PB178335

November, 1967

CHOICE, COMMUNICATION, AND CONFLICT

A System's Approach to the Study of Human Behavior

Russell L. Ackoff
Management Science Center
Wharton School of Finance and Commerce
University of Pennsylvania
Philadelphia, Pennsylvania

The National Science Foundation supported the research herein under Grant GN-389.

FORWARD

This work was prepared in connection with the first of two grants from the National Science Foundation. It presents the conceptual system developed under that that grant. Applications of this system being carried out under a second grant and other sources of support will be reported on in a companion volume which will be a continuation of this one.

PREFACE

The thinking of which this work is a product began in 1941 and for many years was carried on jointly with C. West Churchman, Thomas A. Cowan, and Leon Pritzker. Their contribution has been a major one. For this I am very grateful. Unfortunately, however, I cannot hold them responsible for the errors and misconceptions that are contained herein.

I have also been greatly assisted by Anthony C. Scoville and the many students over a quarter of a century who would not take anything for granted.

Russell L. Ackoff
Philadelphia
August 2, 1967

TABLE OF CONTENTS

Chapter 1

PROLOGUE: ON HUMAN BEHAVIOR AS A SYSTEM

> DELIBERATION, n. The act of examining one's bread to determine which side it is buttered on (Ambrose Bierce, The Devil's Dictionary).

This book presents neither a theory of behavior nor a set of generalizations that explain why people do what they do. Nor does it describe their behavior. Nor is it another of the increasing number of efforts to mathematize or formalize the study of human behavior. What this book does attempt to do is provide a way of looking at human behavior as a system.

Why bother to do this? I have been struck for some time by the fact that in an era that is so systems-oriented--an era in which we are becoming increasingly more interested in wholes than in their parts--that human behavior is still conceived, observed, analyzed, experimented on, and otherwise treated in a piecemeal way.

Human behavior is studies by psychologists, social psychologists, anthropologists, sociologists, psychiatrists, philosophers, and others; and within each of these disciplines there are points of view as distinct and disparate as there are between the disciplines. For example, among psychologists there are those who study only perception, or conception, or traits, or attitudes, or learning, or communication, and so on. Most psychologists make little or no effort to relate their work to that of others outside their area of specialization but in their discipline, let alone to work in other disciplines. As a result, we have a very large number of very thin slices made through the sphere of human behavior, but nothing approaching a conception of it in the round. Consequently, I try in this book to provide a system of concepts in terms of which all aspects of human behavior can be interrelated.

What I have said of the study of human behavior in the large is also true of it in the small. For example, consider the study of human communication. I began this work intending to deal only with this subject. There was no better place to begin than with Colin Cherry's <u>On Human Communication</u> (1957). In reading this book I was impressed both by the large number of ways in which human communication has been studied and by the almost complete unrelatedness of these, despite Cherry's considerable effort to pull them together.

Alfred Smith's more recent effort (1966) is not much more successful. Using two classification schemes, he attempts to relate a large number of individually useful contributions to our understanding of human communication. These schemes allow him to organize his selections but not to relate or synthesize the findings contained in them.

Communication itself has been divided into smaller and smaller subsystems; for example, coding, transmission, data processing, storage and retrieval, indexing, and so on. Not only is human communication too much of a system to treat its parts in isolation from each other, but it is also too much of a <u>sub</u>system to treat it fruitfully in isolation from other aspects of human behavior. Put another way, in order to develop a system of concepts that would relate the wide variety of studies of human communication, I found it necessary to imbed such a system in a more inclusive one involving all of purposeful behavior.

Although there has been a great deal of attention paid by a wide variety of scientists to the "systems point of view," much of the literature is frustratingly vague on the meaning of this expression. Fortunately, E. A. Singer, Jr. (1959) has provided a comprehensive and clear formulation of such a point of view in what he called an "objective teleology." "Teleology" has traditionally been a naughty word in science, to a large extent because of the way Aristotle used it centuries ago.

Many of the Greek philosophers, including Aristotle, invoked the concept of purpose to explain a wide variety of observable phenomena, but they never explained purpose itself in terms of observable phenomena. During the Renaissance philosophers and scientists alike reacted against Aristotle's point of view; they deserted the concept of purpose and replaced it with a mechanistic (ateleological) conception of the universe. By the nineteenth century many believed that all of nature could be explained mechanistically.

The preoccupation with mechanism directed analysis to the way things were <u>structured</u>: the material of which they were made and the forces that governed their behavior. This point of view led to a dissection of things and events into their smallest parts: atoms, molecules, electrons, quantum jumps, and so on. Mechanistically oriented research takes things apart, analyzes the parts, and tries to put them back together again, often unsuccessfully. In the past, knowledge of the whole has almost always been derived from knowledge of the parts.

The mechanistic approach was fruitful but its insufficiency became increasingly apparent early in this century. Perhaps no one made as strong a case for a teleological approach to research that was complementary (not antithetical) to mechanism, as did Singer. He began early in the century to develop the basis for studying purposeful behavior in an objective and experimental way. (See Singer [1924] in which some of his early essays are reprinted.) His motivation derived from recognition of the failure of the mechanistic point of view to deal adequately with the nature of life, mind, society, and values. The fruitfulness of his point of view was not generally recognized until it was stated in another context by Rosenblueth and Wiener (1943 and 1950) who, as far as I know, were unaware of Singer's work.

Rosenblueth and Wiener, and later Wiener (1961), began to see the fruitfulness of looking at mechanisms as functional entities. They became preoccupied with how mechnisms functioned and how and why they kept doing so. Their concern was with mechanisms that served a function, teleological mechanisms. They found it more useful to proceed conceptually from function to structure than from structure to function. (Singer had shown that "structure" itself was a functional concept.)

Prior to the work of Rosenblueth and Wiener, designers of mechanisms tended to get their conception of the whole by composing parts. Since then, designers increasingly tend to get their conception of the parts by decomposing the whole. They derive the structural characteristics of necessary parts from the functional characteristics of the whole. This functional (or teleological) orientation gave rise to what has come to be known as the systems point of view.

Put another way, before the revolution in thought which made it possible to use teleological concepts as a methodological key to open doors previously unlocked by science, scientists tended to derive their understanding of the functioning of the whole from the structure of the parts and the structural relationships between them. Today we increasingly tend to derive our understanding of the structure of the parts of a system from an understanding of the functioning of the whole.

In this book I take a holistic and functional point of view of human behavior. Yet, following Singer, and Rosenblueth and Wiener, I have tried to make all the functional concepts employed as objective, as measureable, as capable of use in experimentation, as any structural concepts that have been produced by the mechanistic point of view.

The objective teleology which is developed here is not intended to replace the objective _a_teleology (mechanism) which preceded it; it is meant to supplement it. Following Singer I shall try to show that the mechanistic and teleological points of view are completely compatible, or, as Neils Bohr said, there is a _complementarity_ between them. I only argue that the mechanistic point of view is not as fruitful as is the teleological in the study of human behavior.

Let me consider the characteristics of an objective teleology in more detail. Centuries ago Aristotle invoked teleological concepts to explain why (inanimate as well as animate) things behaved as they did. Among those who carry on in his spirit on the contemporary scene are some psychologists, for example, who try to explain human behavior by invoking such concepts as "beliefs," "attitudes," and "traits," let alone "instincts" and "drives." To do so is to employ a _subjective_ teleology. In an objective teleology the converse is done: beliefs, attitudes, and traits are attributed to an individual because of what he does; these properties are derived from perceived regularities of behavior under specified conditions. Such concepts do not lie behind behavior; they lie _in_ behavior. Hence, in an objective teleology functional characteristics of human behavior are not treated as intervening variables which are subjectively fabricated to conceal our ignorance; they are objectively derived from what we can observe.

The objective part of objective teleology refers not only to the derivation of functional properties from observable behavior, but also to the fact that the observations involved are reproducible by different observers. Introspection is not required. This opens the study of the "inner workings" of the mind to public examination. In order to accomplish this it is necessary to provide operational definitions and measures of functional concepts, definitions which provide standards in the same sense in which the ateleologically oriented sciences provide standards for structural concepts (e.g., length, density, and energy in physics).

An operational definition of a concept provides a standard if it consists of an explicit specification of the conditions under which, and the operations by which, questions concerning the concept ideally ought to be answered. Even though it may be difficult or impossible to meet the specifications contained in such a definition, they serve an important scientific purpose. They make it possible to compare observations made relative to the same concept but under different sets of conditions. Such observations can be adjusted back to the standard. That is, however research involving a concept is conducted, inferences should be drawn from what was observed to what would have been observed if the idealized specifications contained in the standard had been met. In order to make such inferences it is necessary to formulate explicitly how the conditions under which observations were made differ from those specified in the standard, and to employ appropriate theory to adjust the observations for the effects of these differences. For example, in the idealized conditions formulated in physics for measuring the length of an object, the temperature of the environment in which observations ideally should be made, is specified. If the temperature under which observations actually are made differs from that specified, then the coefficient of linear expansion that is appropriate to the object measured can be employed to adjust the observations. Analogous coefficients and theories on which to base them are rare in the behavioral sciences. The formulation and use of definitional standards points up the need for theories which can be used to adjust data. Without the ability to do so, different researchers on the same subject cannot effectively compare their work, and without the ability to relate different studies of the same thing, results do not build up cumulatively as rapidly as they should.

A standard is idealized relative to our current state of knowledge; it is neither immutable nor absolute. Hence, as our understanding of a concept increases, our formulation of how it ought to be observed and measured changes. This has been the case, for example, with respect to "length." Therefore, at this stage in the development of the behavioral sciences it is not necessary to develop ultimate (or even lasting) definitional standards, but to provide some standards. I cannot hope to provide operational definitions of behavioral concepts that are generally acceptable, but I do hope to provide ones that will provoke constructive discussion that will lead to their rapid improvement.

In developing the content of the definitions that I offer I have tried to take into account both historic and current usage, but usage is frequently confused, obscure, ambiguous, and inconsistent. No one person can arbitrate such conflict but he can hope to provoke work that will reduce it.

The absence of operational definitions and conceptual standards in the behavioral sciences has resulted in much trivial research, perhaps more than is apparent because triviality can easily be concealed by obscure terminology. Charges such as the following by William Gomberg (1966) have seldom been levelled against other areas of science:

> Recently Berelson and Steiner wrote an inventory of scientific findings on human behavior that attempts to summarize those aspects of human behavior that are entitled to the honorific

term "scientific"...*

As the pages of the book are reviewed, what is most striking is the banality of its "scientifically established findings..."

The fruitfulness of their investigation is hobbled because they have failed to distinguish what is needed for a description of social nature from their self-imposed rituals. They have engaged in a decision making ritualistic prescription for scientists to act in certain ways rather than in others. Cowan portrays the corner into which the behavioral scientists have painted themselves beautifully.

> The teleology of decision making is more powerful than its logic in shaping the course of decision; intuition has a more important role to play in even simple and apparently trivial decisions than the rational constraints of present-day decision procedures allow. It seems to me than every true decision, as distinct from an inference, involves an element of individual choice, the constraints imposed by general logic and generalizing mathematics upon decision procedures virtually rule out the study of truly creative decisions and tend to restrict decision science to mechanical, and, therefore, dull and repetitive instances of decision making. **

Professor Henry is even rougher with Berelson and Steiner. He states that the book ought to be called "The Nature of Intellectual Failure in the Behavioral Sciences." He charges the entire field with:

1. An inability to distinguish truism from discovery
2. Insensitivity to platitude
3. Insensitivity to tautology
4. Confusion of causal sequences
5. The delusion of precision
6. The drawing of simple minded parallels***(pp. 9-11).

This quotation reflects how some observers of the behavioral

---

*Berelson, Bernard, and Steiner, Gary A., Human Behavior, An Inventory of Scientific Findings. Harcourt Brace, New York, 1964.

**Cowan, T. A., "Decision Theory in Law, Science and Technology," Science, June 7, 1963.

***Henry, Jules, "Revue of Human Behavior," Scientific American, July 1964, pp. 129-133.

sciences view the results of applying mechanistic and unsystematized concepts, and the methodology derived from them, to the subject of human choice. These concepts and methodology have dictated the kind of studies that have been carried out. But I am not as concerned with the past as I am with the future: with the kinds of study of human behavior that ought to be conducted and with developing the concepts and methods which might make them possible. I hope to show that an objective teleology expressed in the form of a conceptual system can serve as a foundation for significant research into such phenomena as choice, communication, and social interaction; the three interrelated aspects of human behavior to which special attention is given in this book.

The kinds of operational definitions of functional concepts that are developed here suggest general and rich hypotheses about human behavior. Furthermore, they provide a basis for designing adequate tests of these hypotheses. I shall try to support these claims in subsequent chapters.

The point of view taken in this book is derived from the conviction that the principal function of the philosophy of science is to open to scientific investigation types of phenomenon not previously considered to be suited to such inquiry. Historically each branch of science emerged out of philosophical analysis of its subject matter and methods of inquiring into it. This historic role of philosophy--which is the basis for calling it "the mother of the sciences"--can now be carried out with considerably more sophistication than was possible previously because we now have a much better understanding of what science is than our predecessors had. It is possible, therefore, to open up new areas of inquiry, such as will be attempted here, in a rigorous scientific way. Whether such an effort is best called philosophical, scientific, or methodological is a matter of personal preference.

Philosophy has traditionally had another role in science. In the nineteenth century it was commonly believed that its principal function was to synthesize the findings of the various scientific disciplines into one cohesive body of knowledge about natural phenomena. This view was epitomized in the encyclopedic work of Herbert Spenser who attempted to unify science around the evolutionary concept. More recent efforts along these lines have been made by Richard L.Schanck (1954) using the concept of "dynamic equilibrium" and Ludwig von Bertalanffy (1951) and his followers who use the concept of "structural isomorphisms" in the development of General Systems Science.

The need to synthesize findings in the many disciplines of science arises out of the fact that these disciplines have been developed with relatively unrelated conceptual systems. Scientists have carved up the world into smaller and smaller pieces and have created disciplines specializing in each. As disciplines multiply they increase in depth and decrease in breadth. It is estimated that no man has been able to "know everything" since the beginning of the eighteenth century. In brief, the reason there is need to put knowledge of our world together into one cohesive view derives from the fact that it was necessary to take it apart in order to penetrate it in depth.

Nature does not come to us in disciplinary pieces. The disciplines emerge out of points of view, out of how things are looked at, not out of the nature of things. We have broken our concept of nature, like Humpty Dumpty, into bits and pieces and, like all the king's horses and all the king's men, we are having trouble putting it back together again.

Singer tried to see the "whole picture" and show the relationship between the various disciplinary points of view. He observed that if we conceive of science as a system of related points of view we do not have the task of reassembling the view points. Hence, the task of synthesis taken on by some philosophers derives from the confusion of taking the results of disciplinary analysis as the starting points of experience, rather than taking holistic experience as the starting points of disciplinary analysis.

The kind of analysis required to provide a holistic view of nature and science is not the vague speculative type that we have come to associate with most of "professional" philosophy. It must be an analysis that is operationally oriented, one that is directed toward providing science with concepts, measures, and methods that are prescriptive; that is, with instructions, not merely inspiration. In another place (Ackoff, 1962) I have tried to describe in detail how concepts, measures, and methods should be developed in science. The effort here attempts to exemplify that earlier work.

One of the principal hopes behind this effort is that it will facililitate consideration of behavioral variables in the evaluation and design of organized social systems, including those involving machines as well as men. The models of such systems used in system science, management science, operations research, systems engineering, and other systems-oriented inter-disciplines, frequently contain behavioral variables. But these variables are almost always treated ateleologically rather than teleologically. For example, in the study of service-processes (e.g., check-out counters at supermarkets or toll booths at bridges and tunnels) the arrival rate of customers and the service rate of servers are important variables but there is nothing particularly human about the way they are treated. This is not to say that in studying queues the behavioral variables should be treated teleologically, but it is to say that in those processes where it would be desirable to do so, it is seldom done. For example, in models of most communication, advertising, and marketing processes, people's responses are treated statistically at best, not as outputs of individual decision processes. In general we tend to treat behavior collectively, leaving the resulting statistic unexplained and hence do not increase our understanding and potential control over the process under study. To predict behavior is not enough; we must explain it.

For example, even very significant correlations between alcoholism and socio-economic characteristics do not explain this disease and do not help prevent or cure it. Accident statistics and knowledge of associated characteristics of bad drivers does not help us prevent accidents.

Human beings are typically treated by systems researchers as statistics-generating machines, or as entities which respond to stimuli in a mechanical way. In some cases the human is completely excluded. This is reflected in Claude Shannon's (1949) exclusion of the human communicator in his model of the communication process. This is not meant to belittle his contribution, but to point up the need for bringing human purposes into the study of phenomena involving human behavior.

To improve communication processes we must understand <u>why</u> individuals choose to communicate in the way they do. We cannot start our analysis with messages that humans have produced; we must begin with the process by which they are produced. This is a matter of <u>choice</u>. Choice must be an integral part of any complete model of communication.

Finally, I want to caution the reader that what is attempted here is not primarily intended to provide systems scientists and engineers with additional quantitative tools and techniques to put into their kit, but to provide them with a new kit into which old and new techniques and tools can be placed. I try to provide a new way of thinking about and dealing with behavioral variables. I seek to provide more than indices of ill-defined behavioral variables; I seek to provide measures of ones that are well-defined. I will not make it easy to deal rigorously and objectively with the richness, subtlety, and complexity of human behavior, but if I succeed I will have helped make it possible.

BIBLIOGRAPHY

Ackoff, R. L., Scientific Method: Optimizing Applied Research Decisions. John Wiley & Sons, New York, 1962.

Bertalanffy, L. von, "Problems of General System Theory," Human Biology, 23(1951), 302-312.

__________, "Conclusion," Human Biology, 23(1951), 336-345.

Cherry, Colin, On Human Communication. John Wiley & Sons, New York, 1957.

Frank, L. K., et al, "Teleological Mechanisms," Annals of the New York Academy of Science, 50(1948).

Gomberg, William, The University and Business Education. Working Paper No. 16, Department of Industry, Wharton School of Finance and Commerce, University of Pennsylvania, March 1966.

Rosenblueth, A., and Wiener, N., "Behavior, Purpose, and Teleology," Philosophy of Science, 11(1943), 18-24.

__________, "Purposeful and Non-purposeful Behavior," Philosophy of Science, 17(1950).

Schanck, R. L., The Permanent Revolution in Science. Philosophical Library, New York, 1954.

Shannon, C. E., and Weaver, W., The Mathematical Theory of Communication. The University of Illinois Press, Urbana, 1949.

Singer, E. A., Jr., Mind as Behavior. R. G. Adams, Columbus, 1924.

_____, Experience and Reflection. Ed. by C. W. Churchman. University of Pennsylvania Press, Philadelphia, 1959.

Smith, A. G., Communication and Culture. Holt, Rinehart and Winston, New York, 1966.

Wiener, N., Cybernetics. John Wiley & Sons, New York, (2nd ed.) 1961.

CHAPTER 2

STRUCTURE, FUNCTION, PURPOSE, AND CHOICE

DECIDE, v.i. To succomb to the preponderance of one set of influences over another set (Ambrose Bierce, The Devil's Dictionary).

INTRODUCTION

The development of Cybernetics, Information Theory, Communications Engineering, Computer Science, General Systems Theory, Systems Engineering, Operations Research and related scientific and engineering efforts have brought with them a new respectability for such teleological concepts as function, purpose, choice, and communication. They have shown the fruitfulness of conceiving of at least some phenomena in other than a mechanistic framework such as dominated the scientific thought of the nineteenth and early part of this century.

Whenever a set of concepts such as "purpose" and "communication" become critical in many different fields of science and technology, there is a danger that their definitions become oriented to the special interests of their formulators. This restricts their applicability to other types of study. The process goes somewhat as follows: Cyberneticians define "purpose" and "information" so that they are admirably suited to the types of study in which they are engaged. They then suggest that these definitions are equally applicable in other fields. For example, some cyberneticians believe the concept of "purpose" as used in their field is equally applicable in the behavioral sciences. Some psychologists and sociologists, however, realize that the phenomena with which they are concerned are not captured in the Cybernetician's definition and, hence, they look at his offerings simply as metaphors or analogies. This has led some behavioral scientists to ignore work that could at least be very suggestive to them. Others have taken literally the definitions offered by the Cyberneticians and have produced analyses of human behavior that most behavioral scientists feel miss the richness and subtlety of human behavior. Both tendencies mitigate against interdisciplinary studies of human behavior such as Wiener (1961) called for.

To be more specific, consider Rosenblueth's and Wiener's formulation of "some criteria for the distinction between purposeful and nonpurposeful behavior" (1950). These criteria were all concerned with establishing some connection between the purposeful object and its environment and goals. Thus, for them, the purposeful object must be "coupled to" certain features of the environment, as well as "oriented to and guided by" the goal. Tests of purpose must be made by changing the environmental conditions, and so on. The general idea is that an object behaves purposefully if it continues to pursue the same goal by changing its behavior as conditions change.

Although this concept can be applied to some behavioral problems, it clearly cannot be applied to all. For example, the psychologist Kohler observed that simians learn by trial and error how, by use of instruments, to obtain food that is placed out of their reach. Such animals were observed in unchanging environments and yet their actions would generally be regarded as purposeful. Or again, on the social level, a government agency may, under unchanging conditions, try many different tactics to get enacted some legislation that it desires. This too would generally be regarded as purposeful behavior.

In effect, Rosenblueth and Wiener found a useful concept, goal-directed activity in the study of mechanisms. But it is ill-advised to assume this concept captures all the meaning of purpose in human (or even machine) behavior. Finer distinctions are required, as I shall try to show below.

What is needed is a system of concepts and measures which incorporates the findings of Cybernetics and yet which is rich enough to encompass the concerns of the behavioral scientist, psychological or social. A system of concepts which is designed primarily to handle one type of datum (such as that arising in the study of servo-mechanisms or the transmission of messages over wire) runs the risk of being useful in only a metaphorical sense when the type of phenomenon under study changes. Hence, it is critically important to develop a system of teleological concepts (including "purpose," "communication," "information," and others) which is general enough to cover inquiries into many types of phenomena by many different disciplines.

THE CONCEPT OF STRUCTURE

The meaning of "purpose" depends on the meaning of "function" and "function" is used throughout this book in contrast with "structure." "Structure" is a very general concept that includes geometric, kinematic, mechanical, physical, and morphological concepts. Therefore, I treat these aspects of structure first, then derive the meaning of "structure" from them.

Euclid's geometry begins with a set of concepts and properties of which the most elementary are "point" and "line." Concepts of other geometric entities and properties are built up out of these basic building blocks.

2.1. <u>Geometric Class</u>: Two or more sets of geometric points which have one or more geometric properties in common can be said to form a geometric class.

Whether or not such sets are said to be members of the same geometric class depends on whether the property or properties that they have in common are of interest to the one doing the classification. Two sets which are alike with respect to a property of no interest to the investigator, but which differ with respect to another property that is of interest to him will not be said to be members of the same class.

Mechanics, like geometry, begins with certain basic concepts: for example, in classical mechanics these were the Euclidian three dimensional space-coordinate system, a time coordinate, and two mechanical properties: mass and acceleration.

2.2. Mechanical Point: a point which has geometric (spatial), kinematic (temporal), and basic mechanical properties.

In classical mechanics, such points were called "point-particles." The exact nature of these particles (i.e., be they atoms, molecules, electrons, or what not) is not relevant to the concept.

2.3. Mechanical Class: sets of equal numbers of mechanical points whose corresponding members have one or more mechanical properties in common.

2.4. Physical Individual: a set of two or more mechanical points which occupies a geometrically definable volume over a specified period of time.

The objects with which we deal daily, are therefore, physical individuals.

2.5. Physical Environment of a Physical Individual: that part of a specified volume in space at a moment of time which is not occupied by a physical individual contained within it.

2.6. Physical Property: a property of a physical individual which can be expressed as a function of the geometric kinematic, and basic mechanical properties of the mechanical points of which the physical individual is composed.

For example, the temperature of an object is one of its physical properties because it can be expressed as the mean squared velocity of its point particles. Similarly, the mass of a body is equal to the sum of the masses of its point particles.

2.7. Physical Class: Two or more physical individuals which have one or more physical properties in common can be said to be members of the same physical class.

Note that two sets of mechanical points, alike in all respects except their location, must have the same physical properties. But two bodies with the same physical property need not consist of sets of mechanical points in the same mechanical class. For example, two sets of differing numbers of mechanical points in which no pair of points, one from each set, have the same mechanical properties, may nevertheless form physical individuals with the same temperature or mass.

2.8. Morphological Property: Let v represent a value on a scale used to measure a physical property and let k represent some non-zero value on that scale. Then $v \pm k$ defines a morphological property.

Morphological properties are the ones with which the physical sciences usually deal. For example, when we say two bodies have "the same temperature" we do not usually mean "exactly the same temperature," but that their temperatures fall within some specified interval; for example, $70 \pm 0.5°$ F. The size of the interval used depends on our purposes. For some purposes we may want to consider as the same two bodies whose temperature falls within the same 10° F interval; for others, a 1° F interval may be required. When we classify people by age, each class is based on a morphological property. Here too the size of the interval will vary with our purposes. For our purpose it may be sufficient to consider only minors and adults (e.g., in determining who can purchase alcoholic beverages), for another, age at the nearest birthday (e.g., in the census).

2.9. Morphological Class: two or more physical individuals which have one or more morphological properties in common.

Note that two physical bodies with the same physical property must have the same morphological property defined on the scale employed to measure that physical property. Clearly, however, two bodies with the same morphological property need not have the same corresponding physical property.

2.10. Structural Property: any geometric, kinematic, mechanical, physical, or morphological property.

2.11. Structural Class: two or more physical individuals which have one or more structural properties in common.

Thus "structure" is a general concept applicable to geometric, kinematic, and mechanical properties, and any properties which can be expressed as functions of them.

2.12. Structural Behavior of a Physical Individual: a change in one or more structural properties of a physical individual.

For example, when an object "falls" it changes its location. When a body "cools" its temperature changes. A change of an object's properties may also be called an event.

THE CONCEPT OF FUNCTION

The meaning of "function" depends on the meaning of "causality" and the latter has had at least two distinct meanings in science. These meanings are reflected in John Struart Mill's (1862) first two Canons of Induction. In the first, the Method of Agreement, Mill attempted to specify how to determine whether one thing was sufficient for another. In the second canon, the Method of Difference, he attempted to specify how to determine whether one thing was necessary for another. The two types of causality defined below are based on this distinction.

2.13. Deterministic Cause: An object, or its structural behavior, or a structural property of either (hereinafter referred to as a "thing" or "X") in the object's environment (S_1) is the deterministic cause of another thing (Y) in its environment (S_2) if the first is necessary and sufficient for the second.

2.14. Probabilistic Cause or Producer-Product: A thing (X) in its environment (S_1) is the probabilistic cause or producer of another thing (Y, the product) in its environment (S_2), if the first is necessary but not sufficient for the second.

The two environments in these definitions (S_1 and S_2) need not be distinct, but they may be; X may cause a Y in its own or another environment. Furthermore X and Y may be the behavior or property of the same object; a thing may cause something to happen to itself; for example, a person may hurt himself.

The last two definitions presuppose an understanding of the concepts "necessary" and "sufficient." Let me make explicit the sense in which I use these concepts. I will use "$\rightarrow$" to represent "is always followed by." Now, suppose a thing (X) in its environment (S_1) is always followed by another thing (Y) in its environment (S_2); that is

$$X \text{ in } S_1 \rightarrow Y \text{ in } S_2 .$$

Then X in S_1 is sufficient for Y in S_2.

Now let X′ and Y′ represent the nonoccurrence of X and Y. If

$$X \text{ in } S_1 \rightarrow Y \text{ in } S_2$$

and

$$X' \text{ in } S_1 \rightarrow Y' \text{ in } S_2$$

then X in S_1 is both necessary and sufficient for Y in S_2. Y occurs in S_2 only if X occurs in S_1, and always occurs if X does. Hence X in S_1 completely determines the occurrence of Y in S_2. For example, if we can construct an environment in which striking a bell (X) is always followed by a ringing of the bell (Y), then striking the bell in that environment is the deterministic cause of its ringing. Note that whether a phenomen is or is not a deterministic cause depends on how we define or construct it and its environment. For example, we can easily construct an environment (e.g., one virtually without air) in which striking a bell is not sufficient for making it ring. In such an environment striking the bell is necessary but not sufficient for the ringing. Air is also necessary but not sufficient for the ringing.

Now suppose that two things, X_1 and X_2 (e.g., striking the bell and air) are jointly necessary and sufficient in an environment (S_1) for the subsequent occurrence of Y (e.g., ringing of the bell) in the same or another environment (S_2); that is,

$$X_1 \text{ and } X_2 \text{ in } S_1 \rightarrow Y \text{ in } S_2$$
$$X_1' \text{ and } X_2 \text{ in } S_1 \rightarrow Y' \text{ in } S_2$$
$$X_1 \text{ and } X_2' \text{ in } S_1 \rightarrow Y' \text{ in } S_2$$
$$X_1' \text{ and } X_2' \text{ in } S_1 \rightarrow Y' \text{ in } S_2$$

Then X_1 is necessary but not sufficient for Y, and hence is a probabilistic cause or producer of Y. If X_1 occurs in S_1, Y will or will not occur in S_2 depending on whether or not X_2 occurs in S_1. Note that we need not know what are all the necessary conditions for a product (Y) in order to determine that a particular thing (X) is necessary for it. If

(1) Y only occurs in S_2 when X occurs in S_1,

(2) Y never occurs in S_2 if X has not occured in S_1, and

(3) the nonoccurrence of Y in S_2 does not imply that X_1 has not occurred in S_1,

then X can be said to be the producer of Y.

Up to this point I have treated X and Y as though they were specific objects, events, or properties, but they may be considered as members of a structural class of objects, events, or properties. For example, "striking a bell" may be taken as any behavior of a structurally specified class, and so may "ringing of the bell." It is in this sense, for example, that we speak of acorns as producers of oaks. A member of the class of acorns is necessary for a member of the class of oaks. We may also refer to a specific acorn as the producer of a specific oak.

Since not every acorn produces an oak, but some do, we refer to acorns as "potential" producers of oaks.

2.15. Potential Production: All the members of a structural class of things can be said to be potential producers of members of another structural class of things if one or more members of the first class has produced a member of the second class.

Salmon eggs are potential producers of salmon's and robin's eggs are potential producers of robins, but the probabilities of production in these two cases are quite different.

2.16. Probability of Production: (a) The probability that a thing (x) which is a member of a structural class (X) in an environment which is a member of a structural class of environment (S_1) will produce a thing (y) which is a member of a structural class (Y) in an environment which is a member of a structural class of environment (S_2) is the limiting relative frequency with which x's in S_1's produce y's in S_2's.

(b) The probability that a particular individual (x) in an environment which is a member of a structural class of environments (S_1) will produce a thing (y) which is a member of a structural class (Y) in an environment which is a member of a structural class of environments (S_2) is the limiting relative frequency with which that individual (x) in S_1's produce y's in S_2's.

