provides a more accurate representation of the phenomenon we are dealing with and thereby provides a more sensitive means of determining what the next guess should be. This allows the use of more precise measurement techniques so that each reformulation of a hypothesis brings it closer and closer to a scientific law. A scientific law is nothing more than a hypothesis which it is agreed has been substantiated.

The initial formulation of the hypothesis, which may be so loose that we do not deign to call it a hypothesis at all, will probably be based upon uncontrolled observations.

On the basis of this "arm-chairing" variables are formulated. These establish the relationships which we feel exist. Next enough data to be indicative are collected.

It is usually difficult to determine whether or not the data which we collect actually substantiate the hypothesis with any degree of determination but it performs the very useful function of helping us get some idea of what factors are related to the variables. Later, using refined hypotheses and experimental techniques, all selected on the basis of preliminary exploration, more data is collected and on the basis of this we get some indication of whether or not our hypothesis is ultimately going to be tenable or whether it needs revision. The direction the revision should take may also be indicated.

It makes no difference what we call this process. Against any ideal of absolute truth, hypotheses are only guesses. This does not detract from their utility but it does serve to indicate the tentative nature of the hypothesis. Once we have collected data and analyzed them, the reformulated hypothesis becomes an educated guess.

In most scientific investigation the "safe" way is to proceed laboriously through each of the phases for the simple reason that at this pace the size and complexity of the problems which the experimenter must face are usually so reduced that the mind can meet them.

I. INITIATION OF INQUIRY

The first step in the initiation of our inquiry is to develop the background of the problem. Northrop notes that Cohen, Nagel, Dewey, Bacon and many others are all in complete agreement that the first step which we must take in the initiation of any investigation is the laying down of the traditional beliefs on the subject to be studied.[1] The writer would hold that we should go a step further than this and lay down, in addition, the stereotypes which the investigator has about the problem, for every researcher has a body of beliefs about our society. He must have this body of beliefs if he is to live successfully in our society, and having this body of beliefs the researcher is set to perceive many cues in the behavior which he is investigating.

Having made this attempt to gain some initial control over stereotypes, the next step in developing the background of the problem is to set down the social context in which the problem is set at the time the investigation is contemplated. That is, what social factors led the investigator to consider the problem. Is the problem in a social context of stress? Is there governmental or

[1] F. S. C. Northrop, The Logic of the Sciences and the Humanities (New York: Macmillan and Co., 1947), p. 16.

business pressure to undertake this research? Is there monetary gain possible through this research or is the research being undertaken as a fundamental research? What are the practical implications of this research? Finally, what is the purpose in undertaking this research?

If the potential investigator has recorded the social context of the investigation honestly, he should be able to get some indication of the ultimate practicality of the investigation by matching some of these factors. It is obvious that the odds are against getting really fine scientific results when the investigation is undertaken in a stress situation, where there is pressure for results and where these results are very likely to have monetary value to the investigator for one result and none for another result. On the other hand, it is often essential that we get some indication of what the underlying principles explaining practical situations are. If our purpose is only to provide some reliable guide post for industrial or governmental policy making, undertaking research in such a social context may not be undesirable at all.

To further develop the background of the problem, the investigator should attempt to identify as completely as possible the theory system in which his problem is set. Here, a first step traditional with the academician will stand any investigator in good stead, that is, review of the literature and perhaps a series of interviews with the recognized "experts" on the type of problem under investigation, the researcher can do several extremely important things. Most academicians immediately suggest a review of the literature regardless of what problem they face, without a very definite conception of what the purpose of this step can be. For the scientific investigator this step has two important purposes.

First, the researcher should set down every concept, variable, factor, etc., which he can conceivably relate to the problem to be investigated. Second, the investigator should keep a record of sources which contain any methodological techniques which it appears might have any utility in executing the investigation. From this series of concepts, largely extracted from the literature, the investigator should set down the preliminary theory system to be used as the theoretical framework in which to rig his experiment. This theory system, as pointed out above, should contain a hierarchy of concepts progressing from the most specific up to a series of levels of generalization until at the top we have a small number of classes which subsume all of the subclasses in all levels of the hierarchy below it. Next, one establishes the relationship between the concepts at the bottom of the hierarchy and between the subclasses in the other levels of the hierarchy.

Having established our theory system, the next step is to derive the assumption upon which it is based. If the concepts in the theory system each represent small processes which make up larger natural processes, what are the assumptions as to the basic causal forces responsible for the total process we have described?

Churchman notes that we should set down at this point a very detailed definition of our conception of probability.[2] By this, he means that we will define: (1) what we mean by event or observation; (2) what we mean by samples; (3) our elementary probability law; (4) what we mean by universe; (5) what we mean by independence of events; (6) what we mean by randomization, and (7) what we mean by degrees of freedom.

The layman, of course, will probably find that the most sophisticated probability theory which he could understand is the very one suggested here by Churchman and he may find it as much or more than he can cope with mathematically. The essential point is, however, that inherent in a conception of probability and manifest in a specific probability theory are implications about the basic dimensions of nature and meaning and causality. The big thing

---

[2] C. W. Churchman, <u>Theory of Experimental Inference</u> (New York: Macmillan and Co., 1948), pp. 15-19.

which we gain, aside from clarification of one's own thoughts on the matter, is at least the opportunity of locating inconsistencies in experimental procedure. This point brings out the position taken above. The scientific method is not a mechanical thing. While one person sets down a probability theory which is sophisticated and internally consistent but contradictory to the rest of his experimental procedure, another person sets down a probability theory which is completely consistent with the rest of his experimental design but quite unsophisticated in its mathematical implications. Though the results achieved through the use of each of these theory systems are about of the same reliability, weaknesses in these conclusions would be of different kinds.

Finally, in the development of the background problem we should attempt to anticipate those areas where the method will be insufficient to cope with the complexities which we see in the theory system and those areas where new methodological devices need to be developed. The I. B. M. Corporation provides a set of interlocking cards which may well be used to build the theory hierarchy. A careful cateloging of the anticipated trouble areas in our investigation may well lead to a reformulation of our problem or taking of some different tacks in our approach to it.

Precise statement of the problem. While there is general agreement that background development on a problem is useful, there is unanimous agreement that the first positive step in any scientific investigation is a careful statement of the problem to be studied.[3] The simplest purpose of defining the problem is the limitation of the general area of variance to be studied intensively. When one says, for instance, that he is going to investigate work satisfaction, or morale, he has not actually said anything about the specific direction his investigation will take. These ambiguous concepts could include or exclude almost anything. In defining the problem we lay down the types of behavior which we shall call work satisfaction and the type of behavior which indicates high morale or low morale; that is, *the problem is defined in terms of identifiable phenomena in the social milieu.* This amounts to the beginning of the translation of the theory system into behavioral terms and therefore the initial step towards locating the epistemic correlation.

Next, one should set down the purposes of the investigation. In a so-called pure scientific research the only purpose is the creation of a bit of reliable knowledge. It is more probable that an investigation will have

---

[3] Northrop, *op. cit.*, p. 17.

two or more purposes; the creation of reliable knowledge being one of these. Often social science studies are carried out to locate causes of problems or to develop solutions to problems or to learn how to do something which is successful in one social situation with equal success somewhere else. The reason for this definition of purpose is that it helps us determine exactly what variables shall be set up for intensive investigation. In studying morale, for example, it is absolutely impossible for anyone with only several hundred thousand dollars to undertake anything more than the study of a limited number of variables. In such a case the definition of the purpose of the study becomes extremely important in determining what variables would be studied.

The next step is the reduction of the problem to the relevant factual situation. As Northrop points out, in this type of analysis of the problem one is guided to the relevant facts "by way of the theoretical roots . . ."[4] This is an essential step in the organization of any study and represents the beginning of the research activity. The relevant factual situation is the real existent situation which contains the facts which shed some light on the experimental problem which has been established. Northrop is

---

[4] *Ibid.*, p. 30.

well justified in maintaining that at this point the investigation has already begun, for it is here that we make great slashes into the enormous structure of facts which we could conceivably collect. The direction and length of the paths we cut through the jungle of reality determine the gross nature of the measurements achieved in the measurement stage. In any scientific investigation we <u>never</u> make an attempt to describe the total phenomenon. We only make measurements on and observations of those aspects of the total process which are related (in terms of the theory system) to the variable, such as morale, which we are interested in studying.

This points out the importance of the theory system in the definition of the problem. If the review of the literature has been comprehensive, if representative opinions have been secured from experts, and then if the investigator has shown some imagination in filling in the gaps in the theory system, there is probably no better way of determining what variables to investigate than to take those indicated by the theory system. To deny the utility of the theory system as a means of identifying the variables to be studied, merely because it was constructed by using logic is ridiculous if one only secures a substitute from <u>a priori</u>--unplanned, inconsistent variables discovered by meandering around in the literature of the field or at the scene of the behavior.

The most efficient way of identifying the variables which we are going to place under study is to get a match between the limitations of the actual area of behavior made in the first step of defining the problem and the dimensions in the theory system which was constructed in the development of the background of the problem. This holds whether one is concerned solely with prediction or whether one has some interest in the derivation of meaning. This is a most efficient way of identifying predictive items because it provides some sort of subject matter sampling, assuming that the theory system bears some relationship to reality. It is an efficient means of locating variables which have both predictive power and meaning, because it is a way of beginning the long process of establishing the epistemic correlation, at the same time that we begin the location of predictive items experimentally. What has actually been done, if we attempt to match the area to be studied with some area laid out in the theory system, is to lay the background for the epistemic correlation. If we then hypothecate the variables to be studied on the basis of the theory system, there is some basis for feeling that there is a meaningful correspondence between our logical concept and observable behavior--that is, an epistemic correlation.

Some simple devices which may aid in the most meaningful statement of the problem might be mentioned. The first is to attempt an evaluation of the problem in the light of the degree of maturity of the theory system which we are using. Is there a good deal of research behind the theory system? Has it been used successfully as a prognostic device? If so, it may well be satisfactory as the foundation for the study of technical problems. If the theory system is in a developmental stage, it might be wise to restrict ourselves to problems which serve the purpose of testing the applicability of the theory system rather than immediately attempting to develop solutions to practical problems.

Another thing which one can do is to attempt to evaluate the problem as defined, in relation to the total scope of the theory system. That is, how important is the problem in the light of other problems which we can imagine? Does our problem seem to be the first logical step to take in the light of the problems we can imagine from this theory system? It does not make any difference how pressing the practical problems for research may be if the theory system has not ripened. Even the greatest external pressure is not going to squeeze out a dependable conclusion.

Third, we might attempt to anticipate what answers we are going to get to the problem, for we will have to do this sooner or later as we refine our educated guesses.

If the problem is actually undertaken one makes very exact statements of what is anticipated in developing the alternative hypotheses. Certainly if the problem as defined appears to have few inconsistencies with the practical situation, if the results which are anticipated appear consistent with what we believe is actually known on the basis of experience, then we have good reason to feel that our problem as stated is meaningful. As was pointed out above, because science is a self-corrective process there is no great risk involved in undertaking the investigation even if the problem does not appear to be defined with the greatest finesse. A poor definition of the problem is simply inefficient and uneconomical.

The final step in the statement of the problem is the development of a brief summary plan of the methodology which is anticipated. This provides another source for evaluating the form in which the problem is stated. For example, a problem concerning work satisfaction which required investigation of the processes in the subconscious mind might be very meaningful theoretically and appear to have great utility as a management device, but it is impractical because we have no reliable way of determining exactly what is going on in the subconscious. Investigations in this area must wait until methodological studies have produced devices which are reliable for this purpose.

If it is impossible to develop a methodological sequence of experimental devices which appear to be satisfactory techniques for the analysis of the various parts of the problem, we had best revise our problem.

The formulation of hypotheses. One of the two or three critical steps in any research is the development of hypotheses. The purpose of the background study and the statement of the problem is to lay a basis for the formulation of a hypothesis which is as meaningful and as experimentally effective as possible. It is at this point that we actually make our "educated guess." A hypothesis represents an attempt to state the variables and the relationship between the variables which we expect to be reflected in the data which are collected. Obviously, if we are dealing with some extensive complex of variables, a precise statement of the nature and relationship of the variables involved is impossible in one or two sentences. What actually happens in this case is that previous experiments are used as a basis for the development of sub-hypotheses, each representing a statement of the nature and relationship of two of the variables involved in the larger hypothesis and enough sub-hypotheses are utilized to pair off all of the related variables involved. It is obvious that this is impossible if we do not have previous research which can be used to select sub-hypotheses which represent the true relationships of the parts. If our purpose is to develop highly reliable knowledge, there is nothing but to carry out the experiment if sub-hypotheses are unavailable. If there is some pressing practical demand for a solution and there is some experimentation upon which to base our sub-hypotheses, one may feel the formulation of the relationships between the pairs of variables justified. If there are no previous experimental efforts upon which to base our sub-hypotheses, it would seem that probability of meaningful results is so infinitesimal that research is uneconomical.

The hypothesis serves a very important function in the total picture of research. It is important in the reduction of the problem to obtain more precise and concrete dimensions. In setting up our theory system and defining the problem, one begins the definition of the epistemic correlation he hopes to establish. In the translation of the problem into behavioral reality, variance hypothecation of our epistemic correlation is continued.

Definition in terms of the theory system leads to identification through the connection of definitions with real behavior. Having stated the hypothesis in the context of a specific problem situation one runs the risk of being proven wrong because this cuts off any possibility of defending one interpretation of the hypothesis through the manipulation of inherent logical ambiguity. Cohen and Nagel take the position that a hypothesis which does not establish an epistemic correlation is no hypothesis at all.[5] A hypothesis will here be said to be irrelevant if it does not perform this function.

The second major function of the hypothesis is that it provides a dependable basis upon which to select the method that will be used in the investigation. A hypothesis involving variables which it is impossible to measure through the use of tests or other psychometric devices necessarily indicates the need for interviews, projective techniques or some other method. A hypothesis which includes as one of its variables a factor for which there exists widely standardized tests obviously calls for the use of these instruments. The statement of the variables and their probable relationships eliminates groups of techniques which might be used in an investigation.

The third and the most important function of the hypothesis is that it indicates what data to collect. The function of the hypothesis is to direct our search for the order among the facts.[6]

---

[5] Morris R. Cohen and Ernest Nagel, An Introduction to Logic and Scientific Method (New York: Harcourt, Brace and Co., 1934), p. 202.

[6] Ibid., p. 201.

> It is to be noted that it is the analysis of the hypothesis which provides the criteria for selecting out of the infinite number of facts in the world the few that are relevant.[7]

The layman immediately asks exactly how the hypothesis determines the data to be collected. The answer is that the hypothesis by itself does not determine the facts we must collect. The hypothesis when coupled with the statement of the problem and integrated with the background of the problem indicates the data to be collected. It must be admitted that in the early stages of the investigation of any complex of variables the hypothesis does not tell, with any conclusiveness, just what data to collect and what data to disregard.

The hypothesis directly identifies the variables about which we wish to collect data and the relationships which we are to test. The problem is to collect all possible data which represent the nature of the variables involved and which represent the relationship, if any, between the variables. The central problem is to determine what data actually reflect the phenomenon. In many cases the decision as to what "adequate" data are, is purely an incidental or even accidental thing. The investigator arbitrarily determines what he feels is required in the way

---

[7] Northrop, op. cit., p. 33.

of data, he collects this and then he assumes that it is adequate. This is a mediocre procedure. Many investigators go through the motions of reviewing the literature, extracting data, classifying these data and even collecting items for tests. To the writer this is little better than the accidental procedure. The only satisfactory means of deciding what data to collect is to work back through the statement of the problem to the theory system, first laying down the broad types of data required from the statement of the problem, then proscribing the specific relationships which can be inferred from the theory system.

For example, in that part of the theory system referring to status, we might find the broad heading, status, divided into at least three sub-classes: status symbol, status striving, and status structure. Each of these might in turn be divided into sub-classes and so on with an identifying statement under each. Under status symbol we should have a concept which points out that status symbols are not necessarily functional but that they may be articles of apparel, equipment or mannerisms which indicate the relative position of a person in a social structure. By the mere presence of these three sub-classes under the concept "status system," we would assume that in studying status symbols we should also investigate status structure and status striving. We would also assume that we should investigate each of the sub-classes and the sub sub-classes under the concept "status symbol."

The proponent of the accidental system of selecting data or the person who favors a review of the literature as a sufficient basis for the selection of data types might immediately suggest that the description or measurement of some phenomenon which we call status striving does not mean that we have described or measured anything which is inherently status striving, for the epistemic correlation is not established until after the hypothesis has been substantiated. Therefore, it is more efficient and probably more useful to infer our theory system after data has been collected. The answer here is that the theory system, especially after it has some experimental evidence as a foundation, has the great advantage that it is a systematic means of selecting the type of data which we will collect and thus provides a basis for data selection which can be evaluated and controlled.

This is one of the real methodological weaknesses of experimental design. While many experimentalists utilize the most sophisticated of statistical devices, because of their opposition to the hypothetico-deductive method, they overlook completely any control on the type of data collected. The whole method rests upon the relevancy of data. The careful background investigation and the precise statement of the problem are the basis of the initiation of any inquiry.