Therefore, the probability that an acorn in a particular type of environment will produce an oak is equal to the limiting relative frequency with which acorns in such environments produce an oak. The probability that a particular clock will strike twelve in a particular environment is the limiting relative frequency with which that clock strikes twelve at twelve o'clock in that environment.

The question concerning the probability that a particular thing will produce another arises only because of the uniqueness of that thing. If it were considered as a member of a class, and hence not unique, its probability of production would be determined by virtue of its class membership (2.16a). If the relevant properties of the thing in question (those that affect its capability for production) remain constant over time then that thing at various times in the same kind of environment constitutes the class with respect to which probabilities are determined (2.16b). However, if these relevant properties change over time it becomes more difficult to determine its probability of production.

Consider, for example, a cigarette lighter which wears with use. In the first 100 tries it may light 100 times; in the second 100 tries it may light 90 times; in the third 100 tries it may light 80 times. If we knew this and wanted to estimate its probability of producing a flame on a try beyond the 300^{th} try, common sense indicates 0.7, but this is not its limiting relative frequency, which is approximately zero. Therefore, if the lighter's probability of producing a flame is a function of amount of previous use it has had, we must take its previous use into account in determining this probability. If this lighter is not significantly different from others with the same amount of use, then we can revert to determination of its probability of production on the basis of class membership which is based on usage. If, however, it differs from other members of its class with respect to a relevant property then the probability of production of class members can be used as a base, but it must be adjusted for the difference in probability produced by the difference in the relevant property. For example, if this lighter has a different fuel in it than other lighters with the same usage, then the probability of production of fire by members of the relevant class of lighters must be adjusted for the effect of the fuel. Therefore, we must conceptually construct a a class of things similar to the unique one in question and infer the limiting relative frequency of its production from what we can observe about available things similar to the one in question. Notice that inference is required even where we do not have to adjust observations because the limiting relative frequency itself is never observed, but is inferred from a finite number of observations.

The concept of production is used extensively in the pure and applied physical sciences as well as in the behavioral sciences. For example, the reliability measures that are used in engineering are fundamentally probabilities of production or non-production. The reliability of a generator, for example, can be measured as the probability that "turning it on" (X) under specified conditions (S) will produce electric current (Y).

Now let us consider several important special aspects of the producer-product relationship.

2.17. Coproduction: If X_1 and X_2 are both producers of Y (i. e., both are necessary), then X_1 and X_2 are co-producers of Y.

Since no producer is ever sufficient for its product, every producer has at least one coproducer. The set of all coproducers of Y is a deterministic cause of Y. The environment of a producer (X) is always a coproducer of its product (Y) since the environment can always be changed so that X has no probability of producing Y. That is, certain properties of the environment are always necessary; for example, the presence of air and the striking of a bell are coproducers of the bell's ringing. Similarly, water and seeds are coproducers of plants in certain environments.

2.18. Reproduction: If an x_1 which is a member of a structural class (X) produces an x_2 which is a member of the same class (X), x_1 is a reproducer.

Thus oaks are reproducers. Oaks produce acorns and acorns produce oaks. But production is a transitive relationship; that is, if X produces Y and Y produces Z, X is a producer of Z as well as Y, since X is necessary for Z.

Now we can define "function" in terms of the producer-product relationship.

2.19. *Functional Class*: A set of things which are not members of the same structural class but which have either (a) a common producer or type of producer, or (b) a common product or type of product can be said to form a *functional class.*

The property that forms such a class is not a structural property, but a common property of production. If the things involved have a common producer they form a functional class but are not said to have a function. If they have a common product, they not only form a functional class but are also said to have a common function: that of producing the common product or type of product.

For example, my books, my children, and my work bench are structurally different but were all produced by me and, hence, form a functional class, but they are not said to have a common function. On the other hand a sundial, a water-clock, spring watch, and electric clock all produce time-telling and, hence, can be said to have this function.

Our concern here will be almost exclusively with things which can be said to have functions.

Now we can distinguish between three types of function: passive, active, and purposeful.

2.20. *Passive Function*: A set of structurally dissimilar objects have a passive function if the behavior of each is essentially invariant over a wide range of structurally different environments, and these behaviors are potential producers of the same kind of product.

For example, the class of time-telling objects (watches, sundials, and water clocks) have a passive function. So do electric fire-starters: matches, and cigarette lighters. These objects have a function by virtue of their membership in a class which has a specific property of production, not because of their own behavior. For this reason a passive function may also be called an extrinsic function.

Most inanimate objects that we use can have passive functions attributed to them. Such objects are called instruments. For example, there are a number of different kinds of writing instruments--pencil, ink pen, ballpoint pen, chalk, crayon, and so on--each capable of producing marks on a surface.

> 2.21. Active Function: An object has an active function if it can display only one type of behavior in any one environment, but can display structurally different types of behavior in at least some structurally different environments, and these different types of behavior are potential producers of the same kind of product.

Most servo-mechanisms have active functions. A thermostat attached to a heating system in a house has three possible behaviors: do nothing, turn the furnace on, or turn it off. Once it is set it can do only one of these in any environment defined by its temperature. Each of these behaviors produce a temperature in the house within a small range around the setting. The thermostat's (active) function, then, is to maintain the house's interior temperature within a certain range. Automatic pilots on aeroplanes and ships have a similar type of function: maintaining a specified course. An active function may also be called an intrinsic function.

Rosenblueth and Wiener refer to behavior involved in an active function as "goal-seeking," and this it is; but they confused goal seeking and purposeful behavior. Purposeful behavior involves goal seeking, but not all goal-seeking involves purpose.

> 2.22. Purpose: An object has a purpose if it can display structurally different types of behavior in the same

environment, two or more of which are potential producers of the same type of product.

Such an individual can display functional behavior in an unchanging environment and, hence, is more than merely responsive to changes in its environment (as objects with only an active function); it displays choice within an environment.

Note that every object is a producer of its own behavior, since it is necessary but not sufficient for this behavior: if it were sufficient, it would always display this behavior.

2.23. Choice: An object displays choice if it produces structurally different types of behavior in itself in a structurally constant environment, and two or more of these types of behavior have a common function.

I turn now to a detailed consideration of what is involved in purpose and choice.

A PURPOSEFUL STATE

The essential characteristic of purposeful behavior is that it involves choice under constant (structural) conditions. The meaning of this characteristic is revealed by an analysis of the nature of a purposeful state.

A purposeful state has four types of components: the object that displays choice, the behavior it can choose, the product of that behavior (the outcome), and the environment.

2.24. A Subject, A: An object that can display choice; that is, a purposeful individual.

Since my concern here is primarily with human subjects, I will frequently refer to subjects as persons, but choice is not restricted to human or even animate object. Computers programmed to play certain games, for example, can display choice as it is defined here; therefore, they can also have a purpose.

2.25. Courses of Action, C_i $(1 \leq i \leq m)$: structurally different behaviors of a subject produced by that subject in a structurally constant environment, which behaviors have one or more common products and, hence, functions.

2.26. Outcomes, O_j $(1 \leq j \leq n)$: the common products (consequences) of courses of action.

2.27. A Choice Environment, S: the set of properties of the subject's structural environment which coproduce the outcomes of his courses of action.

Note that choice environment is a functional concept: it is a set of structural properties which coproduce common products, outcomes.

The relevant relationships between these components are completely specified by three types of measures which are the parameters of a purposeful state. These are as follows:

2.28. Probability of Choice, P_i: the probability that a subject (A) will produce (i.e., select) a course of action (C_i) in the choice environment (S); that is $P_i = P(C_i \mid A, S)$.

This probability measure applies to a specific individual whose relevant properties may change over time. This fact is of concern to many behavioral scientists. For example, in discussing Shannon's measure of information (which I shall consider in Chapter 8) Wilbur Schramm (1966) commented as follows:

> ... this is one of the pitfalls in the way of applying information theory mathematics to human communication. These are probability formulas, and if the probabilities are altered--i.e., if any learning takes place--during the experiment,

> the events can no longer be regarded as a stochastic process and the formula will not apply. It is therefore necessary rigidly to control the learning factor (p. 522).

The discussion following definition 2.16 is relevant to Schramm's concern. As we shall see later because of learning an individual's probability of choosing a particular course of action may increase or decrease. But this presents no great difficulty. In principle this is no different than the effect of usage on the probability of a lighter's producing a flame. Adjustments for such changes are necessary, but awareness of the kinds of adjustment required can stimulate some very fruitful and fundamental research in the behavioral sciences.

2.29. Efficiency, E_{ij}: that probability that, if a subject selects a course of action of type C_i in the choice environment (S), behavior of type C_i will produce an outcome O_j; that is, $E_{ij} = P(O_j \mid A, C_i, S)$.

2.30. Relative Value, V_j, of an outcome (O_j) to the subject (A) in a choice environment (S). *

To define the relative value of an outcome to a subject requires use of some concepts yet to be developed. For the time being the more familiar concept of utility can be substituted for relative value; the relationship and difference between them will be made explicit in Chapter 3 where both are defined.

I would now like to examine the four components and three parameters of a purposeful state more closely.

Courses of Action

A course of action is not to be construed as mechanistically

*See definition 3.16 and the discussion that follows it for treatment of this concept.

or physically defined behavior, but rather as morphologically defined behavior. Variations in an action with respect to certain of its physical characteristics may not change it. For example, "driving a car" may be designated as a course of action. There are many physically different ways of driving a car but it is frequently useful to group these into one morphological class of behavior. Despite the variations within the class, it can be distinguished from other morphological classes; for example, from "using a street car" or "walking." The morphology of a course of action may be specified narrowly or broadly depending on the purpose of the research. For one purpose (e.g., in testing drivers) it may be desirable to distinguish between automatic and manual shifting of gears. For another purpose (e.g., in planning a program of exercise), it may be desirable to group the use of any self-powered vehicle into one course of action.

It should be noted that the problem of defining a course of action is essentially similar to that of defining a physical object. For one purpose an automobile may be considered as a unit; for another it is a composite of many other units (e.g., wheels, transmission, motor, body, and so on), and for still another purpose it may be considered to be a part of a unit (e.g., a fleet of cars).

2.31. Available Course of Action: a course of action (C_1) in an environment (S), for which the probability of choice (P_1) is greater than zero for some subject.

2.32. Potential Course of Action: a course of action (C_1) is a potential course of action for a subject (A) in an environment (S) if his probability of choosing that course of action is greater than zero; that is, if

$$P(C_1 \mid A, S) > 0.$$

An available course of action may have no probability of being selected by a particular subject, and hence not be a potential choice for him. On the other hand, every potential course of action is available. Further, a course of action that is a potential choice for a subject in one environment may not be in another environment. For example, a person may sometimes use a bicycle in the country, but never in the city. He may be aware of the availability of a bicycle in the city (in a sense to be considered in Chapter 4) and still it may not be a potential choice. For example, many are aware of the availability of narcotics, but nevertheless, never use them.

The relativity of courses of action and outcomes should be noted. Courses of action and outcomes are conceptual constructs of an observer of another's behavior; either may be converted into the other depending on the observer's interests. For example, "chopping the trunk of a tree" may be considered to be a course of action and "the felling of that tree" as its outcome. But "felling a tree" may also be considered to be a course of action which can coproduce the outcome: "clearing a path." Such relativity of concepts appears in all areas of science--for example, the effect of one cause may itself cause another effect--and, hence, does not present any unique methodological problem in this context.

Finally, it will be observed that courses of action are frequently called <u>means</u> and outcomes are frequently called <u>ends</u>.

<u>Efficiency</u>

Many different measures of efficiency of courses of action are in current use. It is fairly common to use some measure of the cost, time, and/or effort required to bring about a specified outcome (e.g., to complete a specified task such as "travelling one mile") as a measure of efficiency. It is also quite common to measure efficiency in terms of the portion of an outcome which is realized by the expenditure of a specified amount of money, time, and/or effort. For example, one can measure the efficiency of a machine-tool either in terms of the number of units produced per dollar or in terms of the cost per unit. Thus, efficiency is commonly measured either as (1) units of input required to obtained a specified output, or (2) the units of output obtained by a specified input. Neither type of measure is sufficiently general to be applied in all situations.

The input required for a fixed output and the output yielded by a fixed input are not constant but vary. For example, the number of units made by a machine per hour varies from hour to hour; the miles per gallon obtained by an automobile also varies. Hence, for a fixed input, various possible outputs exist to each of which a probability can be assigned. If an input is specified in the definition of a course of action, then the efficiency of that course of action for a specified outcome can be defined as the probability that the course of action will produce that outcome. This measure, unlike input- and output-measures, can always be applied to a purposeful state. In order to use probability of production as a measure of efficiency, courses of action which are alike in all respects except the amount of input that they involve must be formulated as different courses of action.

This measure of efficiency of a course of action depends on the environment and the subject involved. Use of skis, for example, may be efficient for self-transportation down a snow-covered hill, but not so down an uncovered hill. Different individuals may ski with different efficiencies and the efficiency of the same individual may change over time (e.g., by learning). Consequently, in order to use this measure it is necessary to specify the relevant time period as well as the individual and relevant properties of the environment.

Relative Value of Outcomes

As in the case of efficiency there is no one measure of the relative value or worth of an outcome that is generally accepted. Fortunately, however, such agreement is not necessary for our purposes here. Nevertheless, it is convenient to use some kind of standard measure wherever possible. A dimensionless measure of relative value provides such a convenient standard. If the values ($\underline{v}_j$) assigned to the various outcomes are all positive, a measure of relative value ($\underline{V}_j$) for each outcome may be obtained by the following conversion:

$$V_j = \frac{v_j}{\Sigma v_j} .$$

Then, since

$$\Sigma \frac{v_j}{\Sigma v_j} = 1.0,$$

it follows that

$$\Sigma V_j = 1.0$$

The minimum relative value (0) occurs only when the absolute value ($\underline{v}_j$) is equal to zero. The maximum relative value (1.0) occurs only when all but one outcome has zero value.

If some or all of the measures (V_j) are negative, one can add to each measure the amount required to raise the lowest value to zero, and convert the resulting adjusted values to relative values. For example,

	Unadjusted Values	Adjusted Values	Relative Values
O_1	-100	0	0
O_2	- 75	25	0.25
O_3	- 25	75	0.75

In the discussion that follows, I shall use the concept of relative value and assume that $\Sigma \underline{V}_j = 1.0$. All the results, however, are easily modified to cover the use of either absolute values or the case in which negative values are employed.

In conceptualizing a purposeful state, it is convenient for the researcher to formulate the available courses of action and possible outcomes as exclusive and exhaustive sets. Sets of courses of action and outcomes which are not exclusive and exhaustive can easily be transformed into sets which are, by use of a Boolean expansion. For example, if we have a non-exclusive and/or non-exhaustive set of outcomes--o_1, o_2, o_3--we can formulate the following exclusive and exhaustive set:

$O_1 = o_1$ and not o_2 or o_3

$O_2 = o_2$ and not o_1 or o_3

$O_3 = o_3$ and not o_1 or o_2

$O_4 = o_1$ and o_2 and not o_3

$O_5 = o_1$ and o_3 and not o_2

$O_6 = o_2$ and o_3 and not o_1

$O_7 = o_1$ and o_2 and o_3

$O_8 =$ not o_1 and not o_2 and not o_3

For an exclusive and exhaustive set of courses of action; the sum of the probabilities of choice must be equal to 1.0 : $\sum_i P_i = 1.0$; and the sum of the efficiencies of each course of action over an exclusive and exhaustive set of outcomes must also equal to 1.0: $\sum_j E_{ij} = 1.0$.

Unless otherwise specified, I will consider the sets of courses of action and outcomes to be defined so as to be exclusive and exhaustive.

A purposeful state can now be defined by use of the concepts that have just been considered.

2.33. Purposeful State: A subject (A) can be said to be in a purposeful state in an environment (S) if the following conditions hold:

(1) There is at least one outcome (say, O_1) which is preferred to another outcome (O_2); hence $V_1 > V_2$

(2) There are at least two potential courses of action for A (say, C_1 and C_2); that is, P_1 and P_2 are greater than zero.

(3) The efficiencies of C_1 and C_2 for O_1 are not equal ($E_{11} \neq E_{21}$) and both have some efficiency for O_1 ($E_{11} > 0$, $E_{21} > 0$).

This definition of a purposeful state may be summarized less technically as follows: a subject may be said to be in such a state if he wants something and has unequally efficient alternative ways of trying to get it.

If we consider a subject over a period of time it will be convenient to refer to the purposeful states at the beginning and end of that period as initial and terminal states, respectively.

The conceptual labors which have been involved in defining a purposeful state are necessary in order to make explicit the meaning of "one mind affecting another," and for identifying the ways in which one mind can affect the other. As we shall see it is necessary to understand the meaning of "one mind affecting another" if one is to understand the nature of human communication. As subsequent discussion will show, these effects may be defined in terms of changes in purposeful states.

CHOICE

The essence of communication, as it will be considered below, is that it involves changes in purposeful states of individuals, and the essence of a purposeful state lies in the availability of choice to an occupant who is capable of exercising it. For this reason it is not possible to pursue the analysis of communication in depth without a deeper understanding of the nature of choice. Therefore, we turn now to a conceptual model of the choice-process and the role of communication within it. The model is sketched here with a very broad brush, finer details are provided in subsequent chapters.

A conceptual model of choice is shown in Figure 2.1. Since the choice process has no beginning or end, we can enter it at any point and ultimately return to the same point. Let us begin with Reality, the subject's purposeful state as it is conceptualized by an observer. Obviously, I do not mean Reality in an ultimate metaphysical sense. The researcher's concept of Reality should not be confused with the subject's whose choice-process is under study; the subject's concept of Reality is contained in his model of it.

Reality, the observer's concept of the subject's purposeful state, consists of the subject (A), and what the observer believes to be the possible courses of action (C_i), the possible outcomes (C_j), and the environment (S) which is made up of a set of variables that are not controlled by the subject but which nevertheless affect the outcome of his action. As noted in the earlier discussions, the state can be characterized by three types of parameter:

(1) $P_i = P(C_i \mid A, S,)$: the probability that A will select C_i in S.
(2) $E_{ij} = P(O_j \mid A, C_i S)$: the efficiency of C_i for O_j in S; the probability that O_j will occur if A selects C_i in S.
(3) V_j: the relative value of O_j to A in S.

The subject's conception or model of Reality may correspond to the observer's but not necessarily be identical to it. The subject's model involves three types of component:

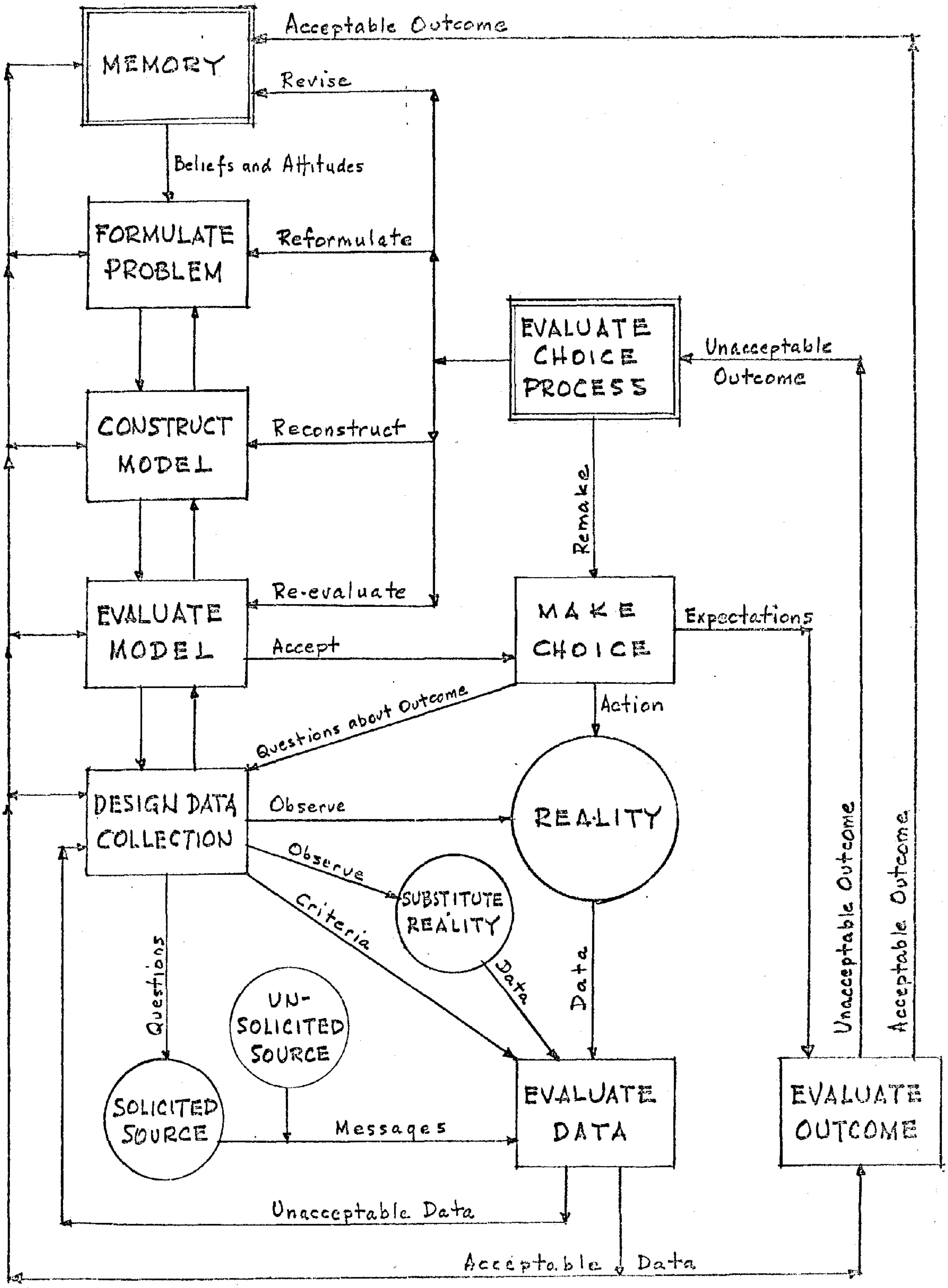


FIGURE 2.1. Conceptual model of choice-process.

(1) The courses of action (c_i) which he believes are available. These may not correspond exactly with the "facts." He may not perceive possibilities that exist, and he may believe certain courses of action are available which, in fact, are not;

(2) The outcomes (o_j) that the subject (A) believes may result from his possible choices. These too may or may not correspond with Reality;

(3) The environment (S) or environments that the subject believes may be the true ones. That is, he may not know what the environment is, but he may consider several possibilities (s_1, s_2, ..) which may or may not include the "true" environment, S.

In addition, the subject's model includes his estimates of three types of parameter:

(1) e_{ij}: estimates of E_{ij}, the efficiencies of the courses of action that he perceives.

(2) v_j : estimates of his relative values, V_j.

(3) $p(s_1)$, $p(s_2)$, ... : estimates of the probability that each "possible" environment pertains.

The subject must also have a concept of the "dynamics" of Reality; that is, how long he can delay before making a choice and what resources are available to him for inquiring into the choice situation.

Confronted by a perceived relevant choice in Reality the subject formulates the problem and constructs (or retrieves from memory) a model of the choice situation using inputs from his past experience (beliefs and attitudes) that are drawn from his Memory and from current observation.

Once a model is formulated, he must decide whether to make a choice now or inquire into the situation; that is, he must evaluate the situation as he sees it. He decides to choose now if he believes that any one, or combination, of the three conditions hold:

(1) his model is adequate and, hence, inquiry is not needed;

(2) choice is necessary now; the consequences of delay are less desirable than the consequences of acting now even if he would prefer to inquire further; and

(3) the situation is not worth inquiring into; that is, the gain that can be obtained by further inquiry does not justify the "cost" of the inquiry.

If he concludes that any of these conditions holds, he proceeds to making a choice and acting. This selection requires use of a criterion of "best" choice, to evaluate the alternative courses of action, and finally selection and action. The action selected may affect Reality and if the subject desires to know how, he "instructs" his Data Acquisition Design function to collect the necessary information and informs his Outcome Evaluator of his expectations. Observations of the effects are made and evaluated. If the data thus obtained are acceptable, an evaluation of the observed outcome is made. If this is acceptable, he simply stores the results in his memory. If the results are not perceived as satisfactory, he must evaluate his choice process and decide whether to modify his beliefs and/or attitudes, his formulation of the problem, his evaluation of it, or re-evaluate the possible courses of action and select another.

If the "feed-back" data are not acceptable, he may either modify his data collection procedure in a way we will consider below, making new observations or re-evaluating the old ones.

Now let us return to the subject's evaluation of his model and examine this process more completely. The subject may find his concept of reality wanting in a number of respects. He may have doubts about the completeness or adequacy of his conception of the possible courses of action, possible outcomes, and possible environments; or he may have doubts about his estimates of any of the parameters involved. Therefore, if (1) he is in a state of doubt, (2) he believes his resources are sufficient for inquiry, and such inquiry has a sufficiently high potential payoff, and (3) he does not feel time pressure, he will decide to investigate further.

He then proceeds to determine what data are required to remove his doubt (i.e., to answer his questions), and how to go about obtaining such data. There are essentially three ways he can go about doing this:

(1) He can inspect (observe) the real situation more closely.
(2) He can conduct experiments on other situations, ones either selected or constructed for the purpose (e.g., a laboratory experiment).
(3) He can address his question to a source which he believes has, or can acquire, the necessary informatinn. The source may be a person or a record of some kind.

Messages from the Solicited Source or from Unsolicited Sources come into a Data Evaluation function which also receives observations of Reality or Substitute. The data are evaluated to determine whether or not they are acceptable. If they are, they are put into his Memory and may be used to either Reformulate the Problem, Reconstruct or Re-Evaluate the Model. If the data are not acceptable, the subject may either redesign the data-acquisition procedure, or he may decide that he must make a choice because of lack of time or resources.

Reality may, and usually, does, contain other persons. Their behavior may be capable of being at least partially controlled by the subject and, hence, his courses of action may be intended to affect their behavior. In such cases, the alternative courses of action available to him may include communicative acts; that is, his course of action may be an act of communication. Responses to his communications are then the products of his actions and may themselves be communicative acts which the subject receives and evaluates. Hence, when the primary intention of communication is to affect another's behavior, he would place the "other" in Reality. If, on the other hand, the "other" is used primarily as a source of information in an inquiry directed toward a choice that does not affect him we would consider him as a Source.

Each rectangle in the graphic model represents a process through which the subject goes, consciously or unconsciously, intentionally or unintentionally. Each process itself may give rise to doubts and the subject may wish to investigate it. For example, he may want to improve his Data Evaluation. This process itself can then become part of the Reality to be investigated. Such a meta-inquiry is methodological in character; that is, an inquiry into the process of inquiry itself. It is possible, of course, to proceed to a still more abstract level and investigate methodological inquiries. These, perhaps, are appropriately called epistemological.

The conceptual model of choice presented here is intended to apply to any type of inquiry, including methodological and epistemological. Hence, the process may involve a nesting of inquiries with the output of one constituting an input into another.

Each phase of the choice model shown in Figure 2.1. is analyzed in detail in subsequent chapters. In these analyses use is made only of those concepts developed in this chapter: a purposeful state and the concepts underlying it; that is, structure, object, behavior, producer-product, and probability of production. The interrelatedness of the concepts to be defined derives from this common conceptual foundation.

BIBLIOGRAPHY

Mill, J. S., A System of Logic. Parker, Son, and Bourn, London, (Fifth Edition) 1862.

Rosenblueth, A., and Wiener, N., "Purposeful and Non-Purposeful Behavior," Philosophy of Science, 17 (1950), 318-326.

__________, __________, and Bigelow, J., "Behavior, Purpose, and Teleology," Philosophy of Science, 11 (1943), 18-24.

Schramm, Wilbur, "Information Theory and Mass Communication," in Communication and Culture, ed. by A. G. Smith. Holt, Rinehart and Winston, New York, 1966, 521-534.

Singer, E. A., Jr., Mind as Behavior. R. G. Adams, Columbus, 1924.

_____, "Mechanism, Vitalism, Naturalism," Philosophy of Science, 13 (1947), 81-89.

Wiener, N., Cybernetics. John Wiley and Sons, New York, (Second Edition) 1961.

CHAPTER 3

PERSONALITY

ME, pro. The objectionable case of I. The personal pronoun in English has three cases, the dominative, the objectionable and the oppressive. Each is all three (Ambrose Bierce, The Devil's Dictionary).

INTRODUCTION

Since the choice process discussed in Chapter 2 can begin at any of its stages and since all stages interact, there is no natural starting point for an examination of this process. I begin at the point where choice is made (i.e., a course of action is selected) because it is here that some of the most general functional properties of the subject are relevant. It will be convenient to have access to definitions of these properties as we probe other aspects of the choice process.

Choice, from the point of view of the subject, consists of (1) deciding which of the courses of action he perceives as available, to select, and (2) carrying it out (implementing the decision). He comes armed for this task with a model of the situation with which he may or may not be satisfied. The exact nature of this model and its development are considered in subsequent chapters. For our purposes here it is sufficient to note that a subject's model (1) identifies the courses of action that he believes are available, and the possible outcomes of these actions; and (2) provides him with estimates of the efficiency of each course of action for each possible outcome.

By adding to this information the relative values that he places on each outcome and a criterion of choice (i.e., his definition of a "best choice") the subject is in a position to make the decision. How good a decision he makes from his own point of view depends on (1) how good his model is, (2) how well he estimates his own relative values, and (3) how effectively he uses his criterion of choice and the model in selecting a course of action.

There appear to be two ways a researcher can approach analysis of the choice process of another person. (1) He can attempt to "see" the situation as the subject does and reconstruct the subject's process of manipulating this view of things so as to reach the decision that the subject does. (2) He can attempt to describe and explain the subject's choice in terms of his (i.e., the researcher's) perception of the choice situation. However, these are not separate ways of studying the subject. In order to determine how the subject views the situation the researcher must use his own view of the situation as a base on which to stand. His own view of the situation, on the other hand, will have predictive and explanatory power only if it has some correspondence with the subject's. This interdependence will become increasingly apparent as we proceed. Different researchers may see the same subject differently, but this is not disastrous providing each researcher formulates his own view in terms of idealized operational concepts and makes these explicit. If they do so it is possible to adjust the different research points of view so as to determine whether or not the results produced by each are consistent. In principle, this situation is no different from that of different observers looking at a pyramid from different positions. One may see a triangle, another may see two triangles, and a third may see a square. But we can easily resolve the apparent differences.

In this chapter I will begin with analysis of the subject's choice process from the researcher's point of view. But as I proceed it will become apparent that we can also use the concepts developed here to define the subject's point of view. For example, in this chapter I will speak of the efficiency of a course of action from the researcher's point of view. In Chapter 5, however, I shall define the subject's belief in efficiency. Similarly, in this chapter I will deal with the courses of action that the researcher believes are available, but in Chapter 5, again, I will try to show how we can determine what courses of action the subject believes are available.

PERSONALITY

Many behavioral scientists and philosophers in the recent past, and even a few in our own time, have thought of personality as a metaphysical or spiritual concept not subject to scientific study. For example, E. E. Eubank (1932) wrote of

> ...the metaphysical nature of the concept, which lies outside the realm of phenomena with which science has been able so far to deal. By its very nature it has up to the present eluded scientific description (p. 105).

But before and after Eubank psychologists and sociologists have made many efforts to so conceptualize personality that it is susceptible to scientific investigation. This has led some, like William Stephenson (1953), to observe

> ...that no one is sure about what to encompass by the term "personality." There are so many meanings for it that it appears almost useless for scientific purposes. Murray [1938], Cattell[1946], Kantor[1933], Kretschmer[1934], Burt [1945], and others, from very different standpoints, regard personality as a rubric for everything that can be found out about a person--his physique, abilities, skills, traits, attitudes, tastes, opinions, knowledge, and all else(p. 273).

It may be possible, however, to find some essential points of agreement in this wide diversity of treatments.

G. W. Allport (1937) surveyed the wide range of meanings that have been assigned to "personality", from ancient Greece forward. He then formulated his own:

> Personality is the dynamic organization within the individual of those psychophysical systems that determine his unique adjustment to his environment (p. 48).

This definition has been widely accepted among psychologists. It has been used recently, for example, by Ross Stagner (1961, p. 8). Allport himself modified his definition slightly in a more recent work (1961) :

> ...personality is the dynamic organization within the individual of the psychophysical systems that determines his characteristic behavior and thought (p. 28).

Apparently he came to identify "unique" and "characteristic." The importance of these concepts is reflected in R. H. Knapp's (1963) observation about Allport's work:

> The result was a view of personality which was centered in the recognition of the peculiar uniqueness of the individual...(p. 154).

This emphasis appears in many definitions of "personality." For example, Stephenson (1953) refers to personality as "the possibility for a distinctive 'character' for a person" (p. 273). Others say the same thing in different ways:

> The distinguishing qualities of an individual taken as a unitary being...(English and English, 1958, p. 382).

> ...entire system of relatively permanent tendencies, both physical and mental, that are distinctive of a given individual and determine his characteristic adjustments to his material and social surroundings (Burt, 1945, p. 107).

> The person is a living whole, individual, unique, striving toward goals, self contained, and yet open to the world around him (Stern, 1938, p. 20).

What is unique about the individual is not the set of stimulae to which he is exposed, but what he contributes through his purposeful responses to his environment. Katz and Schanck (1938) emphasized this point:

> Personality is the concept under which we subsume the individual's characteristic ideational, emotional, and motor reactions and the characteristic organization of these responses. Characteristic in this definition means that the conduct in question is more a function of the individual than of the immediate stimulating situation. Thus we would exclude from personality . . behavior that which is imposed by the exigencies of the present situation. (p. 391).

In addition to the emphasis on uniqueness, a second recurring characteristic of definitions of personality is their emphasis on "totality." For example, Watson (1924) wrote:

> I define personality as the sum of activities that can be discovered by actual observation of behavior over a long enough time to give reliable information (p. 220).

He did not make clear how one sums activities, nor did M. Prince (1924) who similarly wrote:

> Personality is the sum-total of all biological innate dispositions, impulses, tendencies, appetites, and instincts of the individual, and the acquired dispositions and tendencies (p. 532).

Of this and similar definitions Katz and Schanck (1938) observed:

> In other words, personality is the complete term to sum up all the individual's potential responses. The difficulty with this definition is its very inclusiveness. It is very much like defining the world as the sum total of everything in it (p. 390).

Nevertheless, the emphasis on the generality and all-inclusiveness of personality persists.

The third aspect of personality that recurs in many definitions is the relevance of the way an individual responds or adjusts to his environment; for example, see the definitions of Allport and Burt quoted above. Early social psychologists tended to concentrate on responses to the social environment. For example, F. H. Allport (1924) wrote:

> Personality may be defined as the individual's characteristic reactions to social stimuli and the quality of his adaptation to the social features of the environment (p. 101).

As many psychologists have pointed out, however, Robinson Crusoe, even before he acquired Friday, had and displayed a personality. It seems clear that this emphasis on social stimulae reflects the interests of the social psychologist rather than the irrelevance of the non-social aspects of the environment.

Even if no two personalities are alike they are not likely to be different in every respect. Hence there have been many efforts to reduce the diversity by finding basic personality types, drives, and forces. These efforts have been directed at personality in general, rather than at personalities in particular. Brand (1954) noted this dichotomy and commented on an early effort that C. W. Churchman and I made to synthesize these approaches as follows:

> Our goal has been to consider what is the study of personality. We have found two main proposals: the identification of personality as general and the identification of personality as individual behavior. At the present time the former proposal has greater support than the latter one. In contrast to the individual-behavior and the general-behavior definitions, here is the functional definition. [The study of personality is the determinition of the characteristic ways (as measured by a probability function) an individual has of selecting alternative means for a given end.]* The merit of it, at least

*The sentence in brackets is taken from the preceding paragraph in Brand (p. 16).

> as presented by Churchman and Ackoff (1947), is that it offers a proposal for the precise identification of personality within a general-behavior theory. A method is also suggested by which personality may be measured quantitatively. The disadvantage of the proposal is that it requires a methodology not familiar to current research practice in psychology, and it still has to develop an experimental program (p. 16).

The disadvantages to which Brand refers remain but, hopefully, this book will reduce them. The definition of personality to which Brand refers is not the same as that which is developed below; it is the same kind of definition but it has gone through a number of (again hopefully) progressive revisions since 1947.

From this brief analysis of definitions of personality I conclude that a new definition should (1) capture the uniqueness of the individual, (2) provide a very general concept under which all other psychological concepts can be subsumed, and (3) locate personality in the responses of an individual to his environment. The definition that is developed here does, I believe, satisfy these conditions, and unlike the definitions we have examined it provides both a measure of personality and a basis for explicitly relating all other psychological concepts to it. Now let us turn to the task.

From the researcher's point of view, of what can an individual's uniqueness consist? To answer this question we must return to the researcher's model of a choice situation. It identifies the subject (A); the available courses of action; $\{C_i\}$; the possible outcomes, $\{O_j\}$; the environment, $\{S_k\}$*; the subject's probabilities of choice, $\{P_i\}$; the efficiencies of each available course of action for each possible outcome, $\{E_{ij}\}$; and the relative values that the subject places on these outcomes, $\{V_j\}$. The courses of action, outcomes, and environment are characteristics of the situation which are independent

*Recall that the choice environment consists of the set of properties of the subject's physical environment which, with his course of action, coproduce the outcome.

of the subject. The probabilities of choice, efficiencies, and relative values depend on the subject but are not independent of the situation. Therefore, the personality of the subject, his uniqueness, must derive from the way his probabilities of choice, efficiencies, and relative values depend on the properties of the situation.

Let me put the same thing in another way. The "contribution" of a purposeful individual to a choice situation must manifest itself by an affect on what happens in that situation, the outcome. Let $P(O_j)$ represent the probability that an outcome, O_j, will occur in the choice situation. If the probabilities of different outcomes in a choice situation were independent of the subject then there would be no functional difference between different subjects in that situation. But it is just such a difference (i.e., in outcome) that a difference in personality must produce if it esists. Let us pursue this line a bit further.

If $P(O_j)$ represents the probability that an outcome O_j will occur in a particular choice situation. Then

$$P(O_j) = \sum_i P_i \, E_{ij} \qquad (3.1)$$

that is, the probability that O_j will occur is the sum of the products of the probability that each course of action will be selected and the probability that, if selected, it will produce the outcome O_j. For example, in the simple case where $P_1 = 0.6$, $P_2 = 0.4$, $E_{11} = 0.7$, $E_{12} = 0.3$, $E_{21} = 0.1$ and $E_{22} = 0.9$, then

$$P(O_1) = P_1 \, E_{11} + P_2 \, E_{21} = 0.6\ (0.7) + 0.4\ (0.1) = 0.46$$

$$P(O_2) = P_1 \, E_{12} + P_2 \, E_{22} = 0.6\ (0.3) + 0.4\ (0.9) = 0.54$$

Now the subject's probabilities of choice and the efficiencies of these choices depend on the properties of the situation: the available courses of action, the possible outcomes, and the environment. They also depend on the relative values the subject places on these outcomes, but these relative values in turn depend on the properties of the situation.

Hence what the individual "contributes" to a choice situation is a transformation of situational properties into probabilities of choice, efficiencies, and relative values. His personality, then, must lie in this transformation. That is, if two persons are placed in the same choice situation, the difference in their personality must be manifested in the difference in values of their probabilities of choice, efficiencies, and relative values.

An individual's probabilities of choice, efficiencies, and relative values can each be expressed, in principle, as a function of the choice-situation characteristics, that is

$$P_i = f\left[\{C_i\}, \{O_j\}, \{S_k\}\right] \tag{3.2}$$

$$E_{ij} = g\left[\{C_i\}, \{O_j\}, \{S_k\}\right] \tag{3.3}$$

$$V_j = h\left[\{C_i\}, \{O_j\}, \{S_k\}\right] \tag{3.4}$$

In a sense, then, these three functions--f, g, and h--are the three "dimensions" of an individual's personality. It would be desirable, however, if they could be combined into a single function. This can be done as follows.

The subject's expected relative value of a choice situation (EV), as determined by a researcher, is

$$EV = \sum_i \sum_j P_i E_{ij} V_j \tag{3.5}$$

But since the P_i's, E_{ij}'s and V_j's are functions of the choice situation, then so too is the expected relative value. Then we can write

$$EV = \pi\left[\{C_i\}, \{O_j\}, \{S_k\}\right] \tag{3.6}$$

Then we can define personality as follows:

3.1. <u>The Personality of a Purposeful Individual</u> is a mathematical function, π, which relates his expected relative value in any choice situation to the properties

> of the available courses of action, of their possible outcomes, and relevant environmental variables.

Hence, personality is not conceived here as an unobservable intervening variable involved to explain choice, but as an unobservable function which describes how an individual converts a choice situation into an expected value for himself.

This definition of personality is not as operational as it is programmatic; that is, it does not tell us how to find the function Π but it can be used to design a research program that will ultimately yield Π. For example, to evaluate Π we must develop appropriate and general quantitative ways of representing the available courses of action and possible outcomes, and a specification of a set of variables which are sufficient to characterize any choice environment. Such development requires considerable research. However, I will indicate what kinds of research will enable us to "move up on" the personality function. Each of the types of research (and the concepts associated with them) involves an aspect of personality, a "slice" through the multidimensional personality space. These studies and concepts can be grouped into three major categories depending on whether they treat probability of choice, efficiency of choice, or relative value as the dependent variable. Studies involving probability of choice as the dependent variable I shall call familiarity studies; those involving efficiencies of choice, knowledge studies; and those involving relative values, intention studies.

Perhaps the relationship between these three aspects of personality is better understood in the following terms:

1. The measure of familiarity derives from a measure of probability of choice where the choice has no effect on what occurs and, hence on its probability of occurrence.

2. The measure of knowledge derives from a measure of probability of choice where choice affects the probability that a particular outcome will or will not occur.

3. The measure of intention derives from a measure of probability of choice where the choice affects what outcome will occur, but not its probability of occurrence.

PROBABILITY OF CHOICE: FAMILIARITY

The objective here is to determine how an individual's probabilities of choice are influenced by properties of available courses of action and the choice environment, properties which do not affect the efficiencies of the alternatives. Hence, I want to construct a choice-situation in which possible influence of efficiency on the subject has been removed. This can be done as follows:

3.2. Familiarity (Choice) Situation: one in which (1) the possible outcomes are grouped into two exclusive and exhaustive classes, O_1 and O_2, where the subject prefers O_1 to O_2 (i.e., $V_1 > V_2$); and (2) each of the available set of (exclusively and exhaustively defined) course of action has an equal efficiency for each possible outcome; that is,

$$E_{11} = E_{21} = \ldots = E_{m1} = L(E_{i1}), \text{ and}$$

$$E_{12} = E_{22} = \ldots = E_{m2} = L(E_{i2});$$

where $L(E_{ij})$ is the "level of efficiency" of all courses of action for outcome O_j.

Note that

$$O \leq L(E_{ij}) \leq 1$$

and

$$\sum_j L(E_{ij}) = 1.0$$

In such an environment there is not a choice of outcome, only of a course of action.

For example, consider a choice situation in which the subject wants to write a letter (O_1), hence, O_2 is "not writing a letter." Suppose we have a set of ball-point pens identical in all respects except their color. Then the use of each is equally efficient for writing a letter. Choice in such a situation reflects the subject's preference for "color of ink."

3.3. Degree of Familiarity. A subject's degree of familiarity (DF_{ij}) with a course of action (C_i) relative to a preferred outcome (O_j) whose relative value (V_j) is equal to 1.0, and an available set of exclusively and exhaustively defined courses of action $\{C_i\}$ of which C_i is a member, is the probability that he will select C_i in a familiarity situation in which the level of effeciency for the preferred outcome, $L(E_{ij})$, is 1.0; that is, $DF_{ij} = [P_i \mid \{C_i\}, V_j = 1.0, L(E_{ij}) = 1.0]$.

3.4. Familiarity Function. A subject's familiarity function for a course of action (C_i) relative to an outcome (O_j) and an available set of courses of action $\{C_i\}$ is that mathematical function (f_F) which satisfies the equation: $P_i = f_F [\{C_i\}, V_j, L(E_{ij})]$.

The familiarity function describes preferences for courses of action independently of outcomes. In effect, by equating the outcomes of each of the available courses of action, we convert these actions into outcomes. Hence, a subject's preferences among them may reflect their intrinsic value to him, not their instrumental or extrinsic value (i. e., as a means to an end). For example, in the ball-point-pen situation described above if a subject were to select the one with blue ink most frequently, then he may have a preference for this color in this context. Using blue ink may have a "value in itself" for him. On the other hand, he may select the blue ink because he (erroneously) believes it to be more efficient. (I shall consider such beliefs in Chapter 5.) Therefore, if we require that the subject believes the alternative courses of action to be equally efficient, the degrees of familiarity obtained are measures of the intrinsic relative values of the courses of action to him.

If we obtain the degrees of familiarity or familiarity functions for each of a set of courses of action which differ with respect to only one variable (e. g., color or size of instrument employed), relative to a particular preferred outcome, then we can explain the differences between these measures or functions by the variable that produced these differences. The function which relates the differences between familiarity measures or functions to such a variable is an aspect of the subject's personality because it explains how his probabilities of choice are produced by a property of the choice situation.

Variables whose effect on familiarity can be studied in this way can be classified as either structural or functional. Studies of familiarity involving structurally defined variables (e. g., color, shape, size, and texture) relate to what is called an individual's <u>taste</u> or <u>style.</u> Studies involving functionally defined variables relate to what psychologists have called <u>personality traits</u> (e. g., selfishness, generosity, bravery, cowardice, aggressiveness, introversion, extraversion, cooperativeness, and so on). Only traits are considered in detail here.

Traits

Since the early 1930's many psychologists have been concerned with defining, cataloguing, and testing personality traits. However, they have yet to provide an operational definition or a metrical standard appropriate to the concept. For example, some typical definitions are:

> Any enduring or persisting characteristic of a person by which he can be distinguished from another (English and English, 1958, p. 561).
>
> ...dispositions which have differing strengths in different people which persist over a period of time (Fiske, in Heine and Wepman, 1963, p. 454).
>
> We shall use the term "trait" to refer to a consistent feature of personality which has some emotional or ideational content (Stagner, 1961, p. 156).
>
> A trait is the learned tendency of an individual to react as he has reacted more or less successfully in the past in similar situations when similarly motivated (McClelland, 1956, p. 357).

McClelland amplified his definition as follows:

> The trait variable probably ought to be reserved primarily to describe the consistencies in behavior or the modes of adjustment which the subject habitually adopts to meet recurrent situations... In some respects traits are the most obvious aspects of personality; they are nearest the surface, most easily identified, and most often used to describe another person (p. 352).

However, this is only one of the senses in which "trait" is used. English and English (1958) pointed out that

> Usage, even by the same author, fluctuates between reference to a consistently manifested pattern of behavior (= surface trait) and to a part of the enduring structure of the person (inferred from behavior) which is the cause of the consistency (source traits) (p. 561).

Stephenson (1953) also distinguished between these two uses of "traits." Traits, he wrote

> (1) may name a whole behavioral segment as it is observed from the exterior, or (2) may be the name for an inner attribute, imputed to someone as a causal agency...(p. 274).

Stagner (1961) made the following observations about these types of traits:

> How many surface traits ther are, and how many source traits, cannot be definitely stated. Allport and Odbert (1936) counted 17,953 trait names in English, but many of these were synonyms and others represented temporary rather than permanent trends. R. B. Cattell (1945), making an exhaustive study of ratings, found a total of 131 "phenomenal clusters," or common traits. These grouped themselves readily into 50 "nuclear clusters" of related traits, which in turn could be arranged in 20 "sectors of the personality sphere" (pp. 163-64).

Some examples of these surface-trait "sectors," or source traits, are

Egotism, assertion, stubborness	vs.	Modesty, self-effacement, adaptability
Sociability	vs.	Timidity, hostility, gloominess
Amorousness, playfulness	vs.	Propriety

Current interest in traits is largely due to the early work of G. W. Allport (1928) who wrote "A trait of personality is a characteristic form of behavior more generalized than a single reaction or simple habit" (p. 119). Traits are necessary, Allport claimed, for clarifying "the repeated occurrence of actions having the <u>same significance</u> (equivalence of response), following upon a definable range of stimuli having the same personal significance (equivalence of stimuli)" (p. 340). Allport provided no clarification of "equivalence." I have suggested that equivalence means "members of the same functional class," a suggestion which can be read into a later definition of trait offered by Allport (1937):

> A trait is a generalized and focalized neuropsychic system (peculiar to the individual), with the capacity to render many stimuli functionally equivalent, and to initiate and guide consistent (equivalent) forms of adaptive and expressive behavior (p. 295).

It will become apparent that my treatment is fundamentally an "operationalization" of Allport's, one that gains precision because it is imbedded in a system of concepts and because an appropriate measure is developed.

In the study of traits as conceived here the subject's possible responses (courses of action) are functionally classified and traits are associated with the subject's characteristic response (relative to this classification) to a functionally defined stimulus.

3.5. Trait-- relative to a functionally defined stimulus in a familiarity situation and a set of functionally defined alternative courses of action, a greater degree of familiarity with one of the courses of action than any of the others is a trait. The degree of the trait is the degree of familiarity associated with the course of action most often selected.

Since the measure of a trait is a degree of familiarity it is apparent we can also formulate a trait function.

It would undoubtedly be helpful to show how this general definition applies in the case of a specific trait. Furthermore, by developing a definition and measure of a specific trait I can also show how the definition and measure can be used to design a trait test. I use the ascendance-submission trait for this purpose because it is one of the most commonly discussed traits in the psychological literature. The type of trait test I will begin to construct yields information about the trait function as well as estimates of the degree of the trait under certain specified conditions.*

In the ascendant-submissive situation the subject is faced with the following set of conditions: an aggressive act has been committed which decreases the efficiency of the subject's behavior with respect to his objective (i.e., his preferred outcome). In other words, the stimulus is an aggressive act. The response in which we are interested is the subject's attempt to retaliate on the aggressor, that is, to control rather than be controlled by the aggressor. Thus we are interested in whether or not the subject responds to aggression and how he responds.

We can define ascendant-submissive behavior in terms of the following aspects of the subject's behavior:

(1) A response by the subject A to another purposeful individual's (B's) act, when B's act decreases the efficiency of A's behavior with respect to (A's) objective; that is, when B aggresses on A.

(2) A potential producer of a reduction in the efficiency of B's behavior relative to his (B's) objective.

In terms of these aspects of behavior the following exclusive and exhaustive set of courses of action can be defined:

C_1 : To exhibit both (1) and (2), an ascendant act.

C_2 : To exhibit (1) and (not-2), a submissive act.

C_3 : not to exhibit (1) but to exhibit (2), an aggressive (but not ascendant) act.

C_4 : to exhibit neither (1) nor (2), neither an aggressive, ascendant, nor submissive act.

The "degree of ascendance" of an individual can be defined as the probability of his choosing behavior pattern C_1 , and the degree of

---

*The discussion of the trait test which follows is a slightly modified version of one that appeared in The Design of Social Research by this author (University of Chicago Press, Chicago, 1953).

his submission" can be defined as the probability of his choosing behavior pattern C_2. The sum of these probabilities we can call his degree of awareness or consciousness* of the aggression.

From this observation we can immediately discern one loss involved in the use of verbal tests rather than overt behavior in the study of traits. In the verbal test we must ask the subject how often he tends to respond in a certain way when aggression occurs. Such a question will at best elicit information concerning his response to aggression when he is conscious of the aggression and his response to it. But many people respond to aggression without being fully conscious either of the aggressive act or of their response. Hence, a verbal trait test will at best gather evidence on the subject's sensitivity to aggression when he is fully conscious of the aggression (in the sense that he can recall the aggression and his response to it). A more general measure of ascendance would depend on evidence other than the subject's verbal testimony or on a method of inferring from conscious responses to nonconscious ones.

In Appendix I** a verbal test of ascendance-submission is developed. This development demonstrates how definitions of the type formulated here provide instructions for measuring the concept involved.

EFFICIENCY OF CHOICE: KNOWLEDGE

A major aspect of personality to which psychologists have given considerable attention is that of an individual's capabilities or aptitudes. In ordinary language we use three terms in this connection: "knowledge," "understanding," and "intelligence." The first two of these have received more attention from philosophers than from psychologists, but "intelligence" has been a major preoccupation of psychologists. The meanings of these concepts and the difference between them is far from clear in either ordinary or technical usage.

*These concepts are treated in Chapter 4.

**To appear in Volume 2.

"Knowledge," for example, is used in at least two different senses: (1) awareness or possession of a fact or state of affairs (e. g., as in knowing that someone is at home or that water is made up of hydrogen and oxygen), and (2) possession of a practical skill. In the first sense knowledge consists of an individual's true beliefs or what he is aware of; that is, an individual's true beliefs or whatever he is aware of, he knows. I shall pursue this sense of knowledge in Chapters 4 and 5 where the nature of awareness and belief is explored in detail. Here I concentrate on knowledge as a practical skill, on knowing how to do something rather than on knowing about something. Abilities are relevant to knowing how, not to knowing about.

I shall consider knowledge of courses of action, and in this context knowledge is related to the efficiency with which an individual can use a course of action to obtain an objective. In this sense knowledge is clearly a capability.

"Understanding" implies something deeper than knowledge. For English and English (1958, p. 510) and G. W. Allport (1937, pp. 536-37) it involves apprehending the meaning or significance of what is known. According to Dewey (1938):

> ...that which is observed, no matter how carefully and no matter how accurate the record, is capable of being understood only in terms of projected consequences or activities (p. 499).

This reflects Kohler's earlier observation (1929) that understanding implies perception of causal connections between that which is understood and other things; or, as F. H. Allport has put it (1954), "understanding is what one gets as a result of adequate explanation" (p. 11).

Following this lead I shall treat understanding as the ability to adjust one's behavior efficiently to changes in the conditions which affect its efficiency. This implies the ability to explain the effect of changes in one's environment on the efficiency of one's choices.

"Know-how" can be used in a general sense: to designate an individual's ability to obtain what he wants in a given situation by use of any means that are available to him.

I should like to delay discussion of "intelligence" until the concepts just considered are provided with adequate definitions.

Knowledge

As indicated in Chapter 2, courses of action are usually defined morphologically or functionally. Any functionally defined course of action can be broken down into a set of exclusive and exhaustive morphologically defined courses of action; and any morphologically defined course of action can be decomposed either into a similar set of physically defined courses of action or into a set of more finely defined morphological courses of action. For example, if the course of action (C_1) is "to use public transportation" and the relevant outcome is "to go from a to b in a specified time," the course of action may be decomposed as follows. Suppose there are only five possible ways ($w_1, w_2, \ldots, w_5$) of going from a to b by public transportation. Then we can define an exclusive and exhaustive subset of actions:

c_{11} = to select w_1
c_{12} = to select w_2
c_{13} = to select w_3
c_{14} = to select w_4
c_{15} = to select w_5
c_{16} = to select any other way.

Note that c_{16} has no efficiency for going from a to b in the specified time. It is included to make the set exhaustive. Note also that the efficiencies of these subcourses of action for going from a to b in the specified time is independent of the subject. As far as the subject is concerned, the efficiencies are "determined." They may, of course, depend on the environment; for example, the efficiencies of the ways of going from a to b may depend on the weather.

3.6. Knowledge Situation: a choice situation in an environment S in which a set of subcourses of action $\{c_{ik}\}$ is available, a set whose members are exclusive and which exhaust a morphologically or functionally defined course of action, C_i. The efficiency of each subcourse of action for a specified outcome (O_j) is independent of the subject who makes the choice.

Let e_{ikj} represent the efficiency of a subcourse of action (c_{ik}) for an outcome (O_j) in a knowledge situation, and let p_{ik} represent a subject's probabilities of selecting that subcourse of action. Then, using these concepts we can reformulate the definition (2.24) of the efficiency of a course of action, C_i. The efficiency of C_i for O_j for a subject (A) in a knowledge situation is given by

$$E_{ij} = \sum_k p_{ik} \; e_{ikj}. \tag{3.2}$$

Now we can say what "knowing a course of action" means.

3.7. Degree of Knowledge (DK_{ij}) that a subject (A) has of a course of action (C_i) relative to a preferred outcome (O_j) with relative value (V_j) equal to 1.0 in a choice environment (S) is

$$DK_{ij} | S, V_j = 1.0 = \left(\frac{E_{ij} - \min e_{ikj}}{\max e_{ikj} - \min e_{ikj}} \middle| S, \; V_j = 1.0 \right)$$

where 'min e_{ikj}' represents the least efficiency associated with any of C_i's subcourses of action, and 'max e_{ikj}' represents the greatest such efficiency. These are the minimum and maximum possible efficiencies of C_i for O_j in S.

Note that when max $e_{ikj} = 1.0$ and min $e_{ikj} = 0$, $DK_{ij} = E_{ij}$.

It can be seen that the degree of knowledge of a course of action relative to an objective in a specified environment is a measure of the amount of control a subject has over the outcome relative to the maximum amount of control that is possible.

Suppose, in the example involving driving from a to b we have the following information for a subject (A):

	p_{ik}	e_{ikj}
c_{i1}	0.1	0.9
c_{i2}	0.3	0.8
c_{i3}	0.3	0.7
c_{i4}	0.2	0.6
c_{i5}	0.1	0.5
c_{i6}	0.0	0.0

Then A's efficiency would be

$$E_{ij} = 0.1\,(0.9) + 0.3\,(0.8) + 0.3\,(0.7) + 0.2\,(0.6) + 0.1\,(0.5)$$
$$= 0.71.$$

A's degree of knowledge of C_i for O_j would be

$$\frac{0.71 - 0.50}{0.90 - 0.50} = \frac{0.21}{0.40} = 0.51.$$

If a subject were always to select that subcourse of action with maximum efficiency for outcome O_j, then his degree of knowledge of the relevant course of action (C_i) would be maximum and equal to 1.0. If he were always to select the least efficient subcourse of action, then his degree of knowledge would be minimum and equal to zero.

The degree of knowledge is a measure which is made relative to a particular set of environmental conditions (S) and a specific relative value of an outcome (V_j). Therefore, we can generalize as follows:

3.8. The Knowledge Function of a subject (A) for a course of action (C_i) relative to an outcome (O_j) in an environment S is a mathematical function (f_K) which satisfies the equation:

$$(DK_{ij}|S) = f_K(V_j|S)$$

A subject's degree of knowledge of a course of action may be independent of the relative value of the relevant outcome to him, but in general we would expect it to increase as V_j increases and to be maximum when $V_j = 1.0$. It could, however, decrease as V_j increases. The sensitivity of a subject's degree of knowledge of a course of action (C_i) for an outcome (O_j) to V_j can be measured by the derivative of the former with respect to the latter:

$$\frac{d(DK_{ij}|S,\ V_j)}{d\ (V_j|S)}$$

If this derivative has a value of zero for all values of V_j, the subject's knowledge of C_i is insensitive to V_j. If it is positive, he is sensitive to V_j; if negative, he is also sensitive but in a curious way: his knowledge decreases (increases) as the relative importance of the relevant outcome increases (decreases).

The knowledge function can be generalized further:

3.9. The Generalized Knowledge Function of a subject (A) for a course of action (C_i) relative to an outcome (O_j) is a mathematical function (f^*_K) which satisfies the equation:

$$DK_{ij} = f^*_K (S, V_j)$$

Recall that the choice environment (S) consists of a set of properties of the subject's physical environment that affect the outcome of his choice. Hence S may consist of more than one variable ($s_1, s_2, \ldots$). Therefore, the generalized knowledge function describes how the subject's efficiency depends on these variables, and hence is an aspect of his personality function. For example, the efficiency of a subject's choice in going from one place to another may depend on the weather. How it does is an aspect of his knowledge and personality functions.

The concept of knowledge can be applied to instruments as well as to courses of action. To show how, it is first necessary to define 'instrument.'

3.10. Instrument: an object which coproduces the outcome of a subject's behavior, which coproduction is itself produced by the subject.

Oxygen in the air and a match may coproduce a fire but oxygen is not an instrument as is the match. A purposeful individual must strike the match, hence produce its coproduction. The amount of oxygen in the air is not usually controlled by the subject but the behavior of the match is.

Now if 'using a match' is taken as a course of action we can decompose it into subcourses of action, all involving use of a match. We can then define a subject's degree of knowledge of "use of a match" for the outcome, say, "starting a fire." This would then be his degree of knowledge of the instrument relative to the outcome, "starting a fire." By extension we can define his knowledge function and generalized knowledge function of "use of a match."

Knowledge, as I have treated it, is an awareness of the efficiency of alternative subcourses of action under constant environmental conditions. Now consider the effect of environmental conditions on the efficiency of a subject's choices.

Understanding

Understanding is responsiveness to whatever affects efficiency. To make this more precise imagine a set of environments $\{S\}$ which differ from each other with respect to one environmental variable (s) which affects the efficiency of a course of action (C_i) relative to an outcome (O_j). Let $s_1, s_2, \ldots, s_n$ be an exclusive and exhaustive set of values of s over some relevant range of s-values. Let S' represent the set of environmental variables common to all the members of the set $\{S\}$. Now we can define a set of subcourses of action which differ only with respect to values of s:

$$
\begin{aligned}
C_1' &: C_i \text{ under } s_1 \\
C_2' &: C_i \text{ under } s_2 \\
&\vdots \\
C_n' &: C_i \text{ under } s_n
\end{aligned}
$$

If these courses of action are made available to an individual (A) in a choice environment (S') his choice constitutes a selection of an environment. Let E_{ij}' represent the subject's efficiency with C_i' in S' relative to outcome O_j. Then his overall efficiency for O_j, ϵ_{ij}, is given by

$$\epsilon_{ij} = \Sigma P_i' \, E_{ij}' \qquad (3.3)$$

where P_i' is his probability of choosing C_i' in S'.

3.11. <u>Degree of Understanding</u> (DU_{ij}) that a subject (A) has of a course of action (C_i) relative to an outcome (O_j) with relative value (V_j) equal to 1.0, with respect to an environmental variable (s) in a choice environment (S') is

$$(DU_{ij} \mid s, S', V_j=1.0) = \left(\frac{\epsilon_{ij} - \min E'_{ij}}{\max E'_{ij} - \min E'_{ij}} \middle| s, S', V_j=1.0 \right)$$

where min E_{ij}' represents the efficiency of that C_i' in S' which is minimum and max E_{ij}' represents that which is maximum.