In areas where there is little background alternative hypotheses perform a very important function. In planning a study in a new area the various variables can be inter-related in several hypotheses, each providing a unique combination of variables. This has the utility of providing a mechanism for thinking through the nature of the variables involved and indicating the nature of the variables and hence the type of measurements which should be taken.

The more refined use of the alternative hypothesis is in the final phase. Instead of using the null hypothesis, a number of mutually exclusive and comprehensive hypotheses covering all phases of the variance, statistical standards are established under which any particular hypothesis will be accepted.

At this point two truisms should be apparent. One, the development of fruitful hypotheses can never result from the application of this technique alone. Two, the ultimate source of really effective hypotheses is a sensitivity to the reality of the study situation, and the synthetic ability of the generalist to integrate the dimensions of the situation to be studied with the rest of known reality during the growth of his perception.

How, then, does one meet the practical problem of developing effective hypotheses? The logicians have made some very specific suggestions. The most powerful suggestion, in the eyes of the writer, is one made by Cohen and Nagel. The suggestion is to locate a very simple situation which represents the principal elements in the complex situation which the investigator wishes to study. Then to observe, study, and analyze this simple situation by actually working in it and living in it for a short time to thoroughly familiarize oneself with its every aspect. Next, leave the situation and allow a short time as a period for incubation. Following this, sit down and speculate in an attempt to determine the most fundamental independent dimension which will represent all the variance in this simple model of the larger phenomenon. Finally, use these essences as the variables to form hypotheses about the larger situation.[8]

To the pure operationist this might seem like the method of an armchair crackpot. It is, none the less, the method used by Newton in his observation of a rolling ball on an incline and by Watt when he observed the jiggling of the top on a steaming kettle.

Another way of getting insight into a situation is to analyze it in a number of different ways and then compare

---

[8] Cohen and Nagel, _op. cit._, p. 205.

these different analyses in an attempt to locate variables which will explain all of the analyses. This might be called the method of compared analysis.[9]

A technique very similar to this one is to attempt to classify a number of different situations using a set of variables derived from analysis of a single situation, and then to attempt to explain the reasons for the problems of classification. Attempts to reclassify existing sets of data are a variation on this same technique.[10] Of course, another source of fruitful hypothesis is the accidental insight. Inspiration might also be classified here. Needless to say, this is a rather undependable means of developing hypotheses, one that no successful scientist has ever relied upon.

Statement of presuppositions. One of the most subtle steps in the initiation of the investigation and also one which has far-reaching implications is the process of laying down presuppositions. The function of the presupposition is obscure and often so disguised that it is not apparent to anyone but the person who actually conducts the

---

[9] Ibid., p. 220.

[10] Raymond F. Piper and Paul W. Ward, The Fields and Methods of Knowledge (New York: Alfred E. Knopf, Inc., 1929), p. 253.

investigation. For this reason it is very easy for the experimenter to disregard the presuppositions in his investigation and for the critic to make his point in other terms. There is no method of determining just what all the presuppositions for any given experiment are or even should be. Even the most sophisticated proponents of the hypothetico-deductive method admit that many presuppositions represent only a guess.

The presupposition is nothing more than the context in which one makes an "educated guess." The hypothesis is a proposition stating that a given relationship exists between a prescribed set of variables. As we have pointed out many times in this paper, nothing is completely independent of everything else in nature, so when we hypothecate the relevant situation we must of necessity disregard the rest of the factors in the total situation which form the environment in which the test variables are operating. One of the assumptions of every scientific investigation is whether or not there is any change taking place in the relationships of the variables in the experimental environment and if there is change, how much change. In every experiment one assumes an environment which is dynamic in a changing equilibrium. An environment which is dynamic in a chaotic manner in the sense that we cannot determine any equilibrium along the variables operating in it naturally prohibits experimentation.

In any investigation of politics or government we must deal with at least three fundamental types of phenomena, individuals, the social group and the formal organization. In any investigation of the individual or the social group, the formal group is inevitably the context and conversely in any investigation of the formal organization the individual and the social group form the context. As every experienced administrator knows, an investigation which overlooks the impact of a powerful personality on a formal organization would be naive. The really important function of the presupposition is brought out by the point that the only difference between the hypothesis and the presupposition is that the hypothesis is the assumption we are testing and the presupposition is an assumption which we are taking for granted. Since the presupposition is an assumption which we do not test and often one about which we have very little knowledge, each experimenter may have a vitally different set of assumptions. In practice we do find that every investigator will have at least a unique pattern of presuppositions. If we are to make sense of differing experimental conclusions, then, we must know the suppositions from which they spring.

The presupposition for any experiment includes first the whole theory system not included in the hypothesis, and second, the fundamental assumptions upon which are based the foundation of the theory system. One might note that among those fundamental assumptions underlying the theory system which we may find in future experimental studies in political science will be those concerning the basic nature of the American culture. Experimental studies will require fundamental assumptions about the basic nature of man; whether he is motivated primarily by physical or emotional needs, or both; fundamental assumptions about the relation between man and his environment; the impact of heredity upon the determination of personality. On the methodological side a sound probability law which is accepted in its experiment-and semeiotic functions will include the nature and meaning of definition, what constitutes a satisfactory definition, and the like.

Experimentally one of the most important functions of presuppositions is that they aid in identifying those variables which are actually operating in a situation but which are assumed irrelevant. If the hypothesis is not substantiated by the data, we have a means of guessing which assumedly irrelevant variables actually must be included in the hypothesis. If one were testing a hypothesis in relation to leadership the theory system which became the presuppositions might include the concepts of communication, coordination, control, integration, and centralization. The fundamental assumption might assume a dynamic environment with a variable and a moving equilibrium. If the hypothesis provided no significant results, after checking the method used one might turn to the theory system and attempt to determine whether or not it was satisfactory to presuppose communication was not an integral part of leadership, and, if after evaluating this, we determined that our presuppositions were right, we should turn to our fundamental assumption about the general nature of the administrative environment. Many times difference in the pattern of the data of reliable results and unreliable results will point out, in conjunction with our system of presuppositions, just where the "educated guess" was wrong. In this situation, it is a simple thing to reformulate our hypothesis and repeat the experiment.

It is certainly not general practice at this time for an investigator to set down a careful statement of presuppositions. Some do not do this because they say they have no definite presuppositions. The point of view which is taken here and which is substantiated elsewhere in that an investigator who conducts his studies without stating his presuppositions is acting upon suppressed premises actually whether or not he recognizes it as such. A complete evaluation of any investigation should require that the critic derive the suppressed premises from the record of the investigation.[11]

II. EXPLORATORY PHASE

If the function of the first phase is the definition of the problem, the function of phase two is the development of the measurement devices and the function of the third phase is the collection of data. In the exploratory phase we face the question of what type of measurements will provide the data identified by the hypotheses. We carry the epistemic correlation along in the second phase by determining exactly what data we shall collect, and exactly how we shall collect it. Nearly everything one does in the exploratory phase concerns the identification of variables or the measurement of dimensions and thus involves the transfer of the operation from the realm of logic to reality.

In the initiation of inquiry one might determine that he wishes to measure voting behavior and that he wishes to find out what causes are associated with voting behavior. Before one can intelligently gather data about voting behavior the investigation must be transferred to some behavioral setting. This is a difficult and tedious

---

[11] Howard S. Bartley, Beginning Experimental Psychology (New York: McGraw-Hill Book Co., 1950), p. 33.

step. The days of the frantic data gatherer are gone. Data is meaningless except as it is a dependable representation of behavior. The device commonly used to organize data is to lay down dimensions of the behavior being studied and then make measurements on these. Measurements representative of all of the points on the dimensions which we have hypothecated are taken and these constitute the data. All of the measures along a given dimension are then intercorrelated with each other in phase three.

Measurement is a process of describing according to a set of rules. The measurement of household objects is so much a matter of habit that we are unconscious of the rules by which they are made. Who would think of taking a foot ruler and measuring a bedroom in any other way than measuring the two adjoining walls. Meaningful measurement in any field requires obedience to prescribed rules. The difference between the measurement of household articles and of voting behavior is that in the latter case we do not know either what aspect of the behavior to measure or what instruments to use in making the measurements. What is required is a set of rules for making measurements before the study ever enters the measurement phase.

<u>Selection of dimensions to be measured</u>. The point of departure in phase two is the selection of the dimensions

to be measured. There is a good deal of help from the efforts in phase one on this for if we set down a comprehensive theory system and if we prepare pertinent sets of alternative hypotheses we have the major classes of potential dimensions contained in these. The next step is the location of these dimensions in reality. This involves some sort of flexible exploratory study such as the use of undirected interviewing on a representative sample of voters. The object of these interviews is to determine to what degree our dimensions actually represent behavior, to determine what types of behavior might be classified as being along a given dimension and what behavior is to be disregarded because it is not related (is on no dimension) to our problem. No two questions measure the same thing. No two tests measure the same thing either. The scalogram technique and factor analysis become important in dealing with this problem.