Min E_{ij}' and max E_{ij}' represent the worst and the best that the subject could have done.

The degree of understanding has a maximum value of zero and a maximum value of 1.0.

The efficiency of the use of slides to convey information, for example, depends on the level of illumination in the room in which they are projected. If we subclassify "using slides" by appending various levels of illumination, a test can be designed to determine how well a subject understands the effect of illumination on conveying information by use of slides.

This measure, as that of knowledge and know-how, can be generalized into an understanding function (f_U) where

$$(DU_{ij} \mid s, S') = f_U \, (V_j \mid s, S');$$

and a generalized understanding function (f_U^*) where

$$DU_{ij} = f_U^* \, (s, S', V_j).$$

Intelligence

One might well point out that a person's knowledge or understanding constantly changes. At times it changes rapidly and at other times slowly. Many psychologists have felt it important to characterize the individual's ability to extend his knowledge and understanding; that is, to learn. The ability to learn has traditionally been called intelligence. This ability has been described in many ways. Katz and Schanck (1938) said that intelligence "is generally defined as the ability of the individual to adjust to new problems and conditions of life" (p. 418). They go on to point out that "adjustment" is difficult to define adequately and therefore re-define intelligence as "the learning and thinking abilities of the individual" (p. 419).

Thorpe and Holliday (1928) combined both concepts in their definition: "By intelligence we mean principally the capacity for learning, for applying what has been learned, and for making appropriate adjustments to life's problems" (p, 5).

Allport (1937) refers to it as a "capacity to solve novel problems," (p. 406) and as innate individual equipment (p. 108).

These definitions suggest that intelligence has to do with learning in problem situations that are in some sense "novel." Now it is obvious that one individual may learn more rapidly than another, perhaps because of a better formal education, richer experience, encouragement of companions, and so on. The rate at which an individual learns may be influenced by any number of such things. But Allport and many others have suggested that intelligence is "innate" and is therefore independent of such influences. If so, then intelligence is not merely the measure of the rate at which an individual learns, for his potentiality may be enhanced or diminished in a specific environment. The task of measuring intelligence, then, seems to be one of determining the individual's rate of learning independent of any "outside" influences.

Although most psychologists assert that intelligence is, or is dependent on, the ability to learn, it is nevertheless true that the ability to learn is <u>not</u> measured in standard intelligence tests. Dearborn (1928) observed:

> Defining intelligence, as many have done, as the ability or capacity for learning, we then noted that, for practical reasons, the tests in common use are not tests of the actual process of learning but are tests of what has been learned. The assumption is made that if one samples the results of learning in matters where all the individuals tested have had an equal chance at learning, he may arrive at an estimate of the capacity to learn. But since it is difficult to find even simple experiences which are common to all individuals of a given age period, actually, again one tries by sampling a large range of fairly common experiences to strike an 'average' which, despite the fact that a given individual may have missed this or that experience, will still be representative of the individual's learning (p. 99).

Woodrow (1946) has also pointed out that most psychologists have a "tendency to confuse achievement with gain" (p. 156), and that intelligence tests measure achievement (past learning) rather than gain. The measurement of gain is involved in the measurement of the ability to learn. "The ability to learn," he concludes, "cannot be identified with the ability known as intelligence." (p. 148). Since "Learning, as we measure it" according to McGeoch (1942), "is a <u>change</u> in performance with practice" (p. 3). Woodrow comes to this odd conclusion: "intelligence" is what is measured by intelligence tests rather than what it is defined to be, and shows that it is not even "significantly correlated" with tests of learning ability. He demonstrates thereby that the inference of capacity for learning from intelligence tests as they are given is not justified. However, he fails to show why the form of the present tests should constitute the basis for defining intelligence.

In other words, common intelligence tests, whether those which seek a general measure such as the Binet tests, or tests of specific mental abilities that make up intelligence, attempt to measure what has been learned and infer therefrom the ability to learn. Such inference is based on a chain of unjustified assumptions. What an individual <u>has</u> learned is <u>now</u> his knowledge or understanding, thus the tests, if they measure anything, measure knowledge or understanding.

Getzels and Jackson (1962) made the same observation in another way:

> In short, the conventional I. Q. test tends toward the evaluation of those processes that have been called convergent, retentive, and conservative more than those that have been called divergent, innovative, and constructive (p.).

Stagner (1961) has put it more simply:

> The I. Q. far from being a measure of innate capacity is, as early as age six, a composite of capacity and achievement. By the time the child has reached age ten, the achievement component is probably somewhat larger (p. 473).

The assumption on which inferences from achievement to learning-ability are based is as follows: "if one samples the results of learning in matters where all the individuals tested have had an equal chance at learning, he may arrive at an estimate of the capacity to learn." This assumption is built on even less secure foundations than Dearborn indicates. What of specifying the meaning of "an equal chance at learning"? Chance for learning is usually taken to be exposure to a formal educational system. But what of the non-formal education of the home, church, and so on. In what sense could the chances for learning in homes of the same economic class, let alone different ones, be said to be equal? In what sense does the child with an oppressive home life, whatever form such oppression takes (financial, economic, etc.), have an equal chance in school with those more fortunate? If, as Dearborn indicates, tests are given based on this assumption "for practical reasons, such as the great length of time required for the observation of significant learning", (p. 68) consider how much time would be required to evaluate quantitatively the chances for learning, the equality of which for different individuals is so blithely assumed on a common sense level.

The above assumption is, supposedly, bolstered by a second, "by sampling a large range of fairly common experiences to strike an 'average' which...will be representative of the individual's learning". The quotes Dearborn places about the term "average" enclose a multitude of sins, since here again common sense rather than an experimental method is called upon to tell us what such an average is. And what of the individual who has not been subjected to these vague "average" experiences?

Various types of special intelligence tests have been devised. For example, there is a very heavy emphasis in most intelligence tests on linguistics and in many others on mathematics. Those lacking in training in these fields suffer in the tests; consequently, supplementary performance tests are used when this lack is detected. How does one determine whether or not the subject lacks this training? In most cases by inspection. We recognize a mute or a child who cannot write at all when we see one. The extreme cases offer little trouble. It is the less extreme cases which are difficult. If an individual's learning, say in linguistics, is inadequate for taking a standard test, it may be for (1) lack of ability to learn linguistic manipulation and (2) lack of opportunity to learn (where the ability is present). To know which is the case we have to know what we are trying to measure. There is nothing wrong in making assumptions concerning the subject's abilities before testing him, but the assumptinns should be made explicit and the tests designed to provide confirmation or denial of the presuppositions. Such assumptions are not consciously formulated at present.

Tests in the Binet tradition do not attempt to measure specific abilities that make up intelligence, but do not ignore them. There is recognition that individuals may differ in these special abilities. As Dearborn (1928) observed:

> One individual may be characterized by an extraordinary plasticity or strength of memory, another be conspicious for rare powers of the imagination, and a third for an unimaginative but acutely logical mind. Further specialization of abilities may be found within these general divisions of the mind; one person's memory may be much better for some things than for others; another may reason well in mathematics and poorly in finance. These differences cannot be neglected in describing the intellectual development of any given individual; yet the fact remains that the most important recent advance in our knowledge of the growth of the intellect has come about through a method which obscures these differences by striking a balance or average of the individual's abilities to find a measure of his general, or as it might perhaps better be called,'average' intelligence (p. 66).

Once again that catch-all "average" is called on, this time to justify the method of measuring general intelligence. It is difficult to grasp precisely what is being averaged since there is still considerable disagreement as to what the primary mental abilities are, how many there are, and whether or not they are independent. Trying to average concepts which are in this state is like trying to average the number of chairs, people, glasses, windows and walls in the room. For an average of such abilities to yield a meaningful general measure we have to go considerably beyond the test designer's judgment.

The _feeling_ that we are covering all the abilities is not sufficient. Biases are constantly being uncovered in intelligence tests. For example, linguistics were found to play too important a role in the Binet, the Terman, the National, the Otis, and the Thorndike tests. Efforts have been made to correct for this on a piecemeal basis. However, if the average of abilities is to be meaningful we must have at least the following information:

1. An exhaustive and exclusive list of primary abilities.
2. An experimental definition for each ability.
3. A common scale along which to measure each ability.

Needless to say, none of the current tests are built on such a basis. The most Terman (1916) could say, for example, in praise of the Binet scale is, "It is capable of bounding _roughly_ the vocational territory in which an individual's intelligence will probably permit success, _nothing else preventing_"(p. 49 italics mine).

Whereas the Binet and other standard intelligence tests have attempted to infer the ability to learn from what the individual has learned in the past, Woodrow's tests have been designed to measure the ability to learn directly. * A set of performance tests are constructed and given to an individual repeatedly, so that progress can be measured with respect to the accomplishment of each task. According to Woodrow (1946)

> The performances practiced represented a wide variety of activities, and were the following: horizontal adding, substitution, reproduction of spot-patterns, rearranging letters to make words, cancelling letters with complex instructions, estimating lengths, and speed of making 'gates' (making four horizontal lines and one diagonal one in each square of a page divided into one thousand squares). The improvement score used was the difference between final raw score and initial raw score (p. 151).

---

*As I previously noted, Woodrow does not consider these to be intelligence tests.

There is no evidence that the list of tests that Woodrow offers are representative of a general learning ability. The emphasis in the list is on the visual, with the auditory indirectly implied in the word-tests; but ability to learn with respect to the tactile senses, olfactory senses, and so on, are not included. Even relative to the visual learning ability the tests are restricted; for example, to two-dimensional rather than three-dimensional problems and the performance method is paper-and-pencil throughout.

Woodrow points out that "the practice was long enough so that for the most part the individual learning curves showed a pronounced flattening out towards the end of the practice" (p. 151). This he takes to be a general characteristic of the learning process; that is, when a course of action gets to be known well, the rate of improvement decreases. This flattening of the curve, however, may be correlated with the individual's loss of interest in the problem; that is, with boredom. But if an individual were examined relative to an objective for which his interest did not change, no such flattening might be observed. In other words, a fundamental weakness in the design of Woodrow's tests is the lack of specification of the influencing variables which can influence the learning process. At least one of these variables, relative value of objectives defined by the tests, is not controlled. It should be measured in order to eliminate its effect from the experimental results.

Let us turn now to consider the nature of learning. For Allport (1937) "Taken broadly, the field of learning includes every form of acquisition and modification that occurs in the course of growth" (p. 151). But Allport does not make explicit what is acquired or what is modified. Katz and Schanck suggested that "ways of adjusting to new problems" are acquired. Now we can define a problem situation with respect to the means which an individual can choose; it consists of an environment, an end, and alternative means. What can "new" mean with respect to such a situation? It cannot mean a "new" environment; one who upon his first arrival in Alaska if presented with the problem of addition of numbers is not presented with a "new problem". Nor is it merely changing the goal to a new one; that is, to one the subject has never faced before. Few people have ever measured the circumference of a wheel, but if asked to do so, they would not be confronted with a "new" problem, as we use the term "new" in common parlance. But to one who is unfamiliar with geometry and who has no measuring tape, measuring the circumference of a wheel, may be a "new" problem. What I am suggesting is that a problem is "new" if in an environment where a goal is assigned to or accepted by a subject, he has no knowledge or understanding of the alternative courses of action. Therefore, to speak of the acquisition of new courses of action by an individual is to speak of the increase in his degrees of knowledge and understanding of that course of action with respect to a valued outcome. "Learning," then, may be defined as the increase of these measures.

Intelligence is the measure of the efficiency with which an individual could learn. I say "could learn" rather than "actually does learn" since we are interested in his innate ability, that is, his ability independent of situational characteristics and previous experience. A definition of intelligence should reflect this independence.

As observed above the concept of learning is applicable to increases in knowledge and understanding. Hence 'intelligence' can be applied to learning rates on each of these two scales. Part of the confusion in the discussion of the meaning of 'intelligence' may arise out of this multi-dimensionality of the concept. An ability to acquire knowledge quickly is not necessarily accompanied by an ability to acquire understanding quickly, and conversely.

Time is normally used as the basis for measuring rate of change. But, since different courses of action require different amounts of time to carry out, it may be preferable to use "the number of trials" (N_i) as a basis for measuring rate of change.

3.12. <u>K (Knowledge) Intelligence Function</u> (I_K). A subject's K-intelligence function, relative to a course of action (C_i) for which his degree of knowledge is zero, and a preferred outcome (O_j) of relative value (V_j) equal to 1.0 in a choice environment S is

$$I_K = \frac{d(DK_{ij} \mid S,\ V_j = 1.0)}{d\ N_i}$$

3.13. <u>U (Understanding) Intelligence Function</u> (I_U). A subject's U-intelligence function, relative to a course of action (C_i) for which his degree of understanding is zero, a preferred outcome (O_j) of relative value (V_j) equal to 1.0, and an environmental variable (s) in a choice environment (S') is

$$I_U = \frac{d(DU_{ij} \mid s,\ S',\ V_j = 1.0)}{d\ N_i}$$

These intelligence functions can be generalized to account for the effect of the "given" variables in each. Nevertheless they remain specific to a particular course of action.

In order to obtain a <u>general</u> intelligence function (of either type) of an individual, it would be necessary to use a set of courses of action. Standardization of such a sample is necessary if individuals are to be compared with respect to intelligence. Note however, that the courses of action should be ones of which the subject has no knowledge or understanding before the test. In practice it may be possible to infer from rates of change of knowledge or understanding of a course of action for which there is some (but not complete) initial knowledge or understanding to what would have been obtained had the ideal conditions been met. The more is known about an individual's intelligence, the more likely it is that such extrapolations can be made.

Now we can see the difficulty of trying to represent intelligence by a single number. First, functions cannot be represented adequately by one number. Secondly, even if they could, it would be necessary to deal with distributions over sets of courses of action and choice environments. A completely general intelligence function is almost as complex as the personality function. Few have tried to represent personality by a single number. Many, however, have not shown equally sound judgment when it comes to intelligence.

RELATIVE VALUE AND INTENTION

Up to this point I have made extensive use of the concept "relative value" as it applies to outcomes, but it has yet to be defined. To do so I shall first consider a subject's degree of intention for an outcome, then its utility for him, and finally its relative value.

As in the case of familiarity and knowledge, it is necessary to construct an appropriate idealized standard situation.

3.14. Intention Situation: one in which (a) there are the same number (m) of available (exclusive and exhaustive) courses of action and outcomes, (b) each course of action has maximum efficiency for one outcome and hence no efficiency for any other, (c) each outcome has associated with it one course of action which has maximum efficiency for it, and (d) the alternative courses of action are equally familiar, known, and understood by the subject relative to the possible outcomes.

It is apparent that in such an environment the only objective basis for selecting a course of action is desire for the one outcome it is certain to yield.

3.15. Degree of Intention (DI_j) of a subject (A) for an outcome (O_j) relative to an exclusive and exhaustive set of outcomes $\{O_j\}$ in an intention situation in a choice environment (S) is the probability that A selects that course of action which has maximum efficiency for O_j.

This measure, since it involves probability, has a maximum value of 1.0 and a minimum of O. Because it measures preference for an outcome relative to a specific set of outcomes it is a relative measure.

The measure is also relative to the choice environment. Thus, of a subject's degree of intention for an outcome (e.g., access to water) depends on the alternatives that are available (e.g., soft drinks, beer, liquor, milk, etc.) and the "time and place."

If a subject can have any one, and only one, of a set of beverages, or none by simply pushing an appropriate button or pushing none, then the relative frequency with which he selects each is his degree of intention for each.

The sum of the degrees of intention over an exclusive and exhaustive set of outcomes must be equal to 1.0. If the degree of intention for any outcome is greater than 0.5, it is necessarily preferred to any alternative since this measure can exceed 0.5 for only one outcome in an exclusive and exhaustive set. That outcome in a set for which this measure is greatest is the subject's <u>preferred outcome</u> or <u>objective.</u>

Degrees of intention are not necessarily additive. For example, suppose the following four outcomes are possible:

O_1 : coffee and milk
O_2 : coffee but no milk
O_3 : milk but not coffee
O_4 : neither

It is not necessary that $DI_1 = DI_2 + DI_3$. DI_1 may be either greater than or less than $DI_1 + DI_2$.

In the intention environment we control the efficiencies of the alternative courses of action $\{E_{ij}\}$; the degrees of familiarity $\{DF_{ij}\}$; knowledge $\{DK_{ij}\}$; and understanding $\{DU_{ij}\}$. Therefore, we can formulate an intention function as follows::

3.16. <u>Intention (Relative Value) Function</u>. A subject's intention function for an outcome (O_j) is that mathematical function (f_V) which satisfies the equation:

$$(V_j \mid \{O_j\}, S) = f_V(\{E_{ij}\}, \{DF_{ij}\}, \{DK_{ij}\}, \{DU_{ij}\} \mid \{O_j\}, S).$$

3.17. <u>Generalized Intention (Relative Value) Function.</u> A subject's generalized intention function for outcome (O_j) is that mathematical function f^*_V) which satisfies the equation:

$$(V_j \mid \{O_j\}) = f^*_V(\{E_{ij}\}, \{DF_{ij}\}, \{DK_{ij}\}, \{DU_{ij}\}, S \mid \{O_j\}).$$

The relationship between the degree of intention for an outcome and its utility is revealed by examining what might be called a "utility judgment." In the Case Method of measuring utility (see Ackoff, 1963, pp. 91-93) the subject is confronted with a choice between two outcomes O_1 and O_2 where if he selects O_1 he is certain to obtain it (and hence $E_{11} = 1.0$), and if he selects O_2 he will obtain it with probability α (and hence $E_{22} = \alpha$). The researcher seeks a value of α such that the subject has no preference between "O_1 with certainty" and "O_2 with probability α"; that is, an α for which $P_1 = P_2$. Then the utility of O_1, U_1, is set equal to $E_{22} = \alpha$, and the utility of O_2 is set equal to $E_{11} = 1.0$.

This procedure, then yields measures of utility which are equal to the efficiencies (E_{11} and E_{22}) for which the degrees of intention for O_1 and O_2 are equal (i.e., $DI_1 = DI_2$). This utility measure makes the same assumptions concerning familiarity, knowledge, and understanding as are made in obtaining the degree of intention. However, it makes an additional assumption: that the subject attempts to maximize expected utility (i.e., $E_{ij}\ U_j$).

Any of the various measures of utility which have been suggested can similarly be interpreted as a special case of what I have called the "intention function." These measures and my degree of intention are all measures of the relative value of outcomes but they need not yield equivalent results. For example, the utility of coffee may be 1.00 and of milk 0.25 which when "normalized" become 0.80 and 0.20 respectively. But the degree of intention for coffee may be 1.0 and for milk 0.

It is easier to obtain estimates of utility than of intention because of the stronger assumptions which are made. For many purposes either may be used with equal efficiency. Both are measures of relative value. For my purposes here, however, "relative value" has been and will be used to refer to degree of intention, unless I indicate otherwise.

CONCLUSION

The personality function developed here expresses an individual's expected relative value in a choice situation as a function of the courses of action which are available, the possible outcome, and relevant environmental variables. Expected relative value can also be expressed as a function of probabilities of choice, efficiencies of courses of action, and relative values of outcomes. Hence the personality function was decomposed into three functions:

1. The familiarity function which relates probability of choice to other characteristics of the choice situation.
2. The knowledge function which relates efficiency of choice to other characteristics of the choice situation.
3. The intention function which relates the relative value of an outcome to other characteristics of a choice situation.

If these three functions were known, the personality function would be also.

The discussion in this chapter has been directed to providing the researcher with a conceptual framework within which to analyze a subject's choice. The subject's conceptualization of the choice situation, however, may differ widely from that of the researcher. We shall consider the subject's conception in detail in Chapter 5. Until we do so it is not possible to make explicit the nature of the expectations which are an output of the "choice box" shown in Figure 2.1. These expectations are fed into the outcome-evaluation function and play an important role in the subject's behavior subsequent to his taking action.

Several aspects of these expectations should be considered here. First, note that the term "expectation" is used in a psychological, rather than in a statistical sense. For example, suppose the subject estimates the efficiency of the course of action he selects as 0.9 for an outcome whose relative value is 0.8, and 0.1 for an outcome whose relative value is 0.2. Then the "statistically" expected relative value is 0.9 (0.8) + 0.1 (0.2) = 0.74. He will in fact obtain an outcome whose relative value is estimated at either 0.8 or 0.2. Psychologically his expectation is the 0.8 units of relative value, not 0.74. Therefore, if he does not meet his psychological expectation (i.e., he obtains only 0.2 units of relative value) he may consider the problem unsolved and reopen the choice situation with the information on his failure as an input. That is, the psychological expectation involves what might be called a satisficing criterion: a relative value of outcome such that if the outcome that occurs is less valuable than this, he reopens the problem, otherwise he closes it.

For example, the subject's statistical expectation of earnings on a certain investment may be $500. He may, however, be dissatisfied with any return less than $750; should he obtain a return of anything less than $750, he will reexamine his choice.

In principle, the minimal acceptable level of outcome, the satisficing point, is a function of the subject's estimate of the cost (in general sense, not necessarily monetary) of reopening the question and the potential returns from so doing. The satisficing point, then, is the minimal relative value of outcome, improvement over which does not appear to the subject to justify the cost of reopening the question.

Satisfaction involves an intention not to change a situation; that is, an individual is satisfied with a situation if he has less intention to change it than to keep it as it is. Therefore, outcomes below the satisficing level are ones the individual intends to change if they occur. I shall consider "satisfaction" in more detail in Chapter 7.

BIBLIOGRAPHY

Ackoff, R. L., The Design of Social Research. University of Chicago Press, Chicago, 1953.

Allport, F. H., Social Psychology, Houghton Mifflin, Cambridge, 1924.

_______, Theories of Perception and The Concept of Structure. John Wiley & Sons, Inc., New York, 1955.

Allport, G. W., "A Test for Ascendance-submission," Journal, Abnormal Social Psychology, 1928. 23, p. 118-136.

______, Personality: A Psychological Interpretation. Henry Holt, New York, 1937.

______, Personality and Social Encounter. Beacon Press, Boston, 1960.

______, Pattern and Growth in Personality. Holt, Rinehart, and Winston, New York, 1961.

Allport, G. W. and Odbert, H. S., "Trait Names: A Psycholexical Study," Psychological Monographs, 47 (1936).

Brand, H., ed., The Study of Personality, A Book of Readings. John Wiley & Sons, Inc., New York, 1954.

Burt, C., "Symposium on Personality, I, The Assessment of Personality," British Journal of Educational Psychology, 15 (1945), Part 3.

Cattell, R. B., Description and Measurement of Personality. World Book Company, Yonkers, New York, 1946.

Churchman, C. W., and Ackoff, R. L., "An Experimental Definition of Personality," Philosophy of Science, 14 (1947), 304-332.

Dearborn, W. F., Intelligence Tests. Houghton Mifflin Company, New York, 1928.

McGeoch, J. A., The Psychology of Human Learning. Longmans Green & Company, New York, 1942.

Murray, H. A., Explorations in Personality. Oxford University Press, New York, 1938.

Prince, M., The Unconscious. Macmillan, New York, 1924.

Stagner, R., Psychology of Personality. McGraw-Hill Book Co., Inc., New York, (3rd ed.), 1961.

Stephenson, William, The Study of Behavior. The University of Chicago Press, Chicago, 1953.

Stern, W., General Psychology from the Personalistic Standpoint. (trans. by H. D. Spoerl) Macmillan, New York, 1938.

Terman, L. M., The Measurement of Intelligence. Houghton Mifflin Co., New York, 1916.

Thorpe, L. P., and Holliday, J. N., Personality and Life. Longmans Green & Company, 1928.

Thurstone, L. L., and Chave, E. J., The Measurement of Attitude. University of Chicago Press, Chicago, 1929.

Watson, J. B., Behaviorism. W. W. Norton & Company, Inc., New York, 1925.

Woodrow, H., "The Ability to Learn" Psych. Rev. Vol. 53, #3, May, 1946, pp. 147-158.

Dewey, J., Logic, The Theory of Inquiry. Holt, Rinehart, Winston, New York, 1938.

English, H. G., and English, A. C., A Comprehensive Dictionary of Psychological and Psychoanalytical Terms. David McKay, New York, 1958.

Eubank, E. E., The Concepts of Sociology. D. C. Heath, New York, 1932.

Fiske, D. W., "Problems in Measuring Personality," (1963) in Heine and Wepman's Concepts of Personality, 1963, pp. 449-474.

Getzels, J. W., and Jackson, P. W., Creativity and Intelligence. John Wiley & Sons, Inc., New York, 1962.

Heine, R. W., and Wepman, J. M., editors: Concepts of Personality. Aldine Publishing Company, Chicago, 1963.

Kantor, J. R., A Survey of the Science of Psychology. Principia Press, Bloomington, Indiana, 1933.

Katz, D., and Schanck, R. L., Social Psychology. John Wiley & Sons, New York, 1938.

Knapp, R. H., "The Psychology of Personality" in The Behavioral Sciences Today. ed. by Bernard Berelson, Basic Books, New York, 1963, pp. 151-164.

Kohler, W., Gestalt Psychology. Liverwright, New York, 1929.

Kretschmer, E., A Textbook of Medical Psychology. (trans. by E. B. Strauss) Oxford University Press, London, 1934.

McClelland, D. C., "Personality: An Integrative View" in Psychology of Personality. ed. by J. L. McCary, Logos Press, New York, 1956, pp. 321-67.

OBSERVATION: PERCEPTION AND CONSCIOUSNESS

SELF-EVIDENT, adj. Evident to one's self and to nobody else (Ambrose Bierce, The Devil's Dictionary).

INTRODUCTION

Observations provide the information on the basis of which individuals formulate problems, select courses of action and evaluate the outcomes of their behavior. This chapter deals with the nature of observation. Chapter 9 deals with the nature of information.

I should like to emphasize at the outset that this chapter does not attempt to provide a theory which explains how one perceives. Rather it provides a conceptual framework over which such a theory can be constructed. The need for such a framework was observed by F. H. Allport (1955):

> Probably no one would maintain that the present situation, in which we have thirteen theories of perception, all aiming in some degree to be general, yet nearly all different, is a happy one (p. 611).
>
> If we could discover a way of conceptualizing such a dynamic structure in clear denotational terms, we might find ourselves in possession of a concept that would bring together the current generalizations of perceptual theory (p. 613).

The effort here is directed toward the development of such a way.

The terms 'observation,' 'perception,' 'sensation,' 'awareness,' and 'consciousness' are often used interchangably in both ordinary and technical discourse. Most dictionaries define these concepts circularly. I am going to distinguish between them in a way which I believe is useful but which may not be completely justified by either common or technical usage. As I shall try to show, however, my usage is not completely arbitrary.

In considering an individual who observes something, X, I shall speak of X as the <u>stimulus</u> and of his <u>response</u> to it as a <u>perception.</u> Unfortunately, the terms 'stimulus' and 'response' have come into ill repute in psychology because they have usually been treated mechanistically; that is, as synonymous with deterministic 'cause' and 'effect.' Here, however, I treat these concepts functionally, as synonymous with 'producer' and 'product.'

4.1. <u>Stimulus-Response.</u> A stimulus is a producer of a purposeful choice; that is, of a course of action by an individual in a purposeful state. The course of action that is produced is the response.

In dealing with these concepts it will be important to consider the <u>intensity</u> of both the stimulus and the response. I have done so briefly in the discussion of the ascendant-submissive trait test in Chapter 3 and Appendix 1, but here I amplify.

4.2. <u>Intensity of a Stimulus</u>: a measure of a property (of a stimulus) which produces a response.

The intensity of a stimulus may be treated either structurally (e.g., the brightness of a color or the size of an object) or functionally (e.g., the decrease in efficiency of the behavior of the victim of an aggressive act). A stimulus may increase in intensity with respect to one of its properties (e.g., the frequency of sound), and decrease with respect to another (e.g., volume). Which property is used as a basis for measuring the intensity of the stimulus depends on the purposes of the observer. When the intensity of a stimulus is used as an independent variable in experimental work, values of the stimulus with respect to properties other than those used to define its intensity are usually held constant.

> 4.3. Intensity of a Response to a Stimulus: a measure of a property (of a response) which is produced by the stimulus of the response.

The property used to measure the intensity of a response may also be either structural (e.g., when frightened, the loudness of a scream, the distance of a withdrawal, or the speed with which action is taken), or functional (e.g., when aggressed upon, the effect of the response on the efficiency of the aggressor's behavior, as in the ascendant-submissive situation).

PERCEPTION AND OBSERVATION

The general class of responses with which we are first concerned may be called perceptions. Wheras all perceptions are responses to stimuli, not all responses to stimuli are perceptions.

> 4.4. Perception: a response to a stimulus which also produces a change in at least one structural property of the respondent.

Thus a perception is a two-stage production process which is shown schematically in Figure 4.1. In perception there are two products of the producing stimulus (X). First, the stimulus produces structural changes in the respondent (Y). I call this a reaction because this change in the respondent is not a purposeful choice. The tendency of the respondent to react to a stimulus I call his sensitivity. The reaction to a stimulus is its effect on the respondent's senses: sight, hearing, touch, taste, and smell. For example, such changes of structural properties as the vibration of the ear drum,

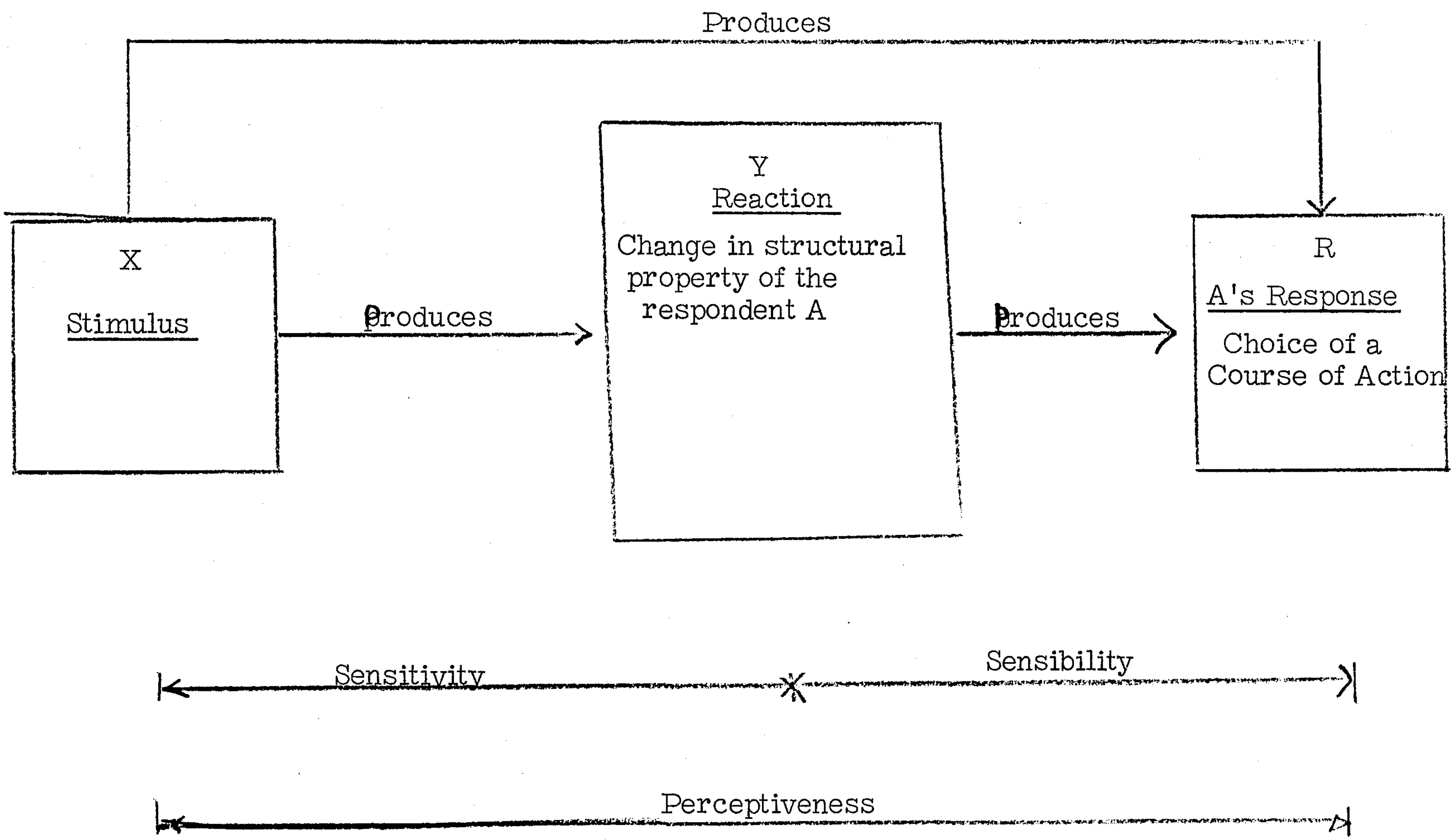


Figure 4.1. Model of Perception.

the formation of an image on the retina, and the associated changes in the nervous system and brain are reactions. These reactions are not under the respondent's control.