<u>Development of measurement techniques</u>. A second purpose of the exploratory phase is the development of measurement techniques which actually take measurement along the dimensions hypothecated as pertinent. The fundamental technique of social research is to direct some sort of a stimulus toward a subject and record his response. The most common of all the stimuli we present to a subject is

a carefully phrased question. In general we shall get one measurement along some dimension for each short answer question submitted to a subject. Since we usually need to get measurements all along a given dimension we must submit a series of questions, which are points along the dimension under question, to each subject to find out where his behavior falls on this dimension.

In the study of something as complex as voting behavior there are hundreds of dimensions which have a significant relationship to voting and probably forty or fifty which will give us some ability to predict. This means that a complete study of voting behavior would involve the analysis and administration of fifty series of questions for each individual. It is because of this sort of problem that we use the undirected type of investigation in the early phases of research. It would be out of the question to find any representative sample of voters who would submit to answering five or six hundred questions.

A central problem in this phase of the research is the selection of exploratory techniques which provide a maximum amount of information. If one is doing some of the first research in a field he must have information which represents as many dimensions of the behavior under study as possible. If he is doing research on behavior where there is a good deal known about the major dimensions the need is for techniques which indicate the bench marks along these dimensions.

Relationship between dimensions. The third step in the exploratory phase is the examination of the relationship between these dimensions. The final description of these interrelationships will come only through the use of more refined measurements, but we have to make some beginning here to develop meaningful measurement devices. The measurements of a bedroom would be without meaning unless we knew that they were taken at right angles to each other and that these two measurements could be used to draw an outline of the complete bedroom because the relationships were such that all of the corners of the room were square and that opposite sides were equal. We make rules to give social science measures meaning also. Many of the rules seem quite complex to the layman. We attempt to make measurements such that when the condition of the economy is X and when the world situation is Y then a belief that the Republican nominee is isolationist will be related to total annual income of the individual in such a way that those with a given size income will tend to vote for a specified type of candidate.

It is on the basis of these suspected relationships that measurement techniques are designed. These techniques require careful definition of the relationships between the different dimensions for several reasons, one of the most important being that we require items and tests and batteries of tests as "pure" in terms of the number of dimensions measured as possible. Pure tests are those tests which have the minimal correlations with any others. One of the values of carefully planned exploratory studies is that they provide questions which may be used in effective measurement instruments.

Perhaps the greatest value of the exploratory phase is that it enables us to select between alternative hypotheses. In this way measurements will be limited to one general hypothesis. In social science, as in any other endeavor, making precise measurements is a much more complex operation than making simple measurements of quantities and relationships. Simple measures cannot provide us with information of much interpretative power.

In terms of the "educated guess" rationale a function of the exploratory phase is to provide an initial check upon the appropriateness of hypothecated variables. From the theory systems, from observation and through synthesis we hypothecate the dimensions which need to be measured. The first exploratory studies, involving the most loose experimental procedures, provide an initial check upon the adequacy of the dimensions developed from speculation. Do the behavior patterns in the situation under investigation correspond to the concept of the dimension? Is the behavior which can be identified through the dimension a type of behavior which is fairly easily differentiated from the behavior involved in other dimensions contained in the study? Do the subjects in the study situation feel that the dimensions used are the most meaningful aspects of the situation in the light of the experimental problem?

The immediate question is, "How does one determine whether there is behavior in the study situation which corresponds to the dimensions which were hypothecated?" What constitutes a "good" dimension? The experimentalist's conception of a good dimension is one which is independent of any other dimension in the study situation in that questions can be framed which reflect this dimension alone. This means that we need dimensions which will produce zero or insignificant correlations when intercorrelated with the other dimensions. No one has determined just what sort of characteristics such a dimension would have. For this reason if the investigation attempts measurements in new areas the writer feels that the insights of those who live in the study situation are invaluable. Perhaps behavior which is simple to identify will lend itself to the definition of the most meaningful dimensions, but this holds only for early stages of study. Many of the most powerful concepts that we have produced are of great subtlety, vis., group syntality in sociology, compensation in psychology, and aculturation in anthropology.

In the end, the process of establishing the epistemic correlation is a matter of making a guess on the most flimsy of foundations and then checking this guess with care through the use of experimental controls. There is nothing especially fascinating about the actual process through which we establish the epistemic correlation; the impressive thing about controls is that we are able to make even these crude checks upon the degree to which our verbalisms represent reality.

The result of the first set of exploratory studies should be the development of an inventory of possible dimensions and the selection of a set which are worthy of further study. The second step of the exploratory phase involves, first, the development of techniques that will provide us with effective measures of the dimensions on which we intend to take measurements and, second, the estimation of the interrelationships between these dimensions. It is in this step that we make some of the most important of all the decisions in the experimental process.

One of the products of this step is a picture of the amount of covariance between each pair of dimensions. It is this structuring that is the goal of the scientific method since it is the structuring of concepts that enables us to represent nature through verbal mechanisms and therefore increase our control over large segments of nature. The whole purpose of science is to develop such a structure.

The question which we face in working with the interrelationships of these dimensions is whether or not the structure appears to be adequate or can be made to be adequate or whether more dimensions are required. Repetition of studies is necessary before dependable structures are developed. The most common test of the efficacy of a structure is to make a prediction using these dimensions. This requires the refinements available only after the measurement phase of research is complete. What we actually have to do is make another educated guess as to the effectiveness of our structure and proceed to the validation phase. It is a simple matter to determine whether the prediction is faulty in respect to the number and nature of dimensions used by executing a factor analysis study. Having this check upon the adequacy of the various dimensions in the structure the patterns of effectiveness and weakness will indicate the efficiency of the exploratory study.

If there is little knowledge of what constitutes a good dimension, there is no knowledge as to what constitutes a good structure. It is held that the structure must be comprehensive, consistent and the like, but there is no indication of how one actually tells whether or not his structure is comprehensive. At this point science is still art.

In an attempt to achieve some universality this discussion has been extremely general as to the specific things one does in the exploratory phase. The answer is that one hardly ever does the same thing in the exploratory stages of any two studies because this stage is so much dominated by the conditions of the study situation. Probably if one is making the first study of a problem, he must begin with some sort of undirected interviewing to obtain first hand impressions of the types of variables he may have to deal with. Since all research is done on a budget the consideration of the measurability of the study phenomenon amounts to a determination of what proportion of the budget shall go into developing techniques and what proportion into the collection of data. In most cases any immediate research in political science will be on the development of techniques.

If the research involves development of techniques a central concern will be cooperation with experimentalists from the other disciplines of social science. Much time will be spent merely in developing practice in bridging cross-discipline barriers. This time will appear to be wasted. The devices produced will impress few political scientists and receive little attention from them. If the research is directed toward results there will be great consternation among experimentalists for they will be sensitive about the fact that the techniques being used will provide only measurements which are meaningless because no one knows what they contain in the way of dimensions. The experimentalists will demand careful adaptation of techniques. Any results will be argued by the political scientists and rejected by the experimentalists.

Nevertheless, we shall be able to create the basis for better guesses through this sort of an operation and we shall some day get to the point where there will be need for consideration of pilot studies with opportunity to consider both of these problems. A continuing problem will be the development of better samples so that the basis for generalization can be extended.

III. MEASUREMENT TECHNIQUES

There is not a perfect statistical design available for each study. A series of compromises on basic principles are necessary in the development of any design. The expert statistician knows only too well the limitations upon any statistical design. This difference in sophistication is illustrated by the tendency of laymen to accept so-called "statistics" on the number of people in urban centers or on the number unemployed as true measures while the statistician may not even accept them as reliable estimates. The political scientist who is undertaking experimental studies needs to be aware of the basic principles of statistical design.

<u>Sampling principles</u>. The principles of sample description and selection are the first statistical principles which come into use in the execution of a study design. What are these principles?

1. Definition of the characteristics of the sample which is to be taken. The definition of the population from which the sample is taken determines the generality of the conclusions which are developed by identifying the individuals to which the conclusions apply. (Note that this process is necessary whether there are a finite or an infinite number of subjects in the population).

2. Identification of the subjects upon whom measurements will be made to supply the raw data of the study. This involves the selection of the basis upon which the sample will be taken, such as random, random-stratified, and so forth.

3. The development of a set of sampling controls such that there is protection against as many biasing factors as can be anticipated before the pilot study is

undertaken. The result of this step is the establishment of a plan for the actual collection of the data which provides the controls which seem necessary.