The subject's responsiveness to a change in one or more of his structural properties is his sensibility. A sensation, his response to the structural change, cannot occur unless there has first been a reaction; for example, a blind person cannot respond to a flash of light because he cannot react to it. Not every reaction is accompanied by a response. For example, when among a large group of people, we may see (react to) someone but not notice (respond to) him. A sleeping person may react to a sound but not respond to it. Reflex actions are reactions, but not responses. Non-purposeful entities can react, but only purposeful entities can respond. A photoelectric cell can display sensitivity (react) to light but it cannot respond to it.

Now let us look at reactions and their related properties in more detail.

4.5. Reaction to a stimulus (X) by a subject (A) is a change in one or more of A's structural properties that is produced by X.

4.6. Intensity of Reaction to a stimulus (X) by a subject (A) is a measure of a structural property of the reaction produced by X.

4.7. Degree of Sensitivity to a stimulus (X) of specified intensity of a subject (A) in a structurally defined environment is the probability that A will react to X in that environment.

4.8. Sensitivity Function of a subject (A) to a stimulus (X) is a mathematical function which relates his degree of sensitivity to X to its intensity and the structural properties of the environment.

Proceeding in parallel, corresponding concepts relating to sensation can also be defined as follows:

4.9. Sensation: a response by a subject (A) to a change in one of his structural properties.

The intensity of response has already been defined in 4.3.

4.10. Degree of Sensibility of a subject (A) to a change in one of his structural properties (Y) in a purposeful state is the probability that A will respond to Y in that state.

4.11. Sensibility Function of a subject (A) to a change in one of his structural properties (Y) is a mathematical function which relates his degree of sensibility to Y to the intensity of Y and the properties of his purposeful state.

As I have treated the terms a 'sensation' is a response to a change in one's own (structural) properties, whereas a 'perception' is a response to something external which produces a change in one's own properties. In this way I have tried to capture the essence of the distinction made between these concepts in Webster's Universal Dictionary (1936):

> Sensation is mere feeling without an object: perception is the mind's apprehension of some external object as occasioning that feeling.

Hence, a perception involves a response to the producer-product relationship between stimulus and reaction; a sensation does not. For example, one may feel (sense) cold without responding to what produced it. On the other hand, in sensing cold one may perceive a draft.

F. P. Kilpatrick (1961) observed that

> A given physiological stimulus pattern may be produced by an infinity of different external conditions (p. 443).

Therefore, in mere sensation, if the stimulus that produced a reaction were to change but not the reaction, the sensation would remain unchanged; but this is not so for perception.

When a psychologist attempts to explain in psychological terms different responses by two persons to the same stimulus, he should first assure himself that they have had the same reactions. For example, a color-blind person may respond to a traffic light differently than a person who is not color-blind because of the difference in their reaction. In studies of perception, however, it is not uncommon to assume that different subjects react similarly to the same stimulus.

The physiologist, rather than the psychologist, is concerned with an individual's reactions and sensitivity. The psychologist is primarily concerned with sensation and perception. The physiologist attempts to determine whether the subject "receives the signal," and the psychologist is primarily concerned with what the subject does with it once he "has it." The psychologist is concerned with how an individual responds to what he can react to.

Measures analogous to those of sensitivity and sensibility are also applicable to perception

4.12. Intensity of Perception of a stimulus (X) by a subject

(A) is the intensity of A's response to the reaction produced by X.

4.13. Degree of Perceptiveness of a subject (A) to a stimulus (X) in a purposeful state is the probability that A will respond to X in that state.

4.14. Perceptiveness Function of a subject (A) to a stimulus (X) is a mathematical function which relates A's degree of perceptiveness of X to the intensity of X and the properties of his purposeful state.

Notice that the degree of perceptiveness of a subject is the product of the probability that he will react to the stimulus and the probability that he will respond if he reacts. The intensity of a perception is also a function of the intensities of reaction and response.

It is apparent from the measures defined above that we can study an individual's perceptions in different ways. First, we can attempt to determine how his degree of perceptiveness of a certain type of stimulus relates to the intensity of that stimulus. Secondly, we can attempt to determine how the intensity of his response to a stimulus relates to the intensity of the stimulus. We can, in addition, combine these considerations. For example, for any intensity of stimulus we can record some function of the intensity of the response and the degree of perceptiveness of the subject; for example, we can plot the average intensity of response or the variance of the intensity. We can also conduct research to determine how these response characteristics of a subject vary in different choice situations. For example, an individual may be very perceptive of noise when he is pursuing an objective of high relative value, but not so when he pursues something of low relative value.

Whether or not an observer can say that another individual has a sensation or perception depends on what properties he uses to characterize the stimulus. Sensations and perceptions, like other psychological events, are not "just there; " that is, in the subject for us to observe. Whether or not we observe them depends on the conceptual scheme we bring to our observations: on what we look for in the stimulus as well as in the subject. This is equally true with respect to observation of physical properties of inanimate objects; for example, in describing a rubber ball we do not usually refer to its taste, but we could.

According to the definitions formulated above an individual can be said to perceive something only if he responds to it. Off hand this may seem to run counter to common sense. We might be willing to assert that an individual perceives the color of, say, a pencil without his having responded to it. This is the type of argument that introspective psychologists have used so often. They argue that only the subject can know whether or not he perceives something, and hence we must ask him and hope he answers truthfully.

What can it mean to say that an individual senses or perceives something but that it may never affect his behavior? When someone tells us that he can discriminate between quarter tones or that despite our concealment he had perceived our presence, we are likely to be skeptical unless we have seen evidence of perception in his behavior. We can, in fact, conduct tests (such as will be discussed below) to determine whether or not he can perceive differences in quarter tones or our presence; tests which involve observing his behavior, and not his testimony alone. What he "says" is, of course, a type of behavior, but it may not be the best type of behavior to use in such tests.

Changes of behavior in a subject may occur which we do not observe because of our conceptual set, or even if our conceptual set is adequate, because they are almost imperceptible to us. Our techniques for observing the responses of others are by no means perfect, but it is more constructive to attempt to improve our techniques of observing responses than to assign another's perceptions to the realm of the unobservable.

Suppose we want to determine whether an individual using a pen with blue ink perceives the blueness of the ink. Clearly we would change only the color of the ink in a way that is undetected by the subject and observe whether his behavior changes in a functional way. If, for example, he discards the pen and selects another with blue ink, we would assert that he had perceived the blueness of the ink. The subject may not change pens but only examine it or inquire about someone "playing around" with it. This would be evidence of his perception of the color of the ink.

An individual may perceive one property of an object or event but not perceive others, or he may perceive different properties of the same thing at different times. What he perceives at any time is related to the conceptual model that he brings to his observations and this depends, in turn, on his desired outcome in his purposeful state. The dependence of an individual's perceptions on the characteristics of the purposeful state in which they occur has been neglected by many psychologists. They conceive of an individual as having perceptions in a psychological vacuum, and hence they think of perception as a type of mechanical response, of marks being made on a blank wax tablet. Data obtained by perception are therefore thought of as "givens" rather than as "takens."

Here I conceive of what is perceived as equally dependent on the observer and the observed. This interdependence will be discussed again below and in detail in Chapter 5 where I consider the subject's model of his purposeful state. The interdependence of a subject's model and his observations enables us to explain why two individuals perceive different properties of the same thing under the same circumstances or why one individual perceives different properties of the same thing under different circumstances. Recognition of this interdependence has become increasingly important in the psychology of perception.

The study of perception was one of the earliest activities in which psychologists engaged. Such research by Fechner and Weber in the last century is frequently cited as the origin of <u>scientific</u> psychology. Psycho-physical experiments, the essential characteristics of which are effectively discussed by Hirsh (1958), are concerned with a subject's response to structurally defined stimuli which either he or the experimenter controls. It is assumed implicitly in most of these experiments that the laboratory in which they are conducted provides a choice situation in which the subject will display his maximum degree of perceptiveness and intensity of response. There is usually no effort to test this assumption or even to identify the parameters of the choice situation. Hence the subject's intentions and the efficiency of his choices relative to his objectives are not taken into account. In effect he is experimented on much as if he were a machine whose sensitivity is being tested. E. A. Singer (1924) cautioned against such treatment. The only essential difference between sensitivity tests and most psychophysical experiments is the reliance on the subject's verbal responses. Care is usually taken, however, to determine the consistency and reliability of these.

In the last two decades increasing amounts of attention have been given to what the subject brings to his perceptive experience. The importance of the individual's set in what he perceives was stressed by the Gestalt psychologists. Cantril (1950), Ittelson (1952), Kilpatrick (1961), and Bruner (1956) have tried to make explicit what is brought to perception. Our concern here, however, is not with how sensations and perceptions occur, but what they are.

The individual who perceives a stimulus (X) may respond to either its structural or functional properties. The set of structural properties of the stimulus to which he responds constitute his description or image* of the stimulus. The set of functional properties to which he responds constitute his explanation or concept* of the stimulus. Thus, if a person responds to the size, color, and weight of an automobile these properties are part of his image of an automobile. If he responds to its capability for transporting and protecting him from the rain, these are part of his concept of an automobile.

There is nothing in this treatment of description and explanation which requires a description or an explanation to be correct. Whether or not they are correct depends on the efficiency of the corresponding set of the observer's responses for his desired outcomes. The more efficient they are, the more correct they are.

As indicated above, what properties are contained in a description or explanation depends on the observer's model of the

*In Chapter 9 I shall explore the nature of images and concepts in more detail.

situation he observes as well as on the situation itself. For example, in any situation there are an indefinitely large number of structural and functional properties to which an individual can respond but he "selects" those that are relevant to his purposes. Thus a longshoreman who loads sugar on a ship, a chemist, and a dietician may all describe sugar differently. We recognize at the commonsense level that what we observe in any situation depends on our "point of view." Our point of view is the model we bring to the situation. Hence there are as many correct descriptions and explanations of a situation as there are different objectives which can be pursued in it.

An individual may describe a situation correctly without explaining it correctly, and conversely. For example, one person may have seen an auto accident and describes it accurately without being able to explain it. A medical examiner who did not see the accident or receive another's description of it may explain it after determining that one of the drivers was drunk.

Not all perceptions are observations; but all observations are perceptions. Observations are a special type of perception:

> 4.15. Observation: a perception of a stimulus (X) by a subject who intended to perceive X.

Hence, 'observation' is used here to connote a deliberate or desired perception. Most of our perceptions are not intended; they occur by chance. When we see something accidently we would not usually say that we had observed it. On the other hand planned data gathered by a scientist are usually called 'observations'.

Some finer distinctions are possible, distinctions on which I do not dwell here. For example, when an individual observes something for the purpose of evaluating it relative to some purpose which he has, he inspects the stimulus. If he looks for a specific property of a stimulus, he examines it.

An individual may respond to a structural or functional property of an object or event without perceiving it. For example, each of us has frequently responded to such properties of people whom we have never seen or heard, or of places to which we have never been. We have done so because information about them has been communicated to us. (I shall explore this use of communication in depth in later chapters.) Those things which a person has not perceived but to which he responds are ones of which he can be said to be aware. But he can also be said to be aware of things which he has perceived.

> 4.16. Awareness. An individual is aware of something (X) if he responds to X.

A person may be aware of things he does not now perceive but once perceived if they are preserved in his memory (a subject to be discussed in Chapter 5). Similarly, he may be aware of things about which he was informed in the past. Hence, to perceive something is to be aware of it, but to be aware of it is not necessarily to have perceived it. Therefore, perception is a special case of awareness.

Consciousness, on the other hand, is a special case of perception, a case to which I now turn.

CONSCIOUSNESS

"Consciousness" has been one of the most enigmatic concepts in psychology and philosophy. One group of psychologists and philosophers have insisted vehemently that there is no such thing: that it is a useless intervening variable. Another group has insisted that it is basic and its meaning is obvious. For example, Freud (1933) wrote, "What is meant by consciousness we need not discuss; it is beyond all doubt" (p. 99). J. G. Miller (1942) collected a large number of definitions of consciousness and showed the difficulty of finding a common ground among them. However, Singer (1929) performed a logico-historical analysis of uses of the term and found recent usage to be returning to its original meaning: thinking with.

Singer went on to analyze the meaning of consciousness in more detail. According to him, one observer (B) can observe a stimulus (X) and the response to it of an individual (A), and hence B can observe A perceiving X. In describing how this can be done

> ...we must...have described all the stimulus-response relations any observer C would have to establish in order to convince himself experimentally that in B's mind existed such knowledge or perception as might be called B's perception of a sensation in the mind of A. In other words, one who has established the only conditions under which an observer B could know that yonder was an organism A possessed of the knowledge called a sensation, cannot but have defined the only conditions under which a second observer C could know that yonder was a first observer B possessed of the knowledge of a sensation in a third mind A (p. 565).

Singer then asked, "What should one call B's perception of a sensation experienced by A?" and answered, "My suggestion would be that just this class of mental state be called conscious" (p. 566). Then Singer pointed out that it is quite possible for B to be conscious of states of mind in A of which A himself is unconscious. Furthermore, B and A may be the same person for

> It will be seen that nothing in this definition of a conscious moment requires the mental state which is to be the stimulus to lie in a mind other than the conscious mind itself; but neither is there anything to exclude this possibility (p. 566).

Singer's concept of consciousness is reflected in the writings of others. For example, E. R. Guthrie (1938) wrote:

> In the inclusive sense of the words, consciousness and awareness are made up by our own secondary responses to our own movements. We may absently brush aside a tickling hair on our forehead, or ease our cramped position on a chair without being aware of it. Awareness of our own movements requires that the movement itself be responded to, be noticed. Noticing our own primary responses to an external situation is itself a secondary response (p. 357).

A similar view was espressed by Y. H. Krikorian (1938):

> If to be conscious means a mental state knowing another mental state, the 'another' can be either my prior mental state or my neighbor's mental state... To be conscious means to respond cognitively to a stimulus which is itself a response (pp. 159-60).

Although Freud deliberately avoided defining consciousness, since he thought its meaning to be obvious, he referred to it as a "seat of awareness" which perceives some mental states but not all. It is like a sensory organ which senses other sensations (1933, p. 224). C. G. Jung (1924) vaguely suggested the same thing:

> ...by consciousness I understand the relatedness of psychic contents to the ego...insofar as they are sensed as such by the ego. Insofar as relations are not sensed as such by the ego, they are unconscious. Consciousness is a function or activity which maintains the relation of psychic contents with the ego (pp. 535-36).

It follows then that one individual (B) is conscious of another individual's (A's) sensation or perception of a stimulus (X) if B perceives A's sensation or perception of X. It is not sufficient for B to respond to A's response for him to be conscious of it; he must respond to the producer-product relationship between the stimulus and the response, and hence to the fact that it is a response. For example, while talking with a friend in my office recently, he rose, put on his top-coat, and sat down again. I then arose and closed an open window in the room. My friend perceived the cold because he responded to it purposefully, putting on a coat. I responded to his behavior by shutting the window. I was therefore aware of the cold and consciousness of his perception of it. He had not perceived the open window and hence was surprised when I closed it. Furthermore, he was not conscious of his response to the cold until he became conscious of my response to his.

Consciousness includes perception of another's perception, but it is not exhausted by such perception; it includes perception of any <u>mental state</u> of another. For example, one can be conscious of another's intentions, feelings, preferences, traits, beliefs, and so on. Hence, in order to define 'consciousness' it is first necessary to define a 'mental state.'

> 4.17. <u>Mental State</u> of a subject A is any one or combination of functional properties of an individual's purposeful behavior.

Definition of mental states is the preoccupation of this book. In fact, it is concerned with the development of a methodology which facilitates one person becoming conscious of another.

Now 'consciousness' may be defined as follows:

> 4.18. <u>Consciousness</u>. One individual (B) is conscious of another individual's (A's) mental state if B perceives A's mental state.

Hence if B perceives what A perceives, remembers, believes, thinks, feels, or any other functional property of A's purposeful behavior, B is conscious of that property (mental state) of A.

4.19. <u>Selfconsciousness</u>. An individual (A) is selfconscious if he perceives one or more of his own mental states.

Peculiarly, there is considerably more agreement as to the meaning of 'selfconsciousness' than there is concerning the meaning of 'consciousness.' A. A. Roback (1933) summarized this general agreement as follows:

> To the philosopher and laboratory psychologist, particularly of the structural school, selfconsciousness means the act or condition of being (or the process of becoming) directly aware of the self or ego during any mental process, or in other words, awareness of what we experience as relating to a self as the subject of these experiences (pp. 1-2).

At first glance it may seem that once 'consciousness' has been defined it should be relatively easy to define 'unconsciousness.' So it seemed to H. H. Goddard (1925) who wrote:

> ...the term 'unconscious' can mean nothing but <u>not conscious</u> ... Everybody knows the state or condition of consciousness and, therefore, the state of unconsciousness... (p. 315).

One wonders why something so universally known has been the subject of so much unsuccessful definitional effort.

The meaning of a negative of a term depends on the universe of discourse to which the term applies. For example, although Miller (1942) observed that the term 'unconscious' has often been applied to inanimate things (p. 22), most psychologists take a position like that taken many years ago by K. Koffka (1929):

> The unconscious as a systematic concept is not synonymous with nonconscious... The movements of a stone are not called unconscious, whereas those of an amoeba might be (p. 43).

The problem of defining 'unconscious' consists first of specifying the universe of discourse to which it applies and then dividing it into the two exclusive and exhaustive domains of consciousness and unconsciousness. The former requirement hinges on the question as to whether unconsciousness refers to nonresponses to stimuli or responses to stimuli that are not conscious. Miller (1942) argues for the former:

> A person is unconscious... when he is one of the states in which the stimuli of the external environment are not affecting his behavior or in which he does not show normal reactions to or discrimination of the stimuli (p. 23).

When we speak of a person being unconscious we sometimes seem to mean that he is in an unresponsive state; for example, when he has been "knocked unconscious" by a blow on the head. However, we also use 'unconscious' in another sense, in a psychoanalystic sense, which is quite different. In this sense, the unconscious is taken to contain experience that is suppressed or hidden in "the dark recesses" of the mind. But if there were no reactions or responses to a stimulus there would be nothing to be hidden. Hence, in the psychoanalytic sense unconsciousness involves receiving inputs which are not readily accessible to the receiver. One can either "receive" a stimulus (i.e., react) but not respond to it, or he may respond but not respond to that response.

Both concepts described are important. I have chosen to label them as follows:

4.20. Unconsciousness: An individual (B) is unconscious of another individual's (A's) mental state if B perceives A but not A's mental state.

For example, if B perceive's A but not A's perception of an X, B is unconscious of A's perception of X.

4.21. Nonconsciousness: An individual (B) is nonconscious of another individual (A) if B does not perceive A.

4.22. Unselfconsciousness: An individual is unselfconscious of his own mental state if he perceives himself but not his mental state.

For example, if an individual perceives an X but does not perceive that he perceives X, he is unselfconscious of that perception.

4.23. Nonselfconsciousness: An individual is nonselfconscious if he does not perceive himself.

Hence, if an individual perceives something of which he is not conscious, I say he is unselfconscious of it. If he cannot perceive it, as when "knocked out," I say he is nonselfconscious of it.

Much of the activity of psychoanalysis can be viewed as bringing into consciousness functional properties of past responses of the subject and others, properties which previously were not responded to, previous perceptions of which the individual was not, or has lost, selfconsciousness.

The Content of Perception and Observation

To perceive or observe objects, events, and their properties is to respond to them. The stimulus which is responded to may produce a change in the responder's probabilities of choice, efficiencies of choice, or relative values. As we shall see, a message may affect an individual in the same ways. A message, of course, is itself a type of stimulus which the subject may perceive. If, for example, he responds to another's utterances as "noise," he responds to its structural properties. Therefore, he may sense a message to whose function he does not respond.

Having perceived something a subject may store it in his memory or use it. To do either he must construct a representation of what was observed, the stimulus. To do so he uses signs, including images and concepts. (The nature of signs is considered in Chapter 9.) Perceptions are thus converted into propositions or statements which are used by the subject to communicate either with others or with himself. In thinking, an individual communicates with himself, often by talking to or writing to himself.

The form of perceptions is revealed by an analysis of the form of statements which represent them. Such an analysis will be made below. It presupposes, however, understanding of the content of perceptions; to which we now turn. Since what we perceive are properties, individuals, and events, we take up each of these in turn.

Properties

We usually think of a property as something belonging to an object or event independently of the observer of that object or event. But when we reflect on the way an observer determines whether or not an object or event has a certain property it becomes clear that what we mean by a property is "what it can do to him under certain circumstances." For example, we say a body is heavy if it requires a great deal of effort to lift it, or, if when it is placed on a scale, a certain reading can be made (i. e., responded to).

> 4.24. Property. A property is a potentiality for producing a specified type of response (R) in a subject (A) in a specified choice environment (S).

Hence, when we say that an object is red, for example, we mean that when it is placed in a certain environment it will produce a particular kind of response from an observer. If it does not produce the specified response then it does not have that property for him.

For example, a color-blind person would not show the specified response for "redness."

In addition to properties which an object or event may have at a moment of time or over an interval of time, there are two types of "derived" properties; changes in or the rate of change of a property (1) under constant environmental conditions, and (2) under changing conditions. In order to define adequately a property of an object or event at a moment of time it is necessary to specity:

(1) the "things" to be observed,

(2) the observer(s),

(3) the environment within which the observations should be made,

(4) the operations (courses of action), if any, which should be carried out in that environment,

(5) the instruments, if any, and the metric standards which are required to carry out the specified operations, and

(6) the observation(s) (responses) which should be produced.

First consider structural properties. Suppose, for example, that we want to define what is meant by the statement that a particular object "is red." We might proceed as follows. (1) We identify the object whose color is to be determined. (2) We identify the observer or kind of observer to be used. (3) We specify the environment in which the color is to be determined; for example, the atmospheric conditions, temperature, and lighting conditions. (4 and 5) We specify where the object should be located in the environment and what instruments (e.g., spectroscope) should be used and how. (6) We designate the spectral range of wave lengths (say 0.00006 to

0.00008 cm) into which the reflected light should fall and how this should be observed.

Note that to observe that an object is red we need not and seldom go through all this. We observe it under normal conditions. If, on the basis of what we do observe and our concept of the effect of the differences between these conditions and the "defining" conditions, we believe that the defining responses would be observed under the defining conditions; then, we conclude the object is red. For example, an object which appears orange under yellow light may nevertheless be said to be red. The property observed in this case is "orange," but the property attributed is "red."

The attribution of a property to an object or event, then, is not an "immediate" mental act; it is an inference from what is observed in one situation to what would have been observed in another situation.

The first type of derived property involves a change in a property over time in a constant environment. Definition of such a property requires the six steps listed above plus a specification of the time interval over which the observations are to be made, the timing of the observations, and the way in which the observations (data) are to be treated. "Rate of dissolution," for example, is such a property. It involves the length of time required for an object to change certain of its structural (e.g., chemical) properties under constant conditions (e.g., while immersed in a specified liquid). "Rate of deformation" of a structural member of a building under constant load is a similar property. The so-called _life properties_ of goods, tools, and equipment fall into this class of structural properties. The life of a lamp bulb, for example, might be defined as the length of time it emits light in a specified constant environment.

The second type of derived structural property involves changes of a property under changing conditions. The form of this definition is similar to the preceding one with the additional requirement for specification of what changes in the environment should be made and how they should be timed. Observations always involve responses to these changes. The coefficient of linear expansion of an object is an example of such a property, as is the coefficient of volume expansion and the coefficient of compressibility. So-called sensitivity properties all fall into this class as well. The sensitivity of photographic paper to light, of an explosive to impact or heat, of a structure to shock, and so on, can all be defined within the form described.

Now consider functional properties. Since the meaning of function is rooted in the meaning of the producer-product relationship, it is not surprising that the essence of all functional concepts lies in a measure of probability of production. This probability may be of either of the following types:

(1) The probability that an individual object or group will select a specified course of action.

(2) The probability that a specified course of action will produce a specified outcome.

These probabilities correspond to measures of preference (familiarity) and efficiency. The measure of every functional property reduces to a measure of one or both of these types of probability. For example, in the concept of ascendancy discussed in Chapter 3, the degree of ascendancy was defined as the probability that an individual would select a type of action which reduces the efficiency of a co-occupant of his environment.

Functional properties are of three general types, corresponding to the types of structural properties already considered:

(a) the property of something at a moment of time,
(b) the change or rate of change in a property under constant conditions, and
(c) the change or rate of change in the property under changing conditions.

Definitions of the first type should contain specification of

(1) the object or class of objects to be observed,
(2) the conditions (environment) under which the observations should be made,
(3) the operations, if any, which should be performed in that environment,
(4) the instruments, if any, which are required to perform the specified operations,
(5) the observations which should be made, and
(6) the treatment of the data obtained.

This content is quite similar to that of the definition of the corresponding type of structural property. If no observational error (the nature of which is discussed below) is involved in the determination of a structural property, only one observation need be made. In the case of a functional property, however, even where no observational error is present, an infinite number of observations are required (in principle) in order to determine the appropriate probabilities without error.

An example of such a property is "the degree of familiarity of an individual with a course of action relative to an outcome." First we identified the subject and the outcome involved. Next we specified the conditions under which the observations should be made. These were:

(a) A set of alternative courses of action (C_1, C_2, . . .) are available in the environment.

(b) Each of the available courses of action has perfect efficiency for the specified outcome.

(c) The individual has interest in only the specified outcome.

In this situation we should observe the frequency with which the individual selects each course of action. Then the degree of familiarity of the individual for a specified course of action, C_i, relative to the outcome in that environment is the probability (limiting relative frequency) of his selecting that course.

As can be seen from this definition, "familiarity" is a preference-type property. All functional properties relating to preferences should have definitions of the form indicated. The same is true for "dispositional" properties; for example, hungry, tired and bored.

The second type of functional property is similar to the first except that our concern is with changes in probabilities under constant conditions over time. These properties are analogous to such structural properties as solubility or rate of deformation under constant load. People, for example, become tired of certain things after a while, or else may become increasingly fond of them. This simply means that their preference patterns change over time. The same may be true of, say, an inspection machine whose probability of rejecting an acceptable item may change with use of the machine. Another set of properties of this type involves changes in the rate of performance of a task with its repetition.

The definitional form of such a property is similar to the first except that the way of measuring the change in the relevant property must be specified, as must the time interval to be covered and the frequency or timing of the observations. For example, one could measure the change of degree of familiarity with a course of action as the change in this degree between two moments of time, t_1 and t_2; or one could measure the average rate of change of this probability measure with respect to time (i. e., the average derivative with respect to time).

The third and last class of functional properties involves probability of choice or efficiency under changing conditions and, hence, parallels the third type of structural property considered: sensitivity properties. The corresponding functional properties are sensibility properties, that is, functionally distinct responses to stimulation. In the case of sensitivity we were concerned with the variations in stimulation necessary to produce certain changes in structural properties. Here we are concerned with the variations in stimulation required to produce certain changes in functional properties. The stimuli may themselves be structural or functional in nature. Sensibility to noise, for example, would involve structurally defined stimuli; whereas sensibility to aggressiveness would involve functionally defined stimuli (i. e., the measurement of aggressiveness is made under conditions where other people's behavior is defined functionally).

The definitional form of this type of property, then, is similar to the second except that it is necessary to specify the stimulus and the operations by which it must be "administered."

Individuals and Objects

A definition of "physical individual" was provided in 2. 5. Now it is possible to generalize that definition.

When we say something is an object or a course of action for a person, we mean it acts as a unit for him, that its properties cohere and act on him as a whole. We don't separate the weight of a table from the table, since we think it is an essential property of the table. We can separate the package lying on top of the table from it, but not its legs. That is, we recognize that the function which the table serves for us always requires its having legs and weight, but not a package lying on it. The table is for us, then, a collection of properties essential to perform a certain job; the table is an instrument we incorporate into a specific type of our purposive activity.

4.25. Individual. A set of properties (p_1, p_2,..., p_n) to which a subject (A) responds in a choice environment (S) is an individual (I) to A if

(1) that set of properties is virtually certain to produce a response (R) by A in S,

(2) if the removal of any one of the set of properties reduces the probability of R by A in S to virtually zero, and

(3) there is no other set of properties which satisfy conditions (1) and (2).

In effect, a set of properties is said to be an individual if we can find an environment and a response (functionally defined) such that the collection of properties has a unique characteristic: it is virtually necessary and sufficient (practically certain) to produce a specified response in that environment.

When a subject treats a collection of properties as an individual (i), he individualizes. For example, suppose we want to determine whether or not the mass, color, shape and texture of a certain piece of paper constitute an individual (i) for some person. We seek an environment in which the presence of these properties is virtually certain to produce an R, say writing, and in which the removal of any one of these properties is virtually certain not to produce writing, and, further, there are no other properties for the person that satisfy these extreme conditions. Thus, when a person individualizes, he responds to a set of properties collectively; if one of these properties changes, then the function of the subject's response changes.

> 4.26. Essential Properties of an Individual: the properties $(p_1, p_2, \ldots, p_n)$ which are individually necessary and collectively sufficient to produce the response (R) which defines the individual.

Not all properties of an individual (I) are essential; that is, I is not merely composed of essential properties. It may have nonessential properties as well. For example, visibility may be an essential property of paper, but now suppose the paper were red instead of white. The defining response R (writing) might change, but not functionally. That is, a different colored ink might be used by the person, but writing might still occur. In this case, "redness" and "whiteness" produce structural (not functional) changes in R, and are properties of I, though not essential ones. For some objects, of course, whiteness or redness may be essential properties. For example, whiteness is an essential property of a flag of truce, and redness an essential property of a danger flag.

A property of an individual which produces nonfunctional changes in the defining response, is a nonessential property of that individual. But these are not the only kind of essential properties of an individual. A property of an individual which may produce a functional change in the defining response R, but does not do so invariably, is also a nonessential property. For example, city sidewalks are ordinarily lightly colored. That the lightness of their color is nonessential is clear, for we would not expect its absence (i. e., a darker color) to assure the non-use of the sidewalk. Nor would we expect its presence to add to the probability of the occurrence of walking.

What constitutes an individual for a person may change in different environments, and different things may constitute individuals for different persons in the same environment. When packing books for shipment, for example, each book is an individual. For the student reading an assignment each page or paragraph may be an individual. For a type setter each letter is an individual. There is nothing absolute about individuality. It is a functional, not a structural, property that lies in the observer as much as in the observed.

Sets and Classes. Collections of individuals may themselves be individuals. For example, an individual may conceive of his library as an (individual) entity, as well as each book in it. Consider a collection of individuals which is an entity to an observer, each element of which is essential. That is, if any element is removed the response to the collection changes functionally. Such a collection constitutes a set. Thus a pair of shoes, a matched pen and pencil and the volumes of an encyclopedia constitute sets..

4.27. Set of Individuals: a collection of individuals that is itself an individual, the inclusion of each member of which is essential.

4.28. Class of Individuals: a collection of individuals each of which may be replaced by any other of the collection in an environment without affecting the subject's response to the one substituted for,

Hence, in the subject's purposeful state the members of a class have the same (relevant) properties for him. Thus a set of objects may constitute a class to an individual in one state but not in another. For one purpose, for example, any volume of a set of books may be as good as another (e.g., to exhibit the format); for another purpose, the content of each volume may be relevant (e.g., for reference).

Classes, therefore, are collections of individuals to each of which an individual responds (e.g., observes) in the same way.

We can, of course, have collections of collections; for example, a set of libraries. There are also classes of classes; for the various species of animals are classes that, for most of us, are included in the class of animals. The class of animals is, in turn, included in the class of living things.

4.29. Object: an individual (I) is an object to a subject (A) if (1) A perceives I, and (2) its essential properties persist over a period of time.

The essential properties are those which are each necessary and are collectively sufficient for producing the defining response, R. A chair, for example, (1) can be sat on by only one person and (2) has a back. It may or may not have arms, may or may not have four legs, and may or may not be mobile. Note that although its essential properties are functional it can nevertheless be perceived; that is, its structural properties may also be responded to.

Individuation and Identification. As noted above, a subject individualizes when he treats a collection of properties as an individual. He identifies two individuals that are observed at different times if, roughly put, he responds to both in the same way. He individuates or differentiates between two individuals observed at the same time, if, again roughly put, he responds to them differently. The processes of identification and individuation warrant closer examination.

Two individuals alike in "all" respects to a subject, in the same environment at the same time, can be differentiated by their location; that is, their relative positions. If it is necessary for the subject to tell "which is which" at a later time he may endow one with a property that he can later respond to, such as an "identifying mark" or name, or, if they are immobile, he may identify them by their location only. An individual may change over the period of time between the subject's exposures to it; for example, a tomato may change its color. The subject takes it to be the same tomato if color is not an essential property to him; if color is essential, the tomato has become something else; for example, a seed becomes a plant. The seed and the plant are not identified but the seed is identified as a producer of the plant.

Under normal circumstances individuals in the same class are individuated by their nonessential properties, of which location may be only one. It is, however, the most general differentiating property of individuals at a moment of time.

When we identify a person whom we meet today with a child we knew many years ago, despite the lack of any intervening meeting, we may sense a similarity of appearance. If not, it requires communication to establish this individual as the one experienced many years ago. The communication may involve revelation of a name, common experiences, common associations, or some such thing. Identification, therefore, may be based on functional properties as well as structural as, for example, "we went to school together."

The process of identification and individuation are illuminated by the following common situation. You have left your car on a parking lot and return a while later to obtain it. You do not remember its exact location which, of course, would be sufficient under normal circumstances to identify your car. You look about and think you see your car. You try to unlock the door with your key. (This response to that car reveals your identification of it with yours.) The key does not work. You examine the key and find it the right one. Then you examine the car more closely and observe it lacks a sticker on the windshield which yours has. Now you have individuated this car from yours, and resume your search. All of this would have been apparent to an observer of your behavior.

If a subject responds to a stimulus (s_2) at a time t_2 in a choice situation (S) in the same way as he would to another stimulus (s_1) at time t_2 in S, then he identifies s_2 with s_1, assuming he has responded to s_1 previously. If he is aware of the presence of s_1 and s_2 in the same environment and is indifferent to which one he uses, then he places s_1 and s_2 in the same class but individuates them, and such individuation is a response to one or more of their non-essential properties.

4.30. Event: a change in a property of an individual.

An event is something which happens to one or more individuals. That which happens can always be described in terms of changes of properties of the individual(s). For example, the meeting of two people can be defined by changes in location and awareness of each other. An object can be said to be dissolving when its particles change their form and location, and so on. When the changes occur to nonessential properties of an object, the object is said to change; when they occur to essential properties the object ceases to exist.

> 4.31. Relation between Individuals: a property of the set of individuals which the individuals taken separately do not have.

For example, if John and Mary are married, then "married" is a property of the pair. Therefore, if they are divorced, an event has occurred since a property of the pair has changed. Marriage is not usually taken as an essential property of the individuals involved, but it is usually so considered for the pair.

The Form of Perceptions and Observations

As we observed in the last section, the form of observations is reflected in the form of messages about them. Such messages contain statements and these in turn contain expressions. Therefore, we examine the form of both statements and expressions which deal with observations. The scheme we will use is the following:

(1) Form of Statements

- (a) Predication--Classification
- (b) Comparative
- (c) Functional

(2) Form of Expressions

- (a) Qualitative
- (b) Quantitative

Form of Statements

A statement may be represented abstractly as

$$F(x_1, x_2, \ldots, x_n),$$

where $x_1, x_2, \ldots, x_n$ represents the things observed and F represents a relationship among them. The things observed are referred to as arguments, F is referred to as the predicate, and n is the degree of the predicate. For $n = 1$ (i.e., a predicate of degree 1), we have a predicational type of statement. For example, the statement

Charles is a male

has the form

$$F(x),$$

where x denotes the subject "Charles," and F denotes the (monadic) predicate "is a male."

For $n>1$, we have a relational statement. For example,

New York is east of Chicago

has the form

$$F(x_1, x_2),$$

where x_1 and x_2 denote "New York" and "Chicago" and F denotes the predicate "is east of." An example of a statement containing a triadic predicate (i.e., a predicate of degree 3) is

Chicago lies between New York and Denver.

which has the form

$$F(x_1, x_2, x_3).$$

It should be noted that the statement

Charles and Tom are males

may be intended as an abbreviation of

Charles is a male and Tom is a male,

which has the form

$$F(x_1) \text{ and } F(x_2)$$

rather than

$$F(x_1, x_2).$$

Predication and Classification. As indicated above, a simple predicational type of statement is one which has the form $F(x)$; for example,

Charles is a male.

Such a statement attributes a property to an object, event, or some combination of these.

A <u>compound</u> predicational statement combines two or more simple ones. For example,

Charles is a male $[F_1(x)]$

and

Charles is an adult $[F_2(x)]$

can be combined into

Charles is an adult male.

This statement can be represented by "$F_1(x)$ and $F_2(x)$." Similarly, the statement

Charles and Tom are adult males

combines two compound predicational statements and can be represented by "$F_1(x_1)$, $F_1(x_2)$, $F_2(x_1)$, and $F_2(x)_2$." This symbolism makes explicit the fact that confirmation of the statement requires four attributions.

In order to confirm simple predicational statements, it is necessary to (a) identify the subject and (b) define the attributed property. Identification, as we have already seen, involves specifying a set of properties which are sufficient to differentiate the subject from any other possible subjects. Hence, identification involves a compound predicational statement, $[F_1(x), F_2(x), \cdots, F_m(x)]$, where $F_1, F_2, \cdots, F_m$ are sufficient to identify x.

It will be noted that the statement

Charles is a male

is equivalent to

Charles is a member of the set of males.

That is, every predicational statement classifies its subject. Therefore, corresponding to each (monadic) predicate (F) defined over a set (X), there is a subset of X consisting of all those members of X having the predicate F. A simple predicate applied to a set, then, creates two classes. If there are m predicates, 2^m classes can be constructed.

Relations and Comparisons. As already indicated, a statement with a predicate of degree greater than 1 is called a relational statement. In $F(x_1, x_2)$ a property is attributed to x_1 and x_2 taken collectively. For example, in the statement

Charles is the brother of Horace,

"is a brother of," the predicate, cannot be attributed to either subject taken separately, as "are male" can. It will be noted that in this statement we can revise the order of the subjects, Charles (x_1) and Horace (x_2); that is,

$F(x_1, x_2)$ implies $F(x_2, x_1)$.

Where the predicate holds for every pair of subjects in a set, the relation is said to be symmetric over the set. Such a relation does not order the subjects, but a relation which is not symmetric may; for example,

Charles is younger than Horace.

Here $F(x_1, x_2)$ does not imply $F(x_2, x_1)$. Charles and Horace are said to be an ordered pair.

For a relation to order more than two subjects it must be transitive, in addition to not being symmetric. A (dyadic) predicate is said to be transitive if and only if, for any triplet of arguments, x, y, and z, $F(x, y)$ and $F(y, z)$ together imply $F(x, z)$. A comparative statement is any statement the principal predicate of which is an ordering relation. For example, the predicate "is less than" defined over the set of real numbers provides an ordering of the real numbers.

Ordering relations are of two types, quasi and strict, depending upon whether the relation is reflexive or irreflexive. A (dyadic) relation F defined over a set X is said to be reflexive if and only if F (x, x) is true for every x in X. It is said to be irreflexive if and only if F (x, x) is false for every x in X.

Examples of quasi-ordering relations are "less than or equal to" over the set of real numbers, "is at least as tall as" over the set of human beings, and "implies" over the set of statements. Examples of strict ordering relations are "is less than" over the set of real numbers, "is the ancestor of" over the set of human beings, and "is a proper subset of" over the set of sets.

There are many different types of ordering relations, some of which are discussed in detail by Ackoff (1962, Chapter 6).

Functions. A particularly important class of relational statements consists of ones involving a functional relation. In a statement of the form $F(x_1, x_2 \cdots, x_n)$, where $n > 1$, if when F and all but one of the x's are specified, the value of the remaining x is completely determined, then F is a strong functional relation. For example, consider the (dyadic) statement

Gloria is the spouse of Charles,

which can be represented as $F(x_1, x_2)$. Once F is specified as "is the spouse of" and either x_1 or x_2 is specified (Gloria or Charles), then the value of the other is completely determined. This statement may be rewritten as either

$$x_1 = f_1 (x_2)$$

or

$$x_2 = f_2 (x_1).$$

Consider the triadic predicate F defined over the real numbers such that $F(x_1, x_2, x_3)$ means "x_1 is the sum of x_2 and x_3." Such a predicate yields a function for all its arguments, and we may write

$$x_1 = f_1(x_2, x_3),$$
$$x_2 = f_2(x_1, x_3),$$
$$x_3 = f_3(x_2, x_3).$$

In this case,

$$f_1(x_2, x_3) = x_2 + x_3,$$
$$f_2(x_1, x_3) = x_3 - x_1,$$
$$f_3(x_1, x_2) = x_2 - x_1.$$

Note the important property of statements involving <u>strong</u> functional relations: if the value of any (independent) argument inside the functional bracket is changed, the value of the (dependent) argument on the left side of the equation must be changed.

Now let us consider a <u>weak</u> functional relation; for example, the dyadic predicate "is the father of" in the domain of human beings. $F(x_1, x_2)$ means "x_1 is the father of x_2." For any given value of x_2, there is only one value of x_1 such that $F(x_1, x_2)$ is true. In this case, however, specifying x_1 does not determine x_2, since x_1 may be the father of several persons. In general, a predicate is a weak functional relation for its $\underline{k}^{th}$ argument if and only if, (a) when the values of all arguments except the $\underline{k}^{th}$ are fixed, precisely one value for the $\underline{k}^{th}$ argument is determined, and (b) a change in an x other than x_k may not necessitate a change in x_k. For example, in the statement

F. D. R. was the father of James Roosevelt

if "F. D. R." is changed, "James Roosevelt" must be also; but, if

"James Roosevelt" is changed, "F. D. R." need not be (if one of his other offspring is substituted for James). In the earlier example in which F denotes "is the spouse of," both x_1 and x_2 were sufficient to completely determine the other. In this example, x_2 is sufficient (relative to the predicate "was the father of") to determine x_1, but x_1 is not sufficient to determine x_2. However x_1 is sufficient to specify a class of subjects any one of which when substituted for x_2 makes the statement true; therefore, x_1 bounds the values of x_2.

When we examine the type of statements which take the form

$$x_1 = f(x_2, x_3, \cdots)$$

we observe three different types which are characterized by the property of the function. Consider first the familiar law of freely falling bodies.

$$s = \frac{1}{2} gt^2,$$

in which s is the distance traveled, g is the gravitational constant, and t is the time from release. We note that (for nonnegative s, g, and t)

$$s = f_1 (g, t), \text{ where } f_1 (g, t) = \frac{1}{2} gt^2$$

$$g = f_2 (s, t), \text{ where } f_2 (s, t) = 2s/t^2$$
$$t = f_3 (s, g), \text{ where } f_3 (s, g) = \sqrt{2s/g}.$$

Clearly, the functional relation involved in this law is strong, since the value of each argument is completely determined by the other two.

Now consider a statement of the form

$$x_1 = f(x_2, x_3, \cdots, x_k),$$

where $x_2, x_3 \cdots, x_k$ is a subset of a set of arguments which is sufficient to completely determine the value of x_1. The subset, then, only partially determines (i. e., bounds) the value of x_1. For example, suppose that in fact (1) $x_1 = x_2 + x_3$, (2) x_2 and x_3 are independent, and (3) x_3 can assume three different values: -1, 0, and 1. Suppose further that we do not know about x_3 but we do know that the value of x_1 depends on the value of x_2 and something else. Then, from observation we could determine that either

$$\text{(a)}\ x_1 = x_2 - 1,$$

$$\text{(b)}\ x_1 = x_2,$$

or

$$\text{(c)}\ x_1 = x_2 + 1.$$

Suppose also that the probabilities of observing each were p(a) = 0.25, p(b) = 0.25, and p(c) = 0.50. We could now compute $E(x_1)$, the expected value of x_1:

$$\begin{aligned} E(x_1) &= 0.25\,(x_2 - 1) + 0.25(x_2) + 0.50\,(x_2 + 1) \\ &= 0.25x_2 - 0.25 + 0.25x_2 + 0.50x_2 + 0.50 \\ &= x_2 + 0.25. \end{aligned}$$

Now, although the expected value of x_1, $E(x_1)$, is completely determined, the value of x_1 is not. We know that a change in x_2 is <u>not</u> sufficient to result in a change in x_1, since a change in x_3 may compensate for it. But we do know that knowledge of the value of x_2 is necessary for determining the value of x_1. Then x_2 is not a deterministic cause of x_1, but (as we have already considered in Chapter 1) it is a <u>probabilistic</u> cause or <u>producer</u> of x_1.

Suppose that we do not know whether the value of x_1 depends on the value of x_2; that is, we know of no necessary connection between x_1 and x_2, but we have observed that x_1 tends to increase as x_2 does. Once again we may express x_1 as a function of x_2, but this is a <u>pseudo</u> function, since x_2 is not sufficient for, and we do not know that it is necessary for, determining the value of x_1. We cannot say that x_2 is either the cause or the producer of x_1, but we may be able to say that they are <u>correlated.</u>

Consider, for example, a person who usually brushes his teeth once a day, just before going to sleep at night. Brushing his teeth is neither necessary nor sufficient for his going to sleep and hence is neither the cause nor the producer of his retiring for the night. And yet the two events usually occur together. To take another example, in one large city it was discovered that people who live in neighborhoods in which there is a heavy soot-fall are more likely to get tuberculosis than people who live in neighborhoods with less soot-fall. Yet medical research has shown that soot-fall is neither necessary nor sufficient for the occurrence of tuberculosis. Hence, the values of two variables may tend to change together, and yet the variables may not be causally connected. Such variables are said to be correlated.

The knowledge that two things tend or do not tend to change together can, nevertheless, be very useful. For example, when we see the person in the above illustration brush his teeth at night, we can predict with some assurance that he is about to retire. That is, we can use our knowledge of the value of one variable to predict the value of another.

Form of Expressions in Statements: Quality and Quantity

Compare the following two statements:

John is heavy

and

John weighs 150 pounds.

Both appear to be simple predicational statements of the form $F(x_1)$, where x_1 denotes "is heavy" in the former and "weighs 150 pounds" in the latter. The obvious difference between these two statements is that the second contains a number. What is not so obvious is that, because the second statement contains a number in what appears to be its predicate, it should be represented as a functional statement of the form $F(x_1, x_2)$, where F denotes "is equal to," x_1 denotes "John's weight," and x_2 denotes "150 pounds." This is a weak function, since specification of F and x_1 completely determines x_2, but F and x_2 do not determine x_1.

A transformation similar to changing

John weighs 150 pounds

into

John's weight is equal to 150 pounds

cannot be performed on

John is heavy.

We can transform this statement into

John's weight is greater than W pounds

or

John's weight is greater than W_1 pounds and less than W_2 pounds.

There is, however, no reasonable transformation of "John is heavy" into a statement containing the relationship of strict equality.

Not all statements which contain numbers are quantitative statements. Numbers may be used in statements for a variety of purposes:

(1) To identify (or name) the subject; for example, This is a prisoner number 59241.

(2) To identify the class in which the subject is placed; for example, He was in the graduating class of 1951.

(3) To identify the number of subjects in a class; for example, Twenty three universities offer courses in this subject.

(4) To identify the rank order of a subject in a class; for example,
General Motors is the largest manufacturer of automobiles.

(5) To identify the number of units on a scale which corresponds to the subject's property; for example,
John is six feet and one inches tall.

Only the last three of these represents what is called "measurement."

Measurement. As we shall see later (in Chapter 6), to think about something is to manipulate a representation of that thing. As we shall also see later (in Chapter 9), such representations are called signs. If a sign that represents what is thought about has some of the same relevant properties as that which it represents--for example, it looks like what it represents--the thought process is usually facilitated. It is possible to go even further to facilitate thought. Man has created systems of signs (e.g., letters and numbers) between whose elements he has established certain relationships (e.g., an order). When such signs are used to represent things which are related to each other in some of the same ways that the signs are taken to be, measurement has taken place.

4.32. Measurement. the use of man-made signs (see 9.1) to represent things which are believed to be related to each other in some of the same ways that the user believes the signs to be related.

This definition has made use of the concept, "belief," which is considered in Chapter 5.

There are four major types of measurement, each employing a different type of scale: nominal, ordinal, interval, and ratio.

1. Nominal Scale. When an individual uses the same sign to represent each member of a class and different signs for members of different classes, then he employs the simplest form of measurement by using a nominal scale. For example, if all males are signified by an "M" and all females by an "F," a two-valued nominal scale is used. The only properties of the signs which are employed here are identity and difference. Numbers or words can also be used to represent class membership (e.g., "male" and "female" or "1" and "2").

Note that use of a nominal scale produces predicational statements (e.g., "X is a male.").

2. Ordinal Scale. Objects can be ordered, ranked, or compared with respect to some relationships that hold between them; for example, they can be ordered with respect to the relationship "is larger than." If $\underline{n}$ objects are so ordered they may be numbered from 1 to $\underline{n}$ in such a way that the order of the numbers and the order of the objects represented by them are the same. To do so is to employ an ordinal scale.

There are a number of different kinds of ordering. The differences between them derive from the properties which the observer believes the ordering relationship has over the set of things observed (the "reference" set). Four properties of relationships are relevant here: reflexivity, symmetry, transitivity, and connectedness. Each of these, and variations thereof, can be defined using the concept "belief." For example, a relationship ($\underline{R}$) is believed to be reflexive relative to a set of entities ($\underline{X}$) if for every member of $\underline{X}$ ($\underline{x}_i$) an individual believes ($\underline{x}_i$ R $\underline{x}_i$). Using definitions of these relationships it is possible to define various types of ordinal scales including the principal ones: partial, weak, and complete. *

3. Interval Scale. Signs can be used to represent the magnitude of differences between elements on an arbitrary scale (i. e., using an arbitrary unit). For example, knowing that a column of mercury rises with temperature we can mark equal distances of any magnitude on such a column and number them consecutively from some arbitrary starting point. This was done to form the Farenheit and Centigrade scales of temperature. These are called interval scales.

If at three successive times we observe 32° F, 64° F, and 128° F we can say the differences between the successive readings are equal and that the difference between the first and last is twice as large as between the first and second. We cannot say, however, that 64° F is twice as hot as 32° F. This is apparent if we were to use a Centigrade scale for the same three observations. In this case we would obtain 0° C, 17.8° C, and 35.6° C. The relative sizes of the intervals do not change, but the relative sizes of the readings themselves do. This characteristic derives from the arbitrary character of the "zero-point" (i. e., where we start to number the units on the scale) and of the units themselves. The units employed do not have to be distances; for example, they can be logarithms of distances.

Hence, when an individual uses arbitrarily numbered constant units on one scale to represent changes in the property of something else, he employs an interval scale. He may use such a scale without being aware of its properties and hence draw inferences from the signs employed (numbers) which are not justified (e.g., 64° F = 2 x 32° F).

---

*For a complete discussion of these and other aspects of measurement, see Ackoff (1962, Chapter6).

Among some of the common properties which we measure on interval scales are position on the earth's surface (using arbitrary longitude and latitude), and time on the calendar (January 1st and lengths of months are arbitrary units).

4. Ratio Scale. Note that in interval measurement, units of the property involved are not measured directly. In measuring length or weight, however, units of these properties are used. The units employed are still arbitrary (inches, feet, centimeters, meters, and so on) but the starting point (the "zero point") is not. The zero-point is natural. Units of this type numbered non-arbitrarily yield a ratio scale. When such a scale is used we can say that 64 units (e.g., inches) is twice as long as 32 units (e.g., that 5'4" is twice as long as 2'8"). Most arithmetical operations are applicable to the numbers obtained from use of such a scale and hence such measurements have the greatest inferential potential. Each arithmetical operation has a physical counterpart; for example, we can add, subtract, multiply, and divide distances as well as the numbers which represent them. On the other hand, we cannot add temperatures; two liquids each at 70° F when added to each other do not yield a liquid with a temperature of 140° F.

Any property which can be quantified can also be treated qualitatively. A quality can be thought of as a range along a scale (i.e., a morphological interval) in terms of which the property can be measured. For example, a person can be said to be "tall" if he is over 5 feet 10 inches, "medium" if he is between 5 feet 6 inches and 5 feet 10 inches, and "short" if he is under 5 feet 6 inches.

It is also true that any qualified property is potentially capable of being expressed quantitatively in terms of such a range along a scale. We may never be able to translate all qualities into such measures, but, as science progresses, it converts more and more qualities into equivalent quantitative expressions. But this is not a one-sided development. As science develops more quantitative measures, it also requires new kinds of qualitative judgments. For example, height can be measured as a vertical distance, but to do so requires our ability to determine verticality. We can convert verticality into a measure of the angle between a straight line and a radius projected from the earth's center of gravity. This requires our ability to determine straightness, and so on. Quantification at any stage depends on qualification. What is qualified at one stage may be quantified at another, but at any stage some qualitative judgments are required. Consequently, improvement of observations not only is a function of an increased capacity to quantify efficiently (i.e., to measure) but also depends on an increased capacity to qualify efficiently.

ERRORS OF OBSERVATION

There are four possible sources of error in observation: (1) the observer himself, (2) the observed, (3) the instruments used in making observations, and (4) the environment in which the observations are made. Furthermore, there are three possible types of error that can be produced by these sources: (a) observing inaccurately (e.g., miscounting or mismeasuring), (b) not seeing something that is there, and (c) seeing something that isn't there. Because of these errors we consider some people to be better observers than others and a number of tests have been developed for evaluation of observers.

Kirk and Talbot (1966) have named these three types of observational error as (a) <u>systematic</u> or <u>stretch</u> distortion; (b) <u>fog</u> distortion, and (c) <u>mirage</u>. Each of these types of error can be produced by any of the four sources of error. (See Table 4.1.)

TABLE 4.1. Sources and Types of Errors of Observation

	Types of Error		
Sources of Error	Systematic	Fog	Mirage
1. Observer			
2. Observed			
3. Instruments			
4. Environment			

Systematic Distortion

> In SD [systematic distortion] no information is lost. Rather, it is changed or recorded in an orderly or systematic way. Distortions of this kind are like the distortions a rubber sheet might undergo, so long as it is not torn. Thus, SD can be eliminated or "corrected for" by the application of a rule specifying the appropriate "topological transformation" (p. 310).

Kirk and Tablot cite the following example of systematic distortion produced by an observer:

> Astronomer Maskelyne fired his assistant, Kinnebrooke, because the latter was clearly incompetent. Charged with clocking upper transits of certain reference stars, Kinnebrooke consistently clocked them "late" (p. 308).

They illustrate instrument-produced systematic distortion as follows:

> Some auto rear-view mirrors are cylindrically convex so that a driver may scan at a glance far more than a "flat-mirror glimpse" of the territory behind him. Again, he sees images which are tall and thin, and they require "getting used to."

A bathroom scale that is improperly set will also produce a systematic bias into readings of persons' weight.

An example of observed-produced systematic distortion is found in a subject being interviewed who always, or almost always lies. If he always lied we could easily correct for this distortion, by attaching a "not" to his main verbs.

Environment-produced systematic distortion is introduced, for example, by a non-white light when we are trying to determine the color of objects. Changing temperatures will also change the length of metal bars and hence may produce distorted observations. These could be corrected if we know the temperature and the coefficients of linear expansion of the metals under observatinn.

Fog. This occurs when an observer does not see what is there.

In such distortion ."...information is lost, mashed out, 'fogged' over..." (p. 313). For example, an observer may not be able to hear sounds above an abnormally low frequency or volume (if he is partially deaf). If he is color blind then, of course, he fails to observe color.

Recording equipment may also fail to pick up low-volume sounds or high frequencies. Film may fail to capture color. (If they distort color, it is systematic distortion, not fog).

Noise in an environment may result in our failure to hear certain sounds. Glare may prevent our seeing objects that would otherwise be clearly visible.

A subject in an interview who lapses into a language or use of words that we do not understand introduces fog into the exchange. Ambiguity is a type of fog. For example, some feel that James Joyce produced an inpenetrable verbal fog in Finnegan's Wake.

Mirage.

> In mirage distortion (MD) we see something that "isn't there." Far from withholding information from us, MD gives us extra, unwanted information (p. 316).

Most of us have seen or heard things that weren't there or tasted ingredients in food that were not there. A subject in an interview can deliberately (or not) produce a belief in us of the occurrence of an event that never took place. A burglar-alarm system may "go-off" because of an internal defect when no intruder is present. A false alarm is a mirage. In a very noisy environment we may hear things that were not said.

Hence, there are four sources and three types of observational error. Implicitly or explicitly each observer has relevant beliefs with respect to each and these determine whether or not use will be made of the data obtained. When the observer believes that error is present he may be able to correct for it if he knows its source and nature. For example, he can correct for the bias of the bathroom scale or the late response of another observer. By interpolation he can fill in missing data and by a wide variety of tests he can eliminate inconsistent data. The theory of data adjustment is frequently used in science for just this purpose. (See Deming, 1943.)

LOCUS OF OBSERVATIONS

In order to determine whether an object has a certain property it must be observed in some environment. Individuals are aware that some environments are not suited for observing certain properties. For example, most of us would not try to determine the color of an object in an environment illuminated by red light. Therefore, for each property to be determined there is an ideal environment in which relevant observations should be made. It is seldom possible, however, to make observations in the "perfect" environment. Therefore, an observer must frequently settle for something less than ideal or construct an environment which closely matches his requirements. In either of these cases the environment may deviate from his "idealized" one and hence he may have to adjust his observations to account for these differences.

A laboratory is the epitome of an environment that is deliberately constructed to facilitate making the type of observations which are desired. Within it relevant variables are controlled so that their effect on what is observed is held constant or can be determined. Even in a natural environment certain variables can be controlled, but usually not all the relevant ones. Hence, the difference between a laboratory and the real world is a matter of degree, with many gradations of control between them.

When an individual seeks to establish a cause-effect or producer-product relationship between two or more things and in so-doing controls the values of other variables that he believes may affect the outcome, he conducts an experiment. Therefore, an experiment is experience under controlled conditions.

4.33. Control of a Property. An individual controls a property of an object, an event, or their environment if either he produces an intended value of that property or he selects an environment in which the property has the intended value.

4.34. Experiment. An individual conducts an experiment if he controls changes or differences of values in one set of properties and observes the values of another set of properties, with the intention of determining whether or how changes of one or more properties of the first set produce or cause changes in one or more properties of the second set.

For example, an experimenter operating in a laboratory may hold temperature constant, change air pressure, and observe the boiling point of water. On the other hand he may select environments which have the same temperature but different air pressures, and observe boiling points in each.

An experiment which is conducted on a representation of the thing being studied, not the thing itself, is a simulation. Simulation is vicarious experimentation in which a representation stands in as a proxy for the thing which it represents. For example, a wind tunnel or tow tank (which represent or model specific environments) may be used to conduct experiments on model aeroplanes or ships. The model may also consist of man-made signs or symbols such as will be discussed in Chapter 5 . Simulation using symbolic models has become commonplace with easy access to electronic digital computers, but they can also be carried out by hand.

A detailed discussion of experimentation and simulation can be found in Ackoff (1962, Chapters 10 and 11).

Now we turn to a consideration of how problems are formulated and models that can be used in solving them are constructed.

BIBLIOGRAPHY

Ackoff, R. L., Scientific Method: Optimizing Applied Research Decisions. John Wiley and Sons, Inc., New York, 1962.

Allport, F. H., Theories of Perception and the Concept of Structure. John Wiley and Sons, Inc., New York, 1955.

Beardslee, D. C., and Wertheimer, M., eds., Readings in Perception. D. Van Nostrand Co., Inc., Princeton, 1958.

Bruner, J. S., "On Perceptual Readiness" (1957) in Beardslee and Wertheimer.

Cantril, H. H., The "Why" of Man's Experience. MacMillan and Co., New York, 1950.

Deming, W. E., Statistical Adjustment of Data, John Wiley and Sons, Inc., New York, 1953.

Freud, S., New Introductory Lectures on Psychoanalysis. W. W. Norton Co., New York, 1933.