4. The last phase of the sampling procedure is to check the data collected, on the basis of a previously established probability theory. (This is contained in the presuppositions of the study). The check insures that the sampling design which was established produced results which are dependable. This step involves the use of reliability and other types of error coefficients.

5. Determination of the size of the sample. This involves estimation through the use of reliability formulae on the basis of arbitrary limits of dependability. In the first place, the sampling design must be integrated with the statistical tests of significance for the form which the probability theory will finally t ake depends upon the nature of the type of test for significance used. The sampling design must also be integrated with the instrument which is used in gathering the data. If one is attempting to collect information about the opinions people have about their mother-in-laws he would probably not use a sampling procedure which involved street corner interviews. The prediction of presidential elections has been carried out effectively using street corner interviewing. If the questions to be used require some sort of free response

from the subject the sampling design must utilize some sampling unit which will allow all or most of the subjects to find a situation where they have sufficient time to respond to the instrument and where they have the facilities for writing.

It is characteristic of the scientific methods being developed in social science today that a careful attempt is made to control the sample, but that after the data on the sample are assembled a check is taken from an independent direction to evaluate the data. The reason for this is that the sampling procedures do not provide absolute controls in any experimental situation.

The development of a sampling design which takes account of these aspects of the measurement operation requires technical skill which few political scientists will have. While the question of whether to use the geographical distribution of subjects or some personal characteristic as a control on the stratification of the sample is largely decided by the nature of the experimental question, there can be little doubt but that the really effective sampling design involves the cooperation of a statistical expert and a political scientist. The statistician presents the technical aspects of the problem and the political scientist presents the substantive aspects, the design being derived by a balancing of the relative importance of these respective factors presented here.

It should be apparent that the sampling design cannot be developed independent of the rest of the statistical design.

One of the most common influences upon the sampling design is the availability of reliable data upon the population to be studied which can be used to establish strata in the population. In many cases the sampling design utilizes units peculiar to the U. S. Census Bureau or some other governmental agency so that known characteristics of the population may be used to set up the stratification.

As has been mentioned, the size of the sample cannot be determined until there is some knowledge of the dependability of the strata used and of the reliability which is desired. The relationship between the sampling design and the rest of the statistical operation has been pointed out to emphasize the interrelatedness of the various statistical devices and the necessity for cooperation between laymen and statistical experts.

Locating dimensions. The desirable characteristics of useful dimensions or variables is that they be independent of as many other dimensions as possible; that they be easily identified in the study situation; that they lend themselves to qualitative if not quantitative measurement;

and that they provide us with interpretative and descriptive power.

The fundamental principle which should be followed in selecting new dimensions is that the dimensions must establish a relationship between the theory system and the behavior which is of interest in the investigation. As was pointed out earlier it is through such connections that the body of a theory system can be related to a specific situation. Through such relationships knowledge becomes cumulative.

The next step in the development of independent dimensions is the location of reliable dimensions. The problem of determining when any two measurements represent measures of the same phenomenon can only be solved by using some of the relatively complicated statistical scaling techniques. If one is measuring honesty in executives and he uses anything but the single question, "How honest are you?" he faces the problem of defining what sort of variance to call honesty. As a practical matter, variance is the difference in responses of individuals to one or more stimuli, in the case of a test the stimuli being a question of some sort. (It should be fairly obvious that the single question, "How honest are you?" is unsatisfactory in measuring honesty since there is great difficulty in finding out what the question means to the subject, and therefore what the response to the question means).

Scaling procedures are often used to evaluate the degree to relationship between dimensions. One of the problems in scaling is the determination of which questions are related to the variance, honesty, for example, which is the object of study. A relatively straight forward means of doing this is to develop a measure which includes most of the variance included in "honest" behavior and then assess the degree to which any given item shows some relationship to the variations in a criterion measure. All of the items which show a significant correlation with the variations in the criterion measure probably are worth further consideration as items for an Honesty Test. The assumption is that since these items were designed to measure what was identified in the criteria as honesty, a relationship which shows itself is probably not a chance relationship. The pragmatic test of the efficacy of the items is whether or not they enable successful prediction of what is defined in the managerial situation as honesty. The other evaluation which can be made involves the use of a theory system. If the theory system is resting upon some knowledge which has experimental foundations it will show some of the boundaries of "honesty" and these can be used to evaluate what some items mean.

Professor Guilford's factorial validity is a second, and very important, means of "nailing down" the dimensions of behavior. Here we have the substantial advantage of overcoming the weaknesses of criterion measures. It has the disadvantage that it is not a practical method of locating dimensions and identifying variance unless there is background work which will provide some of the items for the tests, and some indication of the factors which are present in the field. The power in the factorial validity concept is that it does not rely upon a mature theory system. If there is research which indicates the types of factors which one might expect to find operating in "honest" behavior, the Honesty Test can be correlated with these. Intercorrelations might show that the Honesty Test measured the impact of an ethical factor, a niave-sophisticated bi-polar factor, and a flexibility factor. From this sort of analysis one gains further insight into the behavior he is interested in because of the precise interrelationships which factor analysis provides.

There are several other ways of defining the dimension one is concerned with. The homogeneous test technique described in the last chapter is one of these. Of more concern to the political scientist is the interpretation of the variable once it is located. There are no social variables except in the light of a purpose which means in the light of a value system. There is not now and there never can be anything ultimate or sacred about a common factor or any other dimension of behavior. As was pointed out, the factor analysis technique is designed so that it only produces one type of dimension and the item content of the tests used in the factor analysis process embodies a set of values, these subject to change with each investigator. No political scientist should be discouraged from proposing what seem to him new and more meaningful dimensions merely because he feels there is no experimental evidence for his proposed dimension. Dimensions established through some criterion measure are even more indeterminate since their only justification is that they have been found to vary in some consistent pattern with the criterion, but the only final justification is that they lead to predictive control. There is no known way to locate independent and mutually exclusive dimensions using the criterion technique.

A second aspect of the measurement problem is laying off the units along a dimension. This is termed the scaling problem. Techniques for the determination of scales are much farther along than the techniques for locating variables initially. It is a problem which we are far from solving with complete satisfaction as yet. Dependable scaling methods are paired comparisons, ratings and the like. The technical problems of scaling are not of interest at this point. It is worth pointing out that with the techniques available now make it difficult to say that there is anything which cannot be measured. No social science method measures as accurately as many would like, but this is also true of physical science measurements. There are few elements of behavior where one cannot make some estimation of a true measurement. The scalogram technique is an important addition to the set of scaling techniques.

Tests of hypotheses. The final stage in any concept of the scientific method is the testing of hypotheses. Most political scientists see it as a point at which "the facts" somehow prove or disprove "the hypothesis." Actually, it is the point at which all of the controls built into the experimental design come to bear on the experimental situation and the data selected through these controls is subjected to a statistical test. The product of this statistical test is a contingency measurement of the degree to which the data defined as related to the experimental proposition fell into the anticipated pattern. As was pointed out in the last chapter there are two dimensions to this contingency measurement: (1) the degree to which data conform to the hypothecated pattern, and (2) the probability that the data actually manipulated are accurate.

Summary. Everyone who formulates a generalization ultimately faces the problem of defining just how many people or how many of whatever are included under the statement. The results of tests of the sufficiency of the sample are easily understood and facilitate the sort of logical precision said to be typical of good political science research.

It is the tests of significance which many, perhaps most, political scientists disagree with. Some say that these tests are not complex enough and that the oversimplification of a hypothesis that must be built into a proposition before it can be tested defeats the purpose of the test. Others say that test of significance depends upon purely quantitative data and therefore the pattern of data which is compared is incomplete. Still others maintain that while these tests are satisfactory in every way they test the wrong things. These people maintain the statistical test of significance does not establish the kind of meaning that political scientists value.

One must admit that in each of these criticism there is some truth, more in some than in others. It is true that any given statistical test may be rather simple in itself. But a single technique is only meant to test a small part of the total problem at one time. The results of testing parts of the problem are then combined so that the final test of a substantial hypothesis reflects great complexity. To the criticism that all statistical tests deal with groups of measurements, but give one little insight into a single measure, there is no answer. If one wishes to call this oversimplification there is nothing but to admit that this is one of the characteristics of the statistical tool. This cannot be considered a fatal failure. Case studies will provide this information. It is not within the purpose of measurement.

Those who maintain that statistical procedures are limited because they depend upon measurements have a more tenable position. It is true that at this time there is much that we cannot measure which is important in interpreting and solving present problems. We are incapable of providing a contingency measure of situations where quantitative measurements are not available. At the moment, this means that we are unable to provide reliable knowledge about any of the central problems of political science since we do not have techniques for the measurement of political behavior. This objection is not fatal either since it is eliminated as soon as we develop the techniques for the measurement of political behavior. There is no reason for believing that it will be impossible, ultimately, to develop techniques for the measurement of any political process.