Goddard, H. H.

Guthrie, E. R., The Psychology of Human Conduct. Harper and Bros., Inc., New York, 1938.

Hirsh, I. J., "Sensation and Measurement," in Beardslee and Wertheimer, p. 45-61.

Ittelson, W. H., The Ames Demonstrations in Perception. Princeton University Press, Princeton, 1952.

Jung, C. G., Psychological Types. translation by H. G. Baynes, Harcourt Brace and Co., Inc., New York, 1924.

Kilpatrick, F. P., Explorations in Transactional Psychology. New York University Press, New York, 1961.

Kirk, J. R., and G. D. Talbot, "The Distortion of Information, " in Communication and Culture, ed. by A. G. Smith, Holt, Rinehart and Winston, New York, 1966, pp. 308-321.

Koffka, K., "On the Structure of the Unconscious" in The Unconscious, a Symposium ed. by E. Dummer, New York, p. 43-68, 1927.

Krikorian, Y. H., "An Empirical Definition of Consciousness, " Journal of Philosophy, XXXV, No. 6, March 17, 1938, pp. 156-161.

Miller, J. G., Unconsciousness, John Wiley and Sons, Inc., New York, 1942.

Roback, A. A., Self-Consciousness and Its Treatment, Sci-Art Publishers, Cambridge, 1933.

Singer, E. A., Jr., Mind As Behavior, R. G. Adams and Co., Columbus, 1924.

______, "On the Conscious Mind", Journal of Philosophy, XXVI, 21, October 10, 1929, p. 561-75.

Chapter 5

PROBLEM FORMULATION AND MODEL CONSTRUCTION: MEMORY AND BELIEFS

RECOLLECT, _v._ To recall with additions something not previously known (Ambrose Bierce, _The Devil's Dictionary_).

INTRODUCTION

An individual's conception of a problem is a product of what he perceives in his state and how he feels about it. His feelings will be discussed in Chapter 7. Here I consider the source and nature of the components of the individual's conception of the problem situation.

What an individual observes in a situation is not merely a matter of what is "given" to him by the situation because much more is "offered" than he can possibly receive. Therefore, what he observes is also a matter of what he "takes" and what is "forced" on him. He enters each situation with a "set"; the set is his model of the situation which provides him with criteria of relevance and hence influences what he looks for.

This is not to say that an individual observes only what he looks for. Some stimuli, by the sheer force of their impact on his senses, may impose themselves on him regardless of the criteria that he employs. For example, a person who is reading a book and intends to shut out the conversation around him may, nevertheless, hear a message shouted to him or another. In Chapter I, I called such messages "unsolicited." They may, however, be relevant. For example, the message shouted to him may inform him that the lights are about to be turned out. Imposed or accidental perceptions may play an important role in the process of choice.

In this chapter we begin to examine what a person brings into his observations with him. One's present observations and the conclusions drawn from them are always coproduced by one's past experiences. Past experience, organized in various ways, comes forth from one's memory in the form of beliefs and attitudes. Beliefs are inferences drawn from past and present perceptions, and attitudes are feelings about what was perceived. Attitudes will be considered in the next chapter. Here I consider memory and beliefs.

My objective is to determine how an observer can determine what another remembers and believes.

MEMORY

Inherent in most efforts to analyze the meaning of memory is the question as to whether or not memory is a conscious function. For many early psychologists like Colvin (1915) the answer was "Yes." Colvin took memory to be "conscious phenomenon" which "signifies the modification of present experience in terms of past experiences" (p. 128). Memory for him was "the revival of a past experience with a definite knowledge that this experience belongs to the past" (p. 130). Habitual (or unconscious) responses to past experiences do not involve memory according to Colvin: "Memory easily lapses into mere habitual responses to familiar objects or events without any conscious recognition." To make this position precise it would be necessary to distinguish clearly between memory and habit. That this has not been satisfactorily done is asserted by Dockeray (1932) as follows:

> Memory is usually distinguished from habit in that the former refers to purely mental processes, and the latter to those forms of motor response that have been learned. Here again the distinction is not always clear (p. 351).

Since the distinction was not always clear, many early psychologists evaded the problem by defining memory in "mentalistic" terms, and habit in "behavioristic" terms. For example, Judd (1907) referred memory to "those cases in which phases of experience are recalled from the past and consciously recognized as so recalled from the past" (p. 237). Obviously to define memory in terms of "recall," "recognition," "retention," and so on, gives us little insight into its meaning. Another such typical pseudo-explanation is to be found in Guilford (1939):

> Most psychologists are convinced that learning produces changes in the brain; that those changes are retained for at least some length of time; and that they express themselves later by making the individual behave differently before learning. *This is the modern story of learning in one sentence* (p. 408, italics mine).

Through the influence of psychoanalytic thought, association of memory and consciousness has been considerably weakened. Within Freudian theory one can talk of something remembered by an individual who is conscious neither of what is remembered, nor of the process of remembering it. But for psychoanalysts, memory was so conceived as to make it more susceptible to clinical analysis than to experimentation. They gave it a very subjective tone. That is, in clinical practice the psychoanalyst takes himself to be capable of judging what is remembered, but the basis of his judgment is not made explicit.

Although there is considerable disagreement to be found in contemporary thinking on memory, one can find a common core of agreement, apparent even in the few representative definitions already quoted. The core of agreement consists in recognition that memory involves some kind of response to past experience. This perhaps obvious basis of agreement is expressed by Miller (1942):

> In its widest sense <u>memory</u> is the name for the influence of a person's past upon his present and future thoughts and behavior (p. 210).

Earlier Koffka (1935) had written essentially the same thing:

> The concept of a memory trace is an attempt at explaining the influence of the past by the condition of the present (p. 429).

Later Young (1961) echoed

> Memory is the sum of what can be remembered, the diary of the mind (p. <u>v</u>).

These definitions and what they agree on are so general and so non-operational as to have little value in science. Some contemporary psychologists have tried to be more specific and identify memory with the ability to <u>store</u> and <u>retrieve</u> past experience. English and English (1958) echoing the previous quote from Colvin, add the observation that memory brings with it recognition that what is retrieved <u>is</u> past:

> Memory: the general function of reviving or reliving past experience, with the more or less definite realization that the present experience is a revival (p. 315).

The storage and retrieval capabilities of the human brain have been under considerable investigation as a consequence of the development of information theory and computer memories. In this connection the following observation by Ashby (1966) is relevant:

> The word "memory" is often used to refer to the power of the <u>reproduction</u> of learned material...This power of reproduction seems to be something of a by-product of the brain's activity; the not very intelligent parrot can do it quite well, and the magnetic tape recorder can do it so much better than the human being...(p. 378).

The tape recorder, if not the parrot, reproduces information in a nonpurposeful way. It does not use what is reproduced in a choice process. Reproduction in this sense is certainly not the essence of memory. A person may be able to remember without being able to reproduce it structurally as a recorder can. Furthermore, the tape recorder does not have the ability of the human to selectively forget or, perhaps, to selectively store in the first place. It is apparent that humans store only a small portion of the information they receive. Perhaps attempts to explain memory would benefit from increased attention to the loss of information: non-recording and forgetting. Miller (1956) has been working in this direction. It might also be fruitful to pursue experimentally the line of inquiry initiated by Freud clinically: study of the inability of the human to recall what has been stored in his memory, at least not consciously.

Recent research efforts have been devoted more to retrieval than to storage. This follows from the kind of conviction expressed by Bruner (1962):

> The principal problem of human memory is not storage but retrieval (p. 94).

This emphasis is reflected in the experimental studies reported by Kelley (1964). Memory, however, involves more than storage and retrieval of past experience; it involves a purposeful response to what is retrieved. A computer that stores and retrieves (and even reproduces) information does not remember it unless it uses the information purposefully. If its operations on the retrieved information does not involve choice, as is the case in most programs, then the computer has not remembered. The use of "memory" in such a context is metaphorical at best, and misleading at worst.

A person who stores a letter in a file and later retrieves it does not necessarily remember its content. If he can reproduce its content without examining it he is said to remember it because he is believed to be capable of using it in a choice process.

From this brief survey of what has been said about memory, and from what I have said about it, the following observations can be made.

1. Initially one might assert that unless something has been perceived, it cannot be remembered. It seems preferable, however, to make a weaker statement: unless something has been reacted to (see definitions 4.4 and 4.5) it cannot be remembered. If one accepts the weaker statement then perception is not necessary; that is, the response to the structural change (reaction) produced by a stimulus--the sensation part of perception--may come later when the reaction is recalled. For example, one may see something at time t_0 but not respond to it until a later time t_1.

2. One may remember a stimulus reacted to and/or the response to it. That is, we may not remember <u>all</u> of a perception. If a response is remembered then, as a minimum, the individual must have reacted to his own response. When one responds to one's earlier response and in addition to the stimulus that produced it, the memory is self-conscious (see definition 4.19). Note, however, that response to a response need not be conscious; as when its stimulus is not responded to.

3. What is stored between an initial reaction and its recall is obviously not its stimulus or the response to it, but a <u>representation</u> of these. Since a representation may not be accurate, recall can be in error. The representation must consist of structural changes in the brain because we know that damage to it can produce loss of memory, partially or completely. In this sense, storage of information in a computer's "memory" is analogous to storage in the human brain.

4. Recall--response to a retrieved representation--does not just happen; it is produced by something in the "recall environment."

If this were not the case our consciousness would be inundated by irrelevant memories.

5. Recall, then, is itself a response to at least one stimulus that operates in the present. Recall of a past reaction ordinarily, but not necessarily, involves recognition that the reaction took place in the past. In hallucinations this may clearly not be the case.

6. Recall is selective: it involves a search for the relevant. Otherwise everything stored would be recalled at once. Hence, recall involves an association of something in the present to something in the past, and this association must be based on believed relevance; that is, on what the individual believes will enable him to make a better choice in the present.

7. An individual obviously cannot remember everything. He may not store something either because he believes it is irrelevant to choices he will have to make in the future, or he believes he can retrieve it from some other source when necessary; for example, when he knows that it is recorded in an accessible place. This implies that committing something to memory is a matter of choice, even if an unconscious choice. To deny this requires either that we assert all things reacted to are committed to memory, or that a selection is made nonfunctionally; that is, in a way that an individual cannot control. For example, it has been argued that only strong (structurally intense) stimuli are remembered. But clearly we can remember whispers and forget shouts. A structural explanation of what is committed to memory seems infeasible. On the other hand, we know that a strongly motivated student remembers what he is taught, but one that is poorly motivated (he believes that what he is being taught is irrelevant or that it is not important to remember it whether relevant or not) forgets.

5.1. <u>Memory</u>. An individual who responds at time t_1 to a stimulus (X) to which he reacted at an earlier time, t_o, remembers X.

This concept of memory is a very general one. Since every stimulus precedes a response to it, all stimulus-response phenomena (including sensation, awareness, and consciousness) can be subsumed under it. Clearly, when the interval from t_o to t_1 is very small, a moment, we do not usually associate memory with it; but it is clear that unless the stimulus is "retained" over even a short interval the response could not follow. In practice we apply memory only to situations in which the individual's exposure to the stimulus X is not continuous over the interval $t_o - t_1$.

The definition of memory formulated here does not require that the memory-response be a conscious one. It may be conscious, but is not necessarily so. For example, we can remember how to climb a set of stairs without being conscious of that act. Most of our habitual behavior displays unconscious memory; we frequently are not conscious of why we do things as we do. I wear my wrist watch face down on my left wrist for reasons of which I am not conscious, but, clearly, the original stimulus is still operating on me as I put on my watch each morning.

5.2. Intensity of a Memory: the intensity of the response that defines it.

One may also speak of the durability of a memory as the length of time over which it persists.

5.3. Correctness of a Memory: the efficiency of the memory response for the objective for which it is intended.

For example, a student taking an examination on material that he has read, remembers correctly, if he desires a high grade and his responses are efficient in producing one.

What is remembered--the content of memory--is representable by statements in the same way that observations are. Hence the discussion in Chapter 4 of the form of statements is also applicable here. Memories, it should be noted, are communications to oneself.

BELIEF

The relevance of an individual's beliefs to his model of a choice situation becomes apparent when we reexamine the components of such a model:

(1) A set of courses of action from which he <u>believes</u> he can make a choice.

(2) A set of outcomes which he <u>believes</u> to be producible by the courses of action believed to be available.

(3) One or a number of states (i. e., sets of uncontrolled conditions which can affect the outcome of a course of action) he <u>believes</u> to have some probability of being the true state of affairs.

(4) The efficiency that each course of action believed to be available is <u>believed</u> to have for each outcome that is believed to be producible.

(5) The utility which each outcome believed to be producible is <u>believed</u> to have.

(6) The probability that each state believed to be possible, is believed to have.

In short, an individual's model of a choice situation consists of what he believes and observes to be relevant to his decision.

There has been anything but universal agreement as to <u>precisely</u> what constitutes belief. The historic confusion has led some, like

Bailey (1854), to conclude that belief is an affection "of the mind on which definition can throw no light, but which no one can be at a loss to understand" (p. 1). Although there has been considerable thought given to belief (almost all of it in speculative philosophy rather than in experimental psychology), little advance over Hume's classical treatment (1748) has been made. Prominent experimental work on the concept was done by Lund (1925-26) who felt obliged to foresake defining belief until his experiments were completed, lest he bog down in theoretical difficulties. His definitional effort, made after the completion of the experimental work, led him to conclude that no clear distinction can be made between knowledge and belief.

The confusion is apparent enough in the literature. For example, James (1890) hinted at a behavioristic definition of belief (p. XXI). For Tolson (1941) belief was not necessarily behavioristically defined for it is either a "thought or statement regarded to be true by the person who holds it" (p. 9). For Gurnee (1936) belief was only one kind of behavior: the verbalization of an attitude (p. 250). But none say what kind of behavior, verbal or otherwise, is definitive. What does it mean to say an individual "holds a belief," or "regards a thought or statement as true?" Laird (1930) stated that "Knowledge occurs when a conviction is fully evidenced (or certified in a logical sense) and that mere belief occurs when a conviction is not fully evidenced" (p. 157). What is gained by defining knowledge and belief in terms of "conviction," where "conviction" itself it left undefined? Are we to suppose that the meaning of conviction is somehow better known, and, if so, by what criteria?

The various aspects of belief are considered with equal confusion in the literature. Following Locke and Hume most philosophers and a few psychologists differentiate between knowledge, belief, and opinion along a scale of certainty. Laird (1930) wrote, "In ordinary language, the word opinion is used to signify a weak or dubious assent that is not only not knowledge, but is also far less pronounced than belief" (p. 15). Gurnee (1936) agreed with Laird that "There is obviously a psychological difference between an opinion and a belief. The latter is accompanied by a feeling of certainty, the former not... a person will fight harder to maintain a belief than he will to maintain an opinion" (p. 250). Cardinal Newman (1955) gave opinion a stronger role: "I shall use the word [opinion] to denote an assent, but an assent to a proposition, not as true, but as probably true, that is, to the possibility of that which the proposition enunciates" (p. 64).

Contemporary psychologists, on the other hand, have in the main defined opinion independent of certainty and probability. Thurstone (1929) wrote, "The concept 'opinion' will here mean a verbal expression of attitude...An opinion symbolizes an attitude" (p. 7). Similarly F. H. Allport (1937) insisted that opinions are "instances of behavior" which "involve verbalization." Then, rather than give them a weak role, he insists that "The common stimulating situation [of opinion] must not only be well known, it must be a <u>matter of universal importance</u>. Mere interest is not enough..." (p. 13).

In Fairchild's <u>Dictionary of Sociology</u> (1944) no clear-cut distinction was made between opinion and belief. Belief was defined as "the acceptance of any given proposition as true. Such acceptance is essentially intellectual, although it may be strongly colored by emotion" (p. 23). On the other hand, opinion is defined as "a judgment held as true, arrived at to some extent by intellectual processes, though not necessarily based on evidence sufficient for proof" (p. 208).

It would seem then from this brief survey of the literature on belief, that any effort to make an experimental translation of belief and related concepts will find both support and opposition. The literature on belief is scarce compared with that dealing with most other psychological concepts. In the main, belief has been left to the philosopher. Only recently has the scientist's attention turned to opinion, with the development of public and private opinion polling.

Beliefs in the Presence of Things

First we consider an individual's beliefs in the past, present, and future existence of objects, events, and their properties. This includes what courses of action he believes to be available, what outcomes he believes to be producible, and what conditions he believes to affect the outcome. The word "thing" is used in this discussion to represent either objects, events, their properties, or combinations of these.

An individual only believes in the existence of things when they "make a difference" in his pursuit of his goals. Hence, any attempt to define what is meant by an individual's belief in the existence of a thing, must make reference to the outcome that he seeks. This can be done by constructing an environment in which the individual has intention for only one end. Now in such an environment, what does it mean to say that an individual has some degree of belief in the existence of a thing?

The simplest answer to this question would be that the individual is "acting as though" the thing were present. This is certainly the commonest characterization of belief to be found in the literature. Let us examine the feasibility of this suggestion. The literal translation would run somewhat as follows: when the thing is present, the individual practically always employs a certain course of action (e.g., when my wife is home I always say "Hello" when returning from a day's work). We cannot say that when we observe the individual select such a course of action, that he believes the thing to exist, because he may select the course of action quite regularly when the thing is not present (e.g., I may always call out "Hello" when entering my house). Hence, it appears that we have to add a further stipulation to make the chosen course of action a critical case for inferring belief: when the thing is not present, the individual never employs the course of action. But this suggestion, although it does provide a clear-cut way of determining whether the individual takes the thing to be present, really defines belief out of existence except in the sense of correct belief. Since the individual always acts in a certain way when the thing is present, and never acts in this ways when the thing is not present, then he can never incorrectly display belief in the presence of the thing; that is, he can never choose a behavior indicative of belief when the thing is absent, for when it is absent he never exhibits such behavior.

We can take care of this difficulty as follows. Suppose that when an individual responds to something relative to a certain objective he always (or almost always) displays a particular response, R; for example, when I perceive my wife on returning home and I want her to know that I am home, I always say "Hello." Suppose further that when my wife is not at home and I am aware of this fact, I never say "Hello" when entering the house. Now if I enter the house and do not observe my wife and am not aware of her absence, and say "Hello," an observer could conclude that I believe she is home, assuming, of course, I want her to know that I am home. Under these conditions my belief may or may not be correct. Note that if I do not want her to know that I am home, even when I observe her, I will not say "Hello." Hence belief must always be determined relative to an intended outcome.

5.4. Belief in the Presence (Absence) of Something. An individual believes that something (X) is present (absent) in his environment of type S relative to an objective (O), if he displays a response (R) when the following

> conditions hold: (1) he does not perceive X (or its absence), (2) in other environments of type S in which he has perceived the presence (absence) of X and intended O, he virtually always displayed R, and (3) when he was aware of the absence (presence) of X in environments of type S he virtually never displayed R.

Clearly, our ability to establish an individual's beliefs depends on our ability to find characteristic "belief responses" (R's) which can serve as belief indicators. These responses may be defined either structurally or functionally. As an example of the latter, when I enter a room my response in perceiving another person may not be to say "Hello;" it may be any greeting or just "talking."

One can perceive the absence of something as well as its presence. On returning home I can perceive that a familiar chair has been removed, or on arriving at my office, that my secretary is not there.

Note that when an individual believes that something is present he also believes that selecting the course of action that is the belief indicator (R) is efficient for accomplishing his objective. For example, I believe calling out "Hello" when I return home is an efficient way of letting my family know that I have arrived. Doing so has no efficiency for this end if no one is at home.

Consider the following three situations. In the first, on returning home I call out "Hello" and receive no response. I then go about other business whose efficiency does not depend on the presence of others. In the second, on returning home I call out "Hello" and again receive no response, but this time I hide behind a door to "scare" one of my children whom I expect to come looking for me. In the third, after receiving no response to my "Hello," I start a search of the house, calling out "Hello" periodically. One would conclude that my belief in presence of somebody was stronger in the second and third situations than in the first. In the third, however, I displayed more doubt than I had in the first and second.

"Doubt" seems to imply an intention to investigate the validity of a belief. The term is also used to connote a lack of decisiveness, a lack of belief one way or the other. In this latter sense "doubt" is a redundant concept, hence I will use it in the first sense: an intention to investigate. In this sense one can doubt a strong belief as well as a weak one. For example, a scientist may strongly believe in the existence of a particle that he has not observed, but still want to "prove" it.

The strength of a belief seems to be related to the amount of evidence required to change it. For example, if the belief that someone is at home is changed by one nonresponse to a "Hello," it is weaker than a belief that requires several nonresponses to several "Hellos." A very strong belief may not yield to any amount of contradictory evidence; the evidence is reinterpreted. If I believe strongly that someone is at home and get no response to my "Hello," I may assume that I have not been heard; for example, someone is at home but is in the basement or out back.

> 5.5. Intensity of Belief in the Presence (Absence) of Something. The intensity of an individual's belief in the presence (absence) of something (X) in his environment (S) relative to an objective (O) is one less than the number of times his belief response (R) must fail to produce O before his belief changes to one in the absence (presence) of X.

This number can range from zero to infinity. (I subtract one from the number of failures because it is convenient for intensity to have a minimum value of zero.)

Note that the intensity of a belief may change (usually decrease) with an increase in the number of failures of the belief response to produce the objective.

5.6. Degree of Doubt of the Presence (Absence) of Something. An individual's degree of doubt of the presence (absence) of something (X) in his environment (S) relative to an objective (O) is his degree of intention to become aware of the presence (absence) of X.

A subject's degree of doubt may also decrease with an increase in the number of failures of the belief response to produce his objective. Eventually his doubt may be completely dispelled.

Since the degree of intention can range from zero to one, the degree of doubt can also.

The "strength" of a belief should reflect both its intensity and the degree of doubt associated with it. It should increase as intensity increases and as the degree of doubt decreases. This suggests that the strength of a belief can be taken as the product:

(Intensity of Belief) (1 - Degree of Doubt)

Repeated failures of the belief response to produce the subject's objective necessarily reduces the intensity of belief and may reduce his degree of doubt. For example, suppose that it takes four failures to change a belief. Then the intensity is $4 - 1 = 3$. But after the first failure only three are required to change the belief and hence its intensity is reduced. Therefore, intensity must be reduced more rapidly than doubt if the strength of belief is to decrease.

These and another measure of belief will be discussed below in connection with beliefs in the efficiencies of courses of action. They are not discussed in connection with those types of belief considered in the intervening sections because their application to these concepts is relatively straightforward.

Now let us consider something in an environment other than the one occupied by the subject.

5.7. Belief in the Presence (Absence) of Something in Another Environment. If a subject selects a course of action (C_1) when he desires an objective (O_1), and he is aware that C_1 has no efficiency for O_1 in his environment unless X is present (absent) in another environment, then he believes that X is present (absent) in that other environment.

For example, if I phone a friend at his home when I want to give him some information, I am aware that so doing is efficient only if he is at home. Therefore, when I phone to give him information I believe he is at home. Of course, I may phone him to determine whether he is at home. (Note that the objective has changed.) Hence phoning when I want to give him information shows belief in his presence there; but when I want to find out where he is, phoning only indicates belief in the efficiency of so doing for this purpose, not belief that he is there.

If I phone my friend and am not "certain"--do not believe strongly--that he is home, if I get "no answer" I hang up and change my belief to "he is not home." If I am certain he is, I will assume something to be wrong in my dialing, or in the phone, or even with my friend, and proceed to determine which of these is true. It is apparent then that the amount of evidence one requires to change a belief depends on how strongly he holds that belief. In the situation just described if I re-phone my friend it indicates more doubt of the efficiency of my behavior than I have of his presence at home.

Now suppose on calling a a friend's house I find no one at home and leave a note. From this one is likely to infer that I expect him to return at a later time; that is, I believe he will be present in that environment at a subsequent time. My leaving a note for him would have no efficiency for my desired outcome (e.g., to have him call me later unless he were to return).

5.8. Belief in a Future Event: Expectation. If an individual selects a course of action, C_i, at a time t_o, when he pursues an objective, (O_j), at a later time (t_1), and he is aware (1) that C_i at t_o has no efficiency for O_j at t_1 unless X is present (absent) in the environment before t_1, and (2) X is absent (present) in the environment at t_o; then he can be said to expect (or believe that) X will be in that environment before or by t_1.

Consider another example, suppose I put on a raincoat on a clear morning when I am aware that it is not raining because I want to be dry when I return home that evening. Then it can be asserted that I expect it to rain. If I had a different objective--for example, to leave my coat at a cleaners--this conclusion could not be drawn; or if a raincoat is the only coat available my wearing it would not show an expectation of rain. Examples such as these emphasize the importance of holding the objective constant in tests of belief.

Belief in the past presence of something in an environment is not very different than belief in a future presence. For example, when I go out to the front of my house each morning to get the morning paper I display belief in the earlier presence of the delivery boy.

5.9. <u>Belief in a Past Event.</u> If an individual selects a course of action, C_1, at a time t_1 when he pursues an objective, O_1, and he is aware that C_1 at t_1 has no efficiency for O_1 unless X was present (absent) in the environment before t_1, then he can be said to believe that X was present (absent) in the environment before t_1.

Now that the basic types of belief have been taken care of, we can consider the six previously cited types of belief which are the elements out of which an individual's model of a choice situation is constructed.

<u>Belief in Courses of Action, Relevant Variables, and Outcomes</u>

A course of action is something that an individual does; hence it is an event: a change in one or more of his properties. It may involve use of an instrument (e.g., a car or a telephone), or it may not (e.g., in walking). An individual believes a course of action is available if he believes he is capable of doing what is necessary and any required instruments or environmental conditions are present. For example, he believes he can "use a telephone" if he believes (1) a telephone is available, (2) it is in working condition, and (3) he knows how to use it and is capable of doing so. Note that the required beliefs are beliefs in the presence of properties of the environment and himself, and of required instruments, if any. This kind of belief has already been defined (see 5.4). Therefore, the remaining task is to determine what environmental and personal properties, and instruments a subject believes must be present (i.e., are necessary) if he is to carry out a course of action.

5.10. Belief in Necessity. A subject believes something (X) in environment S_1 at time t_1 is necessary for something else (Y) in environment S_2 at a later time t_2 if he believes (1) whenever Y occurs in S_2 at t_2, X was present in S_1 at t_1; and (2) if X is not present in S_1 at t_1, Y will not be present in S_2 at t_2.

The environments S_1 and S_2 may be the same.

5.11. Belief in Sufficiency. A subject believes something (X) in environment S_1 at time t_1 is sufficient for something else (Y) in environment S_2 at a later time t_2, if he believes that whenever X occurs in S_1 at t_1, Y will occur in S_2 at t_2.

5.12. Belief in Producer-Product. A subject believes something (X) in environment S_1 at time t_1 is a producer of something Y in environment S_2 at a later time t_2 if he believes that X in S_1 at t_1 is a necessary but not a sufficient condition for Y to occur in S_2 at t_2.

5.13. Belief in the Availability of a Course of Action. An individual believes a course of action (C_1) is available to him in a choice environment (S) if he believes all the properties of S and himself, and the instruments that he believes are necessary to take C_1 are present in S.

Belief in Relevant State Variables

An individual's model of a choice situation contains uncontrolled variables, properties of the state which he believes affect the outcome of his choice. Determination of what properties of a state an individual believes to be relevant is closely related to determining what courses of action he believes are available.

5.14. Belief in Control of a Property of the Choice Situation. An individual believes he can control a property of his choice environment if he believes that choice at time t_1 of one of the courses of action believed to be available to him, will produce a change in that property at a later time t_2.

For example, if an individual believes that manipulation of a thermostat will produce a change in room temperature, then he believes "room temperature" is a controllable variable. If, on the other hand, he believes that he can do nothing to affect the weather and that weather will affect the outcome of what he does, then he believes "weather" is an uncontrollable variable.

5.15. Belief in Relevant Uncontrolled Properties of the Choice Situation. An individual believes that a property of a choice situation is a relevant uncontrolled property of that situation if he believes (1) he cannot produce a change of that property, and (2) that property is a (co)-producer of the outcome of one or more of the courses of action he believes to be available to him.

Note that these are the properties of a choice situation about which an individual may want information. Furthermore, his expectations as to what outcomes his behavior will produce are based on his beliefs as to what are the relevant uncontrolled variables and their values.

An outcome of a course of action in a choice situation is the set of changes in the properties of the subject and his environment which are produced by that course of action.

5.16. Believed Outcomes. An individual believes an outcome (O_j) is possible in a choice environment (S) if he believes that one or more of the courses of action that he believes to be available, can produce O_j in S.

Hypotheses and Assumptions

Up to this point we have only considered how to determine that an individual believes that something, X, is or is not, will or will not be, was or was not, present in an environment. The X's can be objects, events, courses of action, or properties of these.

Now we turn our attention to belief in the presence or existence of things which have not been perceived; for example, living things on Mars, the ether, and God. In such cases we clearly cannot determine how an individual responds to such things when he perceives them. Hence, the previously described test of belief does not apply here. The X's involved in such beliefs--be they objects, events, or properties--can be called hypothetical.

5.17. Hypothesis: a belief in the past, present, or future existence of something which has never been perceived.

How can we determine an individual's belief in an object which either has not or cannot be perceived. The answer is that we must determine how the individual would behave if X did exist and he perceived it. To do so does not raise any unique experimental problem.

The determination of what properties an individual actually has in "this" environment is no more direct an investigation than the determination of what properties an individual would have in any specified environment. This should be clear from all that has been said above. To determine, for example, what an individual intends or knows in this environment requires our developing a concept of a model (controlled standard) environment and determining what the individual would do in that environment. Therefore, the determination of what properties an individual has always depends on the determination of what an individual "would do _if._" Even if our task is to determine whether or not an individual "selects" a specified course of action in this environment, we must employ the producer-product model in an idealized environment and relate _this_ environment to it. The selection of a course of action is not determined by so-called "direct observation" any more than is knowledge of it. The process of determining what _is_ and what _would be_ are methodologically similar; the "would" presents no unique problems.

If we know, for example, how an individual responds to various climates, we can find techniques for inferring how he would respond to a climate in which he may never have been. If we know how an individual responds to various forms of authority we can infer how he would behave in response to the presence of so complete an authority as God is defined to be. These problems are analogous to determining how a body would fall _in a vacuum_ on the basis of observations made in something that is never quite a vacuum.

Once we have determined how an individual would respond to a hypothetical X, then the procedure for determining whether he believes that X to be present corresponds exactly to the general description given in definition 5.5 for determining belief in "real" things. For example, with respect to the end of "saving his soul," we could determine how an individual would respond to the existence of God. On the level of common sense, at least, we would say that prayer is a type of behavior that indicates a belief in God. Also we would say in most cases that if the individual does not believe in God, he would not pray. Hence, in this environment, we can perhaps take prayer to be an indication of belief in God. Similarly, we can determine how a scientist would respond to the presence of ether if he observed it, and infer from this whether he believes it to exist.