It is the final objection that has real foundation. Once the social scientist, and the political scientist, has learned to talk the experimentalist's "language" it is this epistemological problem that identifies the point of real disagreement. Those in the experimentally founded social sciences have come to be satisfied with an explanation which enables them to interpret a particular situation at the moment. To the traditional political scientist this sort of meaning is both narrow and shortlived. He is interested in what are often called "developmental factors," threads of history composed of cause and effect chains which he feels are explanations because they tell how our political institution evolved from some past state to its present condition. This is a valid objection on the part of the political scientist. The experimental social scientist does not attempt to satisfy the epistemological conditions of the political scientist. This point will be considered further in the concluding chapter.

This concludes an attempt to make the concept of the experimental design explicit. In the preceding chapters of this paper the attempt has been to lay the foundations of this concept on a bedrock philosophical consideration and then outline the problems and techniques which might be involved in scientific research in political science.

CHAPTER VIII

CONCLUSIONS

It is the conviction of the writer that the question, "Can we use the scientific method in the study of politics and administration?" should be answered with a confident "yes!" The question has been argued since the term, "political science" was invented. Unfortunately for a long time we did not know much about the scientific process. Most political scientists still know nothing of the scientific method and yet they still maintain that the use of the scientific method in political science is impossible.[1]

Political scientists have been fighting a phantom image when they raise their screams, "blasphemy," "soulless," "oversimplified," and "inhuman" to their worn out academic gods. Scientific concepts which approach any real generality are not cold and hard and stereotyped. It is the concepts that we use, developed through the historical method, that are lacking in any individual orientation and warmth. Scientific concepts represent nature, in some

---

[1] It is by no means the best available, but for the latest illustration of the political scientists' ignorance concerning the scientific method, the reader is referred to the first few pages of Hans Kelsen, "Science and Politics," The American Political Science Review, 45:641-662, September, 1951.

degree, and if nature is warm, and if nature emphasizes the great things in many, then a symbolic representation of nature will include these dimensions.

It is perfectly true that science will not give us a value system which can be used as a basis for establishing goals. This is the function of ethics. But if this is what the political scientists are concerned with, let us call our discipline, ~~political~~ ethics, and leave the development of a dependable body of knowledge about our political institutions to the primary social sciences. Ethics in relation to any institution are fundamental. It is time that we take some notice of the necessity of developing some means of bridging the gap between what we are presently dissatisfied with and what we dream of. Prudent men are in agreement that the further development of our culture rests upon the development of knowledge about the things which cause it to react. It is perfectly true that we must have a body of ethics which will enable us to choose between alternative objectives for our culture, but if we lack the knowledge to proceed toward these objectives we shall never have the opportunity of actually putting our mighty ethics to work. Ethics for ethics sake are little better than money for money's sake. It may be true that in the time of Plato a wise man could set down the goals for government and a council of philosopher kings could guide the nation steadily toward them. In American, we are currently attempting a twist very typical of us by dividing this labor and surrounding the president with a number of councils of philosophers in advisory capacities and with little success.

The explanations offered are many and varied, but behind each rationalization of our inability to deal with our social problems is the fact that we are not even certain what our social problems are and few are fools enough to feel that they know what to do about them. A few fools think they know just what should be done and no one pays them any mind because we are accustomed to thinking ourselves and our society so "complex" that we cannot even be studied. This is only true in one sense. We cannot begin by laying down scientific generalizations for the whole political institution. One can develop dependable generalizations about significant parts of the institution, parties, or the administrative process or the legal process for instance. The denial of the possibility of a science of politics rests upon a fundamental misconception of science.

The point brings us back to the epistemological problem mentioned in the last chapter. The political scientist maintains that we have not explained something until we have accounted for it in terms of its antecedent. The experimentalist feels that any pattern of relationships which he finds which will account for the behavior under investigation is sufficient. The point upon which these two orientations can be compared is the purpose of each of these types of research.

The political scientist is attempting to account for and predict the behavior of a whole institution. His fundamental assumption is that there is some limit to the patterns of institutional behavior and that if one studies enough institutions over a substantial time span he will be able to make some prediction and thus maintain some control over the future of a whole nation. At the same time, he is very concerned with the problem of ethics. He spends a good deal of time evaluating the probable effects of following alternative courses of action in the light of his studies of past behavior and provides this information to those who are interested as a basis for their selection of personal values. He sees the political institution as being made up of several processes which are at the same time independent and interdependent. There is the behavior which is identified as political, which includes the selection of the goals of governmental action and their revision; there is the operation of the government itself in achieving its policy objectives; there is the operation of the legal system which is at the same time a process and an institutionalization of the values of the culture.

When the political scientist condemns the use of the scientific method in politics, he often feels that it menaces these traditional operations. It is charged that a scientific approach to politics "takes the fire" out of the subject. The political scientist often maintains that the search for truth cannot be carried out using the scientific method.

Science is not absolute knowledge of truth or falsity. Science is method. Absolute knowledge is only an ultimate goal of science in the same way that pure goodness or happiness or pure anything else is the ultimate goal of man. The practical objective of any scientific study is the production of a bit of knowledge about which we know enough to determine the probability that it is a dependable basis for action in relation to some problem. There is no absolute truth in any body of scientific knowledge at this time and no one living now will see the day the first bit of absolute truth is found. In the physical sciences there is a body of theory of some scope and great detail which has an acceptable contingency factor. In the social sciences there is a body of knowledge of great scope and almost no detail and with the probability that any action taken directly from an analysis of social science theory would backfire. One of the reasons for this is that in social science we have begun with the top of the knowledge hierarchy and laid down the nature of man and society *a priori* and then attempted to verify this while in the physical sciences just the opposite method is used.

This preoccupation with the wrong end of the scientific process as a beginning point is not all wasted. The study of problems of real weight requires a tight theory system and we have developed enough ideas to lay down a satisfactory theory system. Certainly it will be a long time before we shall be able to give a determinate answer to conflicting propositions, but it will not be nearly so long as the eternity it would take to dream up a determinate answer using the orthodox methods of political science research. We need to reverse the orientation of political science research and test the differences between the conclusions arrived at through the analytical method.

The method discussed here assumes that there will be concern with exploratory studies in the first attempt at the scientific method for politics. There must be a method if we are to develop reliable knowledge. The direct transfer of the methods of sociology or psychology or even anthropology will not do except as a starter. These techniques are not sufficient because they were not developed to cope with the central problems of politics. In many narrow areas these methods can be used, of course. This begs the larger problem however.

To the writer, the problems of political science differ from those of the basic social sciences in that they are composites which include individual behavior, group behavior and institutional behavior. This obviously makes the problem of the researcher in political science more complex than the methodological problems of the psychologist or the sociologist or even the anthropologist.

Should many political scientists support this approach it would seem certain that experimentalists who are developing unique research methods for political science will develop also.

BIBLIOGRAPHY

BIBLIOGRAPHY

A. BOOKS

Anderson, William, Research in Public Administration. Chicago: Public Administration Service, 1945. 219 pp.

Andrews, T. G. editor, Methods of Psychology. New York: John Wiley & Son, 1948. 716 pp.

Bartley, S. Howard, Beginning Experimental Psychology. New York: McGraw-Hill Book Co., 1950. 483 pp.

Blankenship, Albert B., Consumer and Opinion Research. New York: Harper & Bros., 1943. 238 pp.

Burt, C., The Factors of the Mind. New York: Macmillan and Co., 1941. 509 pp.

Campbell, Norman, What is Science? London: Methuen and Co., Ltd., 1921. 91 pp.

Chapin, Stuart F., Experimental Designs in Sociological Research. New York: Harper and Bros., 1947. 206 pp.

Churchman, C. W., Theory of Experimental Inference. New York: Macmillan and Co., 1948. 292 pp.

_______, Elements of Logic and Formal Science. New York: J. B. Lippencott Co., 1940. 337 pp.

Cohen, Morris, Reason and Nature. London: Kegan Paul, Trench, Trubner and Co., Ltd., 1931. 470 pp.

Cohen, Morris R., and Ernest Nagel, An Introduction to Logic and Scientific Method. New York: Harcourt, Brace and Co., 1934. 467 pp.

Dewey, John, Logic, the Theory of Inquiry. New York: Henry Holt and Co., 1938. 545 pp.

Dollard, John, Criteria for the Life History. New York: Peter Smith, 1949. 288 pp.

Edwards, Allen L., Experimental Design in Psychological Research. New York: Rinehart & Co., Inc., 1950. 360 pp.

Ezekiel, Mordecai, Methods of Correlation Analysis. New York: John Wiley & Sons, Inc., 1941. 531 pp.

Ferm, Bergilius, A History of Philosophic Systems. New York: The Philosophical Library, Inc., 1950. 642 pp.

Fisher, R. A., Design of Experiment. London: Oliver and Boyd, 1937. 260 pp.

Fry, Luther C., The Technique of Social Investigation. New York: Harper & Bros., 1934. 315 pp.

Garrett, Annette, Interviewing, Its Principles and Methods. New York: Family Welfare Association of America, 1942. 123 pp.

Gaus, John M., Leonard D. White and Marshall E. Dimock, The Frontiers of Public Administration. Chicago: The University of Chicago Press, 1936. 146 pp.

Goodnow, Frank J., Politics and Administration. New York: Macmillan and Co., 1900. 270 pp.

Greenwood, Ernest, Experimental Sociology. New York: King's Crown Press, 1945. 163 pp.

Guilford, J. P., Fundamental Statistics in Psychology and Education. New York: McGraw-Hill Book Co., 1950. 633 pp.

________, Psychometric Methods. New York: Macmillan and Co., 1936. 531 pp.

Holzinger, K. J., and H. H. Harman, Factor Analysis: A Synthesis of Factorial Methods. Chicago: University of Chicago Press, 1941. 417 pp.

Horst, Paul, editor, The Prediction of Personal Adjustment. New York: Edwards Brothers, Inc., 1941. 455 pp.

Kaufmann, F., Methodology of the Social Sciences. Oberlin, Ohio: The Academy Press, 1942. 192 pp.

Kelly, T. L., Statistical Method. New York: Macmillan and Co., 1938. 356 pp.

Larrabee, Harold A., Reliable Knowledge. New York: Houghton Mifflin Co., 1945. 683 pp.

Lundberg, George A., Social Research. New York: Longmans, Green and Co., 1942. 426 pp.

Mill, John I., A System of Logic. New York: Harper & Bros., 1900. 659 pp.

Northrop, F. S. C., The Logic of the Sciences and the Humanities. New York: Macmillan and Co., 1947. 402 pp.

Oldfield, R. C., Psychology of the Interview. Cleveland: The Sherwood Press, 1941. 154 pp.

Parsons, Talcott, Structure of Social Action. New York: McGraw-Hill Book Co., 1937. 817 pp.

Partes, Mildred, Surveys, Polls, and Samples: Practical Procedures. New York: Harper & Bros., 1950. 624 pp.

Peatman, John Gray, Descriptive and Sampling Statistics. New York: Harper & Bros., 1947. 577 pp.

Piper, Raymond F., and Paul W. Ward, The Fields and Methods of Knowledge. New York: Alfred E. Knopf, Inc., 1929. 398 pp.

Pratt, C. C., The Logic of Modern Psychology. New York: Macmillan and Co., 1939. 185 pp.

Stouffer, Samuel A., et al, Measurement and Prediction, Vol. 4. Princeton, New Jersey: Princeton University Press, 1950. 756 pp.

Thurstone, L. L., Multiple-Factor Analysis. Chicago: University of Chicago Press, 1945. 535 pp.

Von Weizsacker, C. F., The History of Nature. Chicago: University of Chicago Press, 1949. 191 pp.

Wallis, Paul, Supplementary Study A. pp. 183-202. Paul Horst, editor, The Prediction of Personal Adjustment. New York: Edwards Bros., Inc., 1941. 455 pp.

Warner, Richard, The Principles of Public Administration. London: Sir Isaac Putnam and Sons, Ltd., 1947. 212 pp.

Westburg, Edward M., Introduction to Clinical Psychology. Philadelphia: T. Balakiston's Son & Co., 1937. 336 pp.

White, Leonard D., The New Social Science. Chicago: University of Chicago Press, 1930. 132 pp.

_______, Trends in Public Administration. New York: McGraw-Hill Book Co., 1933. 365 pp.

Young, Pauline C., and Calvin F. Schmid, Scientific Social Surveys and Research. New York: Prentice-Hall, 1946. 619 pp.

B. PERIODICAL ARTICLES

Abel, Theodore, "The Nature and Use of Biograms," The American Journal of Sociology, 53:111-118, 1947.

Allport, Gordon W., "Scientific Model and Human Morals," Psychological Review, 54:182-191, 1947.

Argyris, Chris, and Gram Taylor, "The Member-Centered Conference and the Research Method, I," Human Organization, Vol. 9, 4:5-14, 1950.

Arrington, Ruth E., "Time Sampling in Studies of Social Behavior: A Critical Review of Techniques and Results with Research Suggestions," Psychological Bulletin, 40:81-124, 1943.

Balc, Albert G., "Relations Between Philosophy and Psychology," Psychological Review, 55:115-125, 1948.

Bierstedt, Robert G., "The Limitations of Anthropological Methods in Sociology," The American Journal of Sociology, 54:22-30, July, 1948.

Biser, Irwin, and Enos E. Witmer, "Methodology of Research and Progress in Science, Philosophy of Science, 14:275-288, 1947.

Boggs, S. Whittemore, "Geographic and Other Scientific Techniques for Political Science," Political Science Review, 42:223-238, 1948.

Bonney, Merle, "A Study of Friendship Choices in College in Relation to Church Affiliation, In-church Preferences, Family Size, and Length of Enrollment in College," Journal of Social Psychology, 29:153-166, 1949.

Boring, Edwin, "Rejoinders and Second Thoughts," Psychological Review, 52:278-281, 1945.

_______, "The Use of Operational Definitions in Science," Psychological Review, 52:243-245, 1945.

Bower, Daniel, "The Problem of Quantification in Psychological Science," Psychological Review, 56:325-333, 1949.

Bray, Douglas W., "The Prediction of Behavior from the Attitude Scale," Journal of Abnormal and Social Psychology, 45:64-84, 1950.

Bridgeman, P. W., "Rejoinders and Second Thoughts," Psychological Review, 52:278-281, 1945.

_______, "Some General Principles of Operational Analysis," Psychological Review, 52:246-249, 1945.

Cantril, Hadley, "The Intensity of an Attitude," Journal of Abnormal and Social Psychology, 41:129-135, 1946.

Cattell, Raymond C., "The Dimensions of Culture Patterns by Factorization of National Characters," Journal of Abnormal and Social Psychology, 44:443-469, 1949.

Churchman, C. W., and R. L. Ackoff, "An Experimental Measure of Personality," Philosophy of Science, 14:304-332, 1947.

Collins, Orvis, Melville Dalton and Donald Roy, "Restriction of Output and Social Cleavage in Industry," Applied Anthropology, Vol. 5, 3:1-14, 1946.

Deri, Susan, Dorothy Dinnerstein, John Harding and Albert C. Pepitone, "Techniques for the Diagnosis and Measurement of Intergroup Attitudes and Behavior," Psychological Bulletin, 45:248-271, May, 1948.

Dickins, Dorothy, "The Problems of Sampling in Connection with the Study of Family Economics," Social Forces, 25:327-332, October-May, 1946-47.

Feigl, Herbert, "Operationism and Scientific Method," Psychological Review, 52:250-259, 1945.

_______, "Rejoinders and Second Thoughts," Psychological Review, 52:284-288, 1945.

"Field Methods and Techniques," Human Organization, Vol. 9, 2:22-34, 1950.

Foreman, Paul B., "The Theory of Case Studies," Social Forces, 26:409-419, October-May, 1947-48.

Fuller, J. L., "Situational Analysis: A Classification of Organism-Field Interaction," Psychological Review, 57:3-18, 1950.

Gengerelli, J. A., "The Dichotomy of Science and Philosophy," Psychological Review, 44:117-137, 1937.

Guilford, J. P., and Andrew L. Comrey, "Predictions of Proficiency of Administrative Personnel from Personal-History Data," Educational and Psychological Measurement, 8:281-296, 1948.

Hadley, John M., "The Relation of Personal Data to Achievement in Radio Training Schools," Psychological Bulletin, 41:60-63, 1944.

Hagood, Margaret J., "Recent Contributions of Statistics to Research Methodology and Sociology," Social Forces, 26:36-40, October-May, 1947-48.

Hall, Evert W., "Some Dangers in the Use of Symbolic Logic in Psychology," Psychological Review, 49:142-169, 1942.

Harper, Robert A., "The Present Status of Questionnaire-Derived Opinion Data," Social Forces, 25:294-297, October-May, 1946-47.

Harris, Marilyn, "On the Determination of Sample Sizes in Designing Experiments," Journal of the American Statistical Association, 43:391-402, 1948.

Hartung, Frank E., "Sociological Foundations of Modern Science," Philosophy of Science, 14:68-95, 1947.

Hempel, Carl G., and Paul Oppenheim, "Studies in the Logic of Explanation," Philosophy of Science, 15:135-175, 1948.

Hertz, Marguerite R., "Rorschach: 20 Years After," Psychological Bulletin, 39:529-572, 1942.

Hillman, Owen N., and Emile Meyerson, "On Scientific Explanation," Philosophy of Science, 5:73-80, 1938.

Hubbard, Henry, "The Elements of a Comprehensive Personnel Program," Public Personnel Review, 1:1-17, 1940.

Hyman, Herbert, "Problems in the Collection of Opinion-Research Data," The American Journal of Sociology, 55: 362-370, January, 1950.

Israel, Harold E., "Two Difficulties in Operational Thinking," Psychological Review, 52:260-261, 1945.

Jacobs, John H., "The Application of Sociometry to Industry," Sociometry, 8:181-198, 1945.

Kattosoff, L. O., "Philosophy, Psychology and Postulational Technique," Psychological Review, 46:62-74, 1939.

Kattosoff, L. W., and J. Thibaut, "Semeiotic and Psychological Concepts," Psychological Review, 49:475-485, 1942.

Kaufman, Harold F., "Member of a Rural Community as Judges of Prestige Rank," Sociometry, 9:71-85, 1946.

_______, "Defining Prestige Rank in a Rural Community," Sociometry, 8:199-207, 1945.

Kay, Lillian W., and Jane H. Schick, "Role Practicing in Training Depth Interviewers," Sociometry, 8:82-85, 1945.

Korner, Annelese, "Theoretical Considerations Concerning the Scope and Limitations of Projective Techniques," Journal of Abnormal and Social Psychology, 45:619-627, 1950.

Krech, David, "Dynamic Systems, Psychological Fields, and Hypothetical Constructs," Psychological Review, 57:283-290, 1950.

Lee, Harold M., "Theoretic Knowledge and Hypotheses," Psychological Review, 57:31-37, 1950.

Lewis, Don, and B. J. Burke, "The Use and the Misuse of the Chi-Square Tests," Psychological Bulletin, 46:433-487, 1949.

Lewis, Oscar, "An Anthropological Approach to Family Studies," The American Journal of Sociology, 55:468-475, March, 1950.

Loevinger, Jane, "The Technique of Homogeneous Tests Compared with Some Aspects of Scale Analysis and Factor Analysis," Psychological Bulletin, 45:507-529, 1948.

Loomis, Charles P., and J. Allen Beegle, "Typological Analysis of Social Systems," Sociometry, 11:147-191, 1948.

Lundberg, George A., and Virginia Beazley, "Consciousness of Kind in a College Population," Sociometry, 11:59-74, 1948.

MacCorquodale, Kenneth, and Paul E. Meehl, "On a Distinction Between Hypothetical Constructs and Intervening Variables," Psychological Review, 55:95-107, 1948.

Manson, G. E., and G. L. Freeman, "A Technique for Evaluating Assembled Evidence of Potential Leadership Ability," Educational and Psychological Measurements, 4:21-33, 1944.

Merton, Robert K., and Patricia L. Kendall, "The Focused Interview," The American Journal of Sociology, 51:541-557, July-May, 1945-46.

Meyer, Herbert H., "Factors Related to Success in the Human Relations Aspects of Work Group Leadership," Psychological Monographs, General and Applied, 65:1-29, No. 320, 1951.

Miller, David L., "The Meaning of Explanation," Psychological Review, 53:241-246, 1946.

Nafe, John Paul, "Toward the Quantification of Psychology," Psychological Review, 49:1-18, 1942.

Pillsbury, Walter B., "Knowledge in Modern Psychology," Psychological Review, 57:328-333, 1950.

Pratt, Carrol C., "Rejoinders and Second Thoughts," Psychological Review, 52:289-290, 1945.

Prescott, Jr., C. H., "The Scientific Method and Its Extension to Systems of Many Degrees of Freedom," Philosophy of Science, 5:237-266, 1938.

Preston, Malcolm G., Emily Hartshorne Mudd, William L. Peltz and Hazel B. Groscher, "An Experimental Study of a Method for Abstracting the Content of Social Case Records," Journal of Abnormal and Social Psychology, 45:628-646, 1950.

Rogers, Carl R., "The Nondirective Method as a Technique for Social Research," The American Journal of Sociology, 50:279-283, July-May, 1944-45.

Sanai, M., "A Factorial Study of Social Attitudes," Journal of Social Psychology, 31:167-182, 1950.

Sargent, Helen, "Projective Methods: Their Origins, Theory, and Application in Personality Research," Psychological Bulletin, 42:257-293, 1945.

Sargent, Howland, "Scientists in Government," Public Administration Review, 2:339-342, 1943.

Schmid, Calvin F., "The Measurement of Public Opinion," Sociology and Social Research, 34:83-90, November-December, 1949.

Seward, John P., "The Sign of a Symbol: A Reply to Professor Allport," Psychological Review, 55:277-296, 1948.

Sedecor, George W., "On a Unique Feature of Statistics," Journal of the American Statistical Association, 44: 1-16, 1941.

Simon, H. O., and William R. Divine, "Controlling Human Factors in an Administrative Experiment," Public Administration Review, 1:485-490, 1940.

Skinner, B. F., "The Operational Analysis of Psychological Terms," Psychological Review, 52:270-277, 1945.

Slone, Eugene H., "Reductionism," Psychological Review, 52:214-223, 1945.

Smucker, Orzen, "Prestige Status Stratification on a College Campus," Applied Anthropology, 3:20-26, 1944.

Solomon, Richard L., "An Extension of Control Group Design," Psychological Bulletin, 46:137-150, 1949.

Stene, Edwin, "An Approach to a Science of Public Administration," Political Science Review, 34:1124-1137, 1940.

Stephan, Frederick F., "History of the Uses of Modern Sampling Procedures," Journal of the American Statistical Association, 43:12-39, 1948.

_______, "Sampling," The American Journal of Sociology, 55:371-375, January, 1950.

Steward, Frank A., "Sociometric Testing at the Adult Level," Sociometry, 9:147-148, 1946.

Stoghill, Ralph M., "The Sociometry of Working Relationship in Formal Organization," Sociometry, 12:276-286, 1949.

Weber, Christian O., "Valid and Invalid Conceptions of Operationism in Psychology," Psychological Review, 49:54-68, 1942.

Wherry, Robert J., and Douglas H. Fryer, "Buddy Ratings: Popularity Contests, Leadership Criterion?" Sociometry, 12:179-190, 1949.

Whyte, William Foote, "A Challenge to Political Scientists," Political Science Review, 37:692-697, 1943.

Wilson, M. L., "Basic Social Science Applied to Administration," Public Administration Review, 5:386-391, 1945.

Wilson, Woodrow, "The Study of Administration," Political Science Quarterly, 2:197-222, 1887.

Witzemann, Edgar J., "The Scope, Objectives and Limitations of Modern Science as Seen in the Light of its History," Philosophy of Science, 14:44-55, 1947.

Yule, G. U., "Critical Notice on William Brown and Dodgrey H. Thomson's The Essentials of Mental Measurement," British Journal of Psychology, 12:100-107, 1921-22.

C. PARTS OF SERIES

Urwick, L., and E. F. Brech, The Making of Scientific Management, Vol. I., Thirteen Pioneers. London: Management Publications Trust, 1949. 196 pp.

_______, The Making of Scientific Management, Vol. II, Management in British Industry. London: Management Publications Trust, 1949. 241 pp.

Urwick, L., and E. F. Brech, The Making of Scientific Management, Vol. III, The Hawthorne Investigation. London: Management Publications Trust, 1949. 225 pp.

D. PUBLICATIONS OF LEARNED ORGANIZATIONS

Committee on Public Administration, The Governmental Research Bureau. New York: Social Science Research Council, preliminary draft, 1938. 304 pp.

Fesler, James W., et al, The University Bureaus of Public Administration. University: University of Alabama, Bureau of Public Administration, 1946. 87 pp.

Gaus, John M., A Study of Research in Public Administration. New York: Social Science Research Council, preliminary draft, 1930. 43 pp.

Grahm, George A., Education for Public Administration. Chicago: Public Administration Services, 1941. 366 pp.

Simon, Herbert A., Measurement Techniques in Administrative Research. Civil Affairs Pamphlet, Vol. 8. 1941. 8 pp.

Spicer, George W., et al, A Critical Appraisal of the University Bureaus of Public Administration. Charlottesville, Virginia: Bureau of Public Administration, University of Virginia, 1948. 55 pp.

Thomas, Dorothy S., Alice M. Loomis, Ruth E. Arrington, Observational Studies of Social Behavior. Institute of Human Relations, Yale University, 1933. Monograph. 273 pp.