We sometimes use "assume" and "believe" synonymously, but this is careless. Clearly, an individual may assume something he does not believe, as well as something he does believe, or he may assume something which he neither believes nor disbelieves. In assuming X one acts as he would if he believed it, but with an important additional condition: he does so for the purpose of determining the consequences (outcomes) of the belief. This purpose may or may not be conscious. If not conscious it is referred to as an <u>implicit</u> assumption. If conscious, it is <u>explicit</u> and frequently takes the form of a supposition, axiom, or postulate. The latter two are linguistic representations of assumptions.

> 5.18. <u>Assumption</u>. An individual (A) assumes something (X) in a choice situation (S) if (1) a belief in X in S would produce different behavior of A in S than would non-belief, (2) A behaves as he would if he believed X in S, and (3) he intends to determine (i.e., perceive) the consequences of this belief (i.e., what outcomes such belief behavior produces).

To <u>pretend</u> something is true is not quite the same as to assume it is true; furthermore, to believe and to <u>make-believe</u> are not equivalent. We distinguish, for example, between the psychotic who believes he is Napoleon and the actor or masquerador who pretends or makes believe he is Napoleon. In ordinary language we would characterize make-believe as "acting as though X were so, but really knowing better." It is the "but really knowing better" which provides the clue for making the distinction between "believe" and "make-believe" more precise.

First, it is to be noted that the behavior of the individual who makes-believe could be interpreted as belief were we to ignore certain aspects of it. For example, the actor who moves heavy furniture about on the stage that he occupies before an audience makes-believe that no one else is present to aid him. If his intention were only to move the furniture (O_1) we would say that he believes no one else to be present; that is, relative to O_1 his behavior could be interpreted as belief in the absence of other people. However, we know his intention is to entertain the audience (O_2), and that his behavior has some efficiency for this second outcome. In addition we know that relative to O_2 the actor senses the presence of other people, that is, relative to O_2 he behaves efficiently in his response to the presence of the audience. He is aware, in addition, that his behavior can be interpreted as belief in the absence of people, and it is precisely for this reason that he performs it, since such interpretation on the part of the audience is necessary for the actor's attainment of O_2.

> 5.19. To Make-Believe or Pretend. An individual (A) makes-believe or pretends that he believes something (X) in a choice environment (S) if (1) he does not believe X in S, (2) he behaves as he would if he believed X in S, and (3) he believes that such behavior will produce a response in one or more individuals that he (A) intends to produce.

This definition appears to be self-contradictory: how can an individual display a characteristic belief response and not believe what is indicated? The answer lies in the fact that R, which is a belief response when the subject's objective is O_1, may not be a belief response when his objective is O_2. In make-believe he pretends to have objective O_1 but doesn't. An actor may pretend to want to harm another actor without actually wanting to do so.

It seems appropriate to bring this section to a close with an illusion.

5.20. Illusion. An individual has an illusion of something (X) in a choice environment (S) if he does not perceive X in S but believes he does.

Beliefs in Efficiency

There are many situations in which an individual has very high intention for an outcome and yet does not select the most efficient course of action for pursuing it. We sometimes "explain" such a choice by saying he believed that the course of action he did select was the most efficient available.

If we observe an individual put on a raincoat on a cold clear day we do not necessarily conclude that he believes wearing a raincoat to be the most efficient way of keeping warm. He may, as a matter of fact, believe that wearing a raincoat has a very low efficiency for this purpose, but he may want to take his overcoat (concerning which he has a higher opinion) to the tailor for cleaning, or he may merely want to take the raincoat to be repaired. As long as there is the possibility that the individual in this environment is pursuing many different ends we cannot use his behavior directly as evidence of what efficiency he believes a course of action to have with respect to any one outcome, for we do not know with respect to which outcome his behavior can be taken as an indicator of such belief.

To determine an individual's belief in the efficiency of a course of action for any outcome, it is necessary for us to isolate the outcome so that his choices cannot be taken to be serving any other objectives. If we know that an individual wants to keep warm, and has no other conflicting objectives, and further that when he wants to keep warm he almost always wears a raincoat, we would then take his behavior as indicating a belief in the efficiency of wearing the raincoat for that purpose.

The first condition, then, to be incorporated in the definition of a "belief environment" is that the individual have intention for one and only one outcome. But where the individual is only interested in, say, keeping dry, the fact that he repeatedly wears a raincoat may not indicate that he believes the act to be the most efficient possible. First, he may not have any other course of action available which he believes to be more efficient. He may, as a matter of fact, believe a woolen overcoat to be much more efficient, but such a coat may not be available to him. Then the repeated choice of wearing a raincoat in such an environment can at best indicate a belief in its <u>relative efficiency</u>; that is, a belief that wearing a raincoat is the most efficient <u>available</u> means for keeping warm in that environment.

An individual who is faced with the problem of making a difficult calculation may repeatedly use a slide-rule, even when a calculating machine is available, and yet we might consider him to believe that use of a calculator is more efficient than use of a slide rule. He may be unfamiliar with the machine and not know how to use it, and refrain for this reason. It is necessary, therefore, to distinguish between his use of the instrument and someone else's. The repeated choice of a behavior pattern in the "belief environment" can only be taken to indicate belief in the relative efficiency of his use of a course of action in that environment.

Suppose now that an individual has only one intended outcome, that of obtaining an answer to a complicated mathematical problem, and further that he has only two potential courses of action in the environment: use of pencil and paper, and use of a slide-rule. Then we can take the relative frequency with which he selected each course of action as an indication of his degree of belief in its <u>Maximum Relative Efficiency</u> (to be designated hereafter as "MRE"). If he always selected the slide-rule in this situation we would say that he believes with certainty in the MRE of the use of the slide-rule relative to his objective in that environment; and that he has absolutely no belief in the MRE of the pencil and paper calculations. If, on the other hand, his probability of choice of the slide-rule is 0.75, and his probability of choice of the other course of action is 0.25, then we would not take him to be absolutely sure of the MRE of the use of the slide-rule for that end. He would be surer, however, of the efficiency of the slide-rule than of the efficiency of using pencil and paper. Where probabilities of choice are equal, then his degrees of belief in the MRE's of both are equal.

Unless we are careful belief in the MRE of a course can be confused with familiarity with it. It is necessary, therefore, to construct the "belief environment" so that the intrinsic values of the courses of action studied are not compounded with beliefs in MRE's.

5.21. Degree of Belief in Maximum Relative Efficiency of a Course of Action. This measure of an individual (A) with respect to a course of action (C_i), an objective (O_j), and a set of alternative courses of action which A believes to be available and for each of which and C_i he has the same degree of familiarity relative to O_j in the choice environment (S), is the probability of A's choosing C_i in S when his degree of intention for O_j is 1.0.

If the course of action in the available set for which the subject has maximum degree of belief in its MRE actually has maximum relative efficiency, we would be inclined to say that the individual's belief is true and that he knows the MRE of that course of action. The relationship between belief and knowledge was commented on by J. S. Mill (1865) as follows:

> We do not know a truth and believe it besides, the belief is knowledge. Belief altogether, is a genus which includes

> knowledge; according to the usage of language, we believe whatever we assent to; but some of our beliefs are knowledge, others only belief. The first requisite, which, by universal admission, a belief must possess, to constitute it knowledge, is that it be true (p. 80 fn.).

The degree of belief in the MRE of a course of action is not equivalent to the degree of knowledge of the MRE of that action, but they are related. Just how becomes apparent in the following definition.

> 5.22. Degree of Knowledge of MRE. An individual's degree of knowledge of the MRE of a course of action (C_i) relative to an objective (O_j) in a choice environment (S) is his degree of belief in its MRE when C_i actually is the most efficient course of action for O_j in S.

Therefore, an individual's degrees of belief and knowledge in the MRE of a course of action are equivalent when the course of action involved is the most efficient available. If it is not, then the degree of belief is false and hence does not constitute knowledge.

An individual's degree of belief in the MRE of a course of action relative to an objective is different from the intensity of his belief in its MRE.

> 5.23. Intensity of Belief in MRE. An individual's intensity of belief in the MRE of a course of action (C_i) for an objective (O_j) relative to a set of courses of action which he believes to be available in a choice environment (S), is one less than the number of failures of C_i to produce O_j in S which are required to change his degree of belief in the MRE of C_i to zero.

The individual's degree of doubt of the MRE of a course of action is another matter.

5.24. Degree of Doubt of Relative Efficiency. An individual's degree of doubt about the relative efficiency of a course of action (C_i) relative to an objective (O_j) in a choice environment (S) is his degree of intention to become aware of that relative efficiency.

An adequate measure of the strength of a belief in the MRE of a course of action should be a function of its degree, intensity, and the degree of doubt associated with it. It should increase with increases in the first two and decrease with increases in the last. It should be zero when either of the first two measures are at their minimum (i.e., zero) or the last is at its maximum (i.e., one). A measure which would satisfy these conditions is

(Degree of Belief) (Intensity of Belief) (1-Degree of Doubt).

This measure of belief can range between zero and infinity. It is shown graphically in Figure 5.1.

Believed Relative and Absolute Efficiencies.

It is to be noted that the probability of choice in the belief environment provides a measure of the degree of belief in the MRE and does not indicate how efficient (in the absolute sense) the individual believes the courses of action to be. Nor does the probability of choice of each of the alternative courses of action indicate how the individual ranks their efficiency; it merely indicates which course the individual most believes to have the MRE.

We can "move up one notch" now and consider how to determine what relative efficiency an individual believes courses of actions of action to have relative to an outcome, O, in a state, S. By relative efficiency of courses of action I mean the ratios of their efficiency. Therefore, if one course of action has an efficiency of 0.08 and another an efficiency of 0.04 for an outcome, then "2" and "1" or "1.0" and "0.5" represent

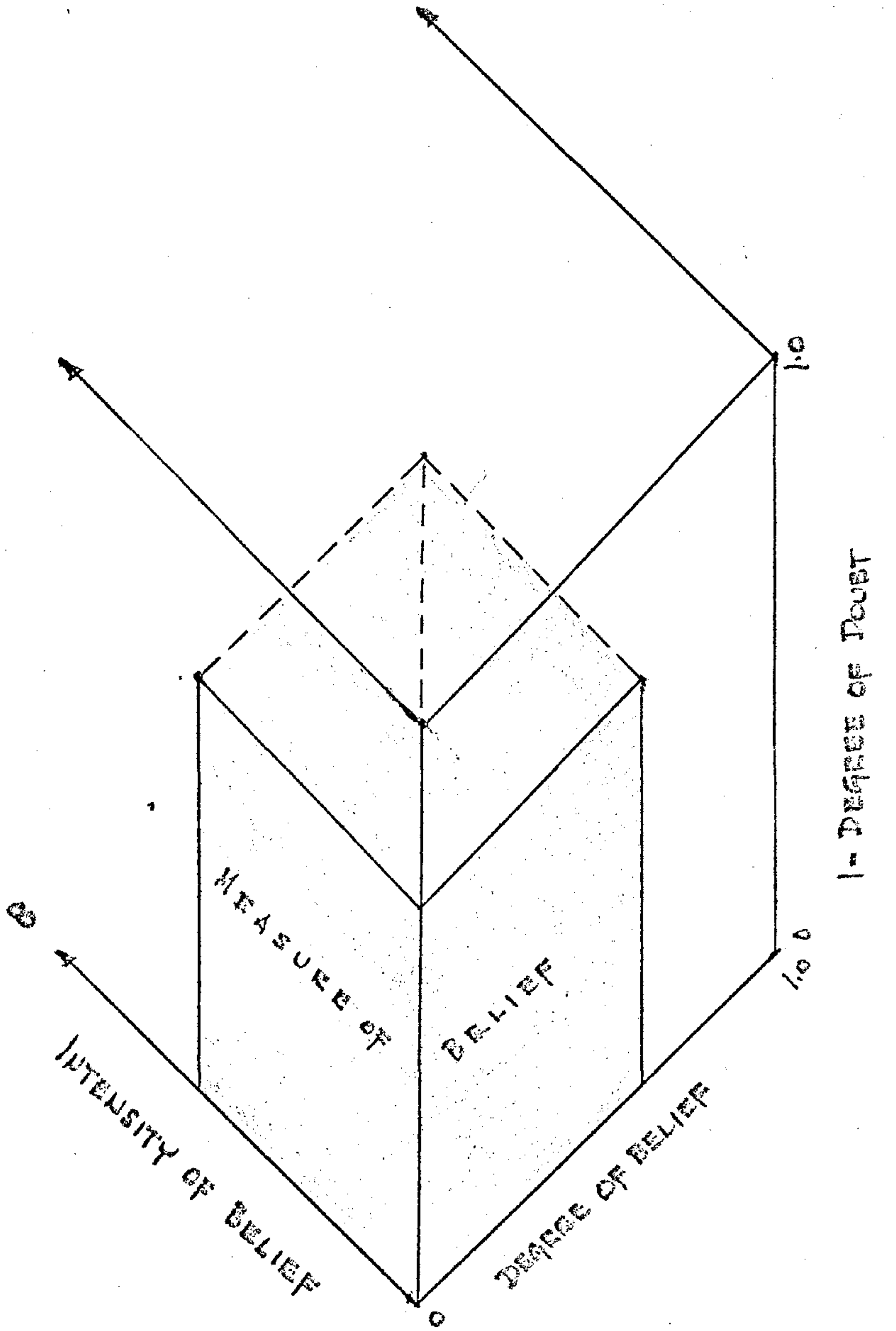


Fig. 5.1. Illustrative measure of belief in the HRE of a course of action.

their relative efficiencies. For convenience I will express relative efficiency as the ratio to the highest efficiency of the set, thereby yielding a scale between 0 and 1.0.

Suppose an individual in S has a degree of intention of 1.0 for a specific outcome, O. This is the same situation we used to determine the degree of belief in the MRE of the available courses of action. In this situation we assumed the subject could select any course of action; that is, the selection and occurrence of each course of action is the same. Now suppose we separate them by having him indicate which course of action he wants to select. Then we can control the probability that the course of action will in fact occur. Let $a_1, a_2, \ldots, a_m$ represent the probabilities of occurrence that we attach to $C_1, C_2, \ldots, C_m$. The subject is made aware of these. We then seek a set of values of $a_1, a_2, \ldots, a_m$ such that the probabilities of choice are equal: $P_1 = P_2 = \ldots = P_m = \frac{1}{m}$. These values of the a's give us the values of the relative efficiencies that the subject believes the courses of action to have. For example, if $a_1 = 1.0$ and $a_2 = 0.4$ we conclude that the subject believes C_2 to be $\frac{1.0}{0.4} = 2.5$ times as efficient as C_1, or C_1 to be $\frac{0.4}{1.0} = 0.4$ times as efficient as C_2.

Note that when the courses of action and outcomes available in a state are defined so as to form exclusive and exhaustive sets although $\sum_{j=1}^{n} E_{ij} = 1.0$, $\sum_{i=1}^{m} E_{ij}$ may take on any value from O to m. For this reason we cannot translate the believed relative efficiencies directly into believed (absolute) efficiencies.

If the number of courses of action (m) in the exclusive or exhaustive set is equal to or greater than the number of outcomes (n) in the exclusive and exhaustive set of these, then the believed absolute efficiencies can be determined. This follows from the fact that we can form m equations in n unknowns.

For example, suppose there are two courses of action, C_1 and C_2, and two exclusive and exhaustive outcomes, O_1 and O_2. Let a_1 and a_2 represent values associated with C_1 and C_2 relative to O_1 (when intention for it is equal to 1.0), which the probabilities of their choice equal (i.e., $P_1 = P_2$); and b_1 and b_2 represent the corresponding values with respect to O_2 (when intention for it is equal to 1.0). Let these values be as follows:

	O_1	O_2
C_1	$a_1 = 0.5$	$b_1 = 1.0$
C_2	$a_2 = 1.0$	$b_2 = 0.33$

Now we can formulate the following equations:

$$a_1 E_{11} = a_2 E_{21} \tag{5.1}$$

$$b_1 E_{12} = b_2 E_{22} \tag{5.2}$$

$$E_{11} + E_{12} = 1.0 \tag{5.3}$$

$$E_{21} + E_{22} = 1.0 \tag{5.4}$$

Then from (5.1) and (5.2) we get

$$E_{11} = \frac{a_2}{a_1} E_{21} \tag{5.5}$$

$$E_{12} = \frac{b_2}{b_1} E_{22} \tag{5.6}$$

Substituting in (5.3) yields

$$\frac{a_2}{a_1} E_{21} + \frac{b_2}{b_1} E_{22} = 2E_{21} + 0.33 E_{22} = 1.0 \tag{5.7}$$

Multiplying (5.4) by 2 we get

$$2 E_{21} + 2 E_{22} = 2.0 \tag{5.8}$$

Substracting (5.7) from (5.8) yields

$$\frac{5}{3} E_{22} = 1.0 \tag{5.9}$$

or

$$E_{22} = 0.6 \tag{5.10}$$

Then, from (5.4), $E_{21} = 0.4$, and from (5.1) and (5.2), $E_{11} = 0.8$ and $E_{12} = 0.2$.

Therefore, in formulating a model of the subject's choice situation if we do so in such a way that $m \geq n$, we can determine what are the subject's believed efficiencies of each course of action for each outcome. These believed values are sometimes called "subjective efficiencies."

This discussion can be summarized in an awkward and lengthly definition:

5.25. Belief in Efficiencies. In a choice environment in which a subject (A) believes

(1) n exclusive and exhaustive outcomes are possible "O_j, where $j = 1, 2, \ldots, n$),

(2) m ($m \geq n$) exclusive and exhaustive courses of action are available (C_i, where $i = 1, 2, \ldots, m$), when his intention for O_j is 1.0, then for a set of probabilities $\{a_{ij}\}$ which are associated with C_1, $C_2, \ldots, C_m$ respectively so that $P_1 = P_2 = \ldots = P_m$; the believed efficiencies of the C_i's for the O_j's are those values of E_{ij} which satisfy the following set of equations:

$$a_{11} E_{11} = a_{21} E_{21} = \ldots = a_{m1} E_{m1}$$
$$a_{12} E_{12} = a_{22} E_{22} = \ldots = a_{m2} E_{m2}$$
$$\vdots$$
$$a_{1n} E_{1n} = a_{2n} E_{2n} = \ldots = a_{mn} E_{mn}$$
$$\sum_j E_{ij} = 1.0.$$

Belief in Probabilities

Suppose an individual (A) believes that two states are possible: S_1 and S_2 (e.g., it will or it won't rain today). Relative to an objective (O_j) for which his degree of intention is 1.0, suppose he believes the efficiency of a course of action, C_1, is 1.0 if S_1 pertains and O if S_2 pertains. Correspondingly, he believes the efficiency of C_2 is O if S_1 pertains and 1.0 if S_2 pertains. No other courses of action are available. Now we construct a choice situation in which probabilities a_1 and a_2 ($a_1 + a_2 = 1.0$) are associated with C_1 and C_2 respectively so that (1) if A selects C_1 he will be able to carry it out a_1 portion of the time with C_2 occurring ($1-a_1 = a_2$) portion of the time; and (2) if he selects C_2 it will "materialize" a_2 portion of the time with C_1 occurring ($1-a_2 = a_1$) portion of the time. Then we find the values of a_1 and a_2 for which A's probabilities of selecting C_1 and C_2 are equal (i.e., $P_1 = P_2 = 0.5$). Where this is so $a_1\ C_1$ and $a_2\ C_2$ are equally preferable to A.

Now we can determine what he believes to be the probabilities of S_1 occurring (p_1) and S_2 (p_2). For example, suppose $a_1 = 0.4$ and $a_2 = 0.6$. Then, since $P_1 = P_2$

$$0.4\ p_1 = 0.6\ p_2. \tag{5.11}$$

Solving, we get

$$0.4\ p_1 = 0.6\ (1-p_1) \tag{5.12}$$

$$p_1 = 0.6 \tag{5.13}$$

$$p_2 = 1-0.6 = 0.4. \tag{5.14}$$

A corresponding procedure can be used when more than two possible states are involved. For example, consider three states (S_1, S_2, and S_3) for which we have found that when $a_1 = 0.5$, $a_2 = 0.3$, and $a_3 = 0.2$; $P_1 = P_2 = P_3 = 0.33$. Then

$$0.5\ p_1 = 0.3\ p_2 = 0.2\ p_3 \tag{5.15}$$

$$p_3 = (1 - p_1 - p_2) \tag{5.16}$$

$$0.5\, p_1 = 0.2\,(1 - p_1 - p_2) \tag{5.17}$$

$$0.7\, p_1 + 0.2\, p_2 = 0.2. \tag{5.18}$$

Multiplying through by 3/2 yields

$$1.05\, p_1 + 0.3\, p_2 = 0.3 \tag{5.19}$$

Adding

$$0.5\, p_1 - 0.3\, p_2 = 0 \tag{5.20}$$

yields

$$1.55\, p_1 = 0.3$$

$$p_1 = 0.19. \tag{5.21}$$

Then, from

$$0.5\,(0.19) = 0.3\, p_2 = 0.2\, p_3 \tag{5.22}$$

we set

$$p_2 = 0.32 \text{ and } p_3 = 0.48. \tag{5.23}$$

Note that the values of the a's essentially reflect the "odds" that the subject is willing to set for each course of action to yield "fair bets".

This discussion, like the last, may also be summarized in a rather awkward and lengthly definition as follows:

5.26. Belief in Probability of States. In a choice environment in which

(1) a subject (A) believes n states are possible (S_1, S_2, ... S_n),

(2) A believes there are n exclusive and exhaustive courses of action available (C_1, C_2, C_n) such that

(3) relative to an objective (O_j) for which A's degree of intention is 1.0,

(4) he believes $E_{1j} \mid S_1 = 1.0$, $E_{1j} \mid S_k \; (k \neq 1) = 0$; $E_{2j} \mid S_2 = 1.0$, $E_{2j} \mid S_k \; (k \neq 2) = 0; \ldots; E_{nj} \mid S_k \; (k \neq n) = 0$

then, for a set of probabilities $[(a_1, a_2, \ldots, a_n)$ where $a_1 + a_2 + \ldots + a_n = 1.0]$ associated respectively with $C_1, C_2, \ldots, C_n$, for which $P_1 = P_2 = \ldots = P_n = 1/n$; the believed probabilities of $S_1, S_2, \ldots, S_n$ (i.e., $p_1, p_2, \ldots, p_n$) are those values for which

$$a_1 p_1 = a_2 p_2 = \ldots = a_n p_n$$

where

$$p_1 + p_2 + \ldots + p_n = 1.0.$$

Believed probabilities are frequently referred to as <u>subjective</u> probabilities. I prefer the term used by Cowan (1947): <u>credibilities</u>.

Note that when an individual is <u>asked</u> what probability he believes X to have he may reply with a two-place decimal; for example, "0.25." But in his behavior he may not discriminate between values from, say 0.10 to 0.40. Therefore, verbal testimony must be treated with care. More will be said on this point below in the discussion of <u>opinions.</u>

<u>Beliefs in Intentions and Utilities</u>

The only aspect of an individual's model of a problem situation that remains to be considered involves the values he places on the outcomes that he believes are possible.

Many would argue that what an individual thinks (believes) he wants and what he wants are the same thing. Such an argument must either be based on fact or on a tautology; that is, what an individual believes he wants and what he wants are defined to be the same thing. To define them as equivalent is to ignore a commonly made distinction between the two. Most would agree that what one individual, A, wants, and what another individual, B, believes A wants, are not necessarily the same thing. We certainly do not want to equate these by definition. If our definition allows them to differ and is general enough to allow A and B to be the same individual, then it becomes a question of fact whether what an individual wants and what he believes he wants are the same.

A parent who believes his child wants to learn how to play the piano behaves differently from one who doesn't. The desire to play the piano--a high degree of intention to do so or a high utility placed on doing so--is a property of the child. To determine whether or not a parent believes the child to have such a property is the same as determining whether he believes the child to have any other property, particularly a functional property. Once the belief indicators have been identified, we would proceed as described above in the discussion of belief in the presence of objects, events, and properties of either (5.14).

One should proceed the same way to determine what properties an individual believes he has. We often say of another that he believes that he knows more than he does or he is less or more generous than he thinks he is, and so on. Of course, an individual may be as smart or generous as he believes he is. The point is that we commonly distinguish what properties an individual believes himself to have and those which he actually has.

Therefore, to determine what relative value or utility an individual believes an outcome, O, has for him in a state, S, we must find a type of behavior he displays almost invariably when he is aware of this utility and which he almost never displays otherwise.

An individual may not be aware that he has a certain illness, physical or psychological; for example, paranoia. If paranoic, he may not believe he is. If not paranoic, he may believe he is. A doctor can make one aware of an illness of which he was previously unaware. An individual can become aware of his own relative values or utilities either by observing himself under appropriate conditions or by being informed by someone who has so observed him.

Suppose that when an individual, A, is aware (relative to an outcome that he intends, O_1) that another individual, B, has a high intention for an outcome O_2, he displays a characteristic response R which he virtually never displays when he is aware that B has low intention for O_2. Then if A is not aware of B's intentions and A has high intention for O_1 and displays R, he can be said to <u>believe</u> that B has a high intention for O_2.

For example, when I (A) want to please my wife (O_1) and am aware that she (B) wants a particular household appliance (O_2), I buy it for her. I never do so when I am aware of the fact that she does not want a particular appliance. Then, if I am observed buying an appliance to please her when I am not aware of her desires, I can be said to believe that she wants it.

5.27. <u>Belief in Relative Values of Outcomes</u>. If when an individual (A) is aware (relative to an objective, O_1) that another individual (B, who may or may not be the same as A) has intention for another outcome (O_2, where O_1 and O_2 need not be exclusive), he displays a characteristic response (R) which he virtually never displays when he is aware that B has low intention for O_2; then if when A has high intention for O_1 and he is not aware of B's intentions for O_2, he displays R, he believes that B has high intentions for O_2.

For example, suppose that when an individual wants to relax (O_1) and is aware of the fact that he also wants "to see a movie" (O_2) he virtually always goes to a motion picture theatre, (R) and virtually never does so otherwise. Then when he wants to relax and does not know (is not sure) whether or not he wants to see a movie, if he goes to a cinema he can be said to believe he wants to see a movie. An observer, noting his restlessness during the movie, may conclude that his belief is in error; that he does in fact not want to watch a movie. Indeed, the individual himself may become aware of this by observing his own restlessness.

The discussion up to this point has been concerned only with two levels of intention, above and below 0.5. It is possible, however, to divide the intention scale into smaller intervals and obtain characteristic responses for each. This would enable one to determine which of smaller ranges of intention an individual believes another or himself to have.

CONCLUSION

In this chapter I have considered the contribution of memory to the content of one's model of a choice situation. It is the source out of which relevant past experience is extracted in the form of beliefs. These beliefs provide the components of the model: courses of action, outcomes, relevant state variables, efficiencies, and relative values. If the individual has doubts about any of these it will be reflected in his evaluation of his model and possibly in the design of data acquisition and evaluation. The intensity of these beliefs affect the amount of data he requires to confirm or disconfirm his beliefs.

OPINIONS: A POSTSCRIPT

In this chapter I have discussed belief in connection with an individual's actions when he is confronted with a real choice situation.

But the typical questionnaire usually asks an individual what he <u>would</u> do were he in a certain situation, or in general, what he considers to be the most efficient course of action to be in an environment different from the one in which he is asked the question. For example, one public opinion poll asked, "If you were advising President Truman on cabinet appointments, what changes in the present personnel would you suggest?"* Sometimes these questions are posed in a different form, one in which the individual is asked whether some other individual, or agency "should" do so-and-so to accomplish some specified end most efficiently: "In order for the United States to continue as a great power, which branch of our service do you think we should spend the most money on after the war, our Army, our Navy, or our Air Forces?"**

A subject's answers to such questions are evidently taken to mean something. It demands some reconstruction to see just what they can mean. The reconstruction I will make here runs as follows: I have already said that <u>if</u> the individual could be observed in a problem situation, and <u>if</u> he had a maximum measure of belief in the maximum relative efficiency of one of the courses of action, then he would choose that course of action. But further, suppose that when the individual has such belief, he invariably responds "yes" in a certain environment to a specific question when it is posed to him. That is, the measure of belief, together with the question-stimulus, are co-producers (invariably) of a certain verbal response. We would also have to be careful to add that when the measure of belief is maximum for one of the alternative courses of action, then the individual will never display the assenting response. Further, since individuals are sometimes "in-between" on issues, it is necessary to add another category: if the individual's measure of belief is neither maximum for, nor maximum

---

*<u>Public Opinion Quarterly</u>, 9, 1945, p. 226, quoted from American Leadership Panel.

**<u>Ibid</u>, p. 254, quoted from a <u>Fortune</u> poll.

against the choice of a course of action, then this indecision, together with the question-stimulus will produce an answer of "undecided" or whatever other word the questionnaire uses to indicate this category. Such "signs"* or belief we shall hereafter call "opinions." Like any signs they may signify what is either true or false.

Many have asked whether actions are better indicators of belief than words. To ask such a question is to assume that belief is something which produces behavior rather than behavior itself. Since I have defined belief as a type of behavior the only relevant question is the following: Are verbal or non-verbal responses in belief environments better as the critical responses to be observed? This is a question that can only be answered empirically.

Consider the following statement by Thurstone and Chave (1929): "But his actions may also be distortions of his attitude. A politician extends friendship and hospitality in overt action while hiding an attitude that he expresses more <u>truthfully</u> to an intimate friend." One may certainly want to ask what "truthfully" means in this example. If the politician tended invariably to extend such hospitality and friendship, then would we not be inclined to say that his expressions to his "friends" were lies rather than truths? This does not mean that we indorse the use of "overt" actions alone as the criteria of measuring any of the psychological properties; speech is itself an "act," and there seems to be little reason for relegating it to a special class.

It may very well be that Thurstone's politician is really <u>making-believe.</u> The distinction between belief and make-believe is one that Thurstone and Chave did not make.

What we really want to know is <u>not</u> whether overt actions must be used, but whether any given response can be taken as critical in the

---

*The nature of "signs" will be considered in detail in Chapter 9 .

measure of his belief. In the case of the politician, we want to know how he acts when he is actually showing a friendly response, and how he will act when he is showing a non-friendly response. The critical behavior may be verbal or overt, but the fact that it is either does not necessarily guarantee its adequacy or inadequacy.

Every opinion poller expects that on occasion, individuals will give false opinions, in the sense that what they actually believe does not influence what they say. Opinions may be in error in two different ways. The type I error of an opinion questionnaire would be measured in terms of the probability of failing to elicit an opinion when the individual actually has a relevant belief. The type II error would be measured in terms of its probability of producing an opinion which does not signify the true belief.

The accuracy of taking the selection of an expression or message M by an individual (A) as his opinion relative to a belief in X with respect to an outcome (O) in a state (S_1) can be determined as follows:

(1) Determine the probability P_1 that when A has a measure of belief in X relative to O in S_1 lying in the range $m_1 - m_2$, he will select M in S_2.

(2) Determine the probability P_2 that when I has a measure of belief in X relative to O in S_1 <u>not</u> lying in the range $m_1 - m_2$, he will select M in S_2.

Then the couplet (P_1, P_2), are measures of the accuracy of the use of M.

The serious student of methodology cannot help but raise certain questions concerning the usual techniques of opinion polling. Perhaps the situation presented to the subject is totally "unreal." This means that what the subject would do in the hypothetical situation is not a producer of his behavior in his present environment. This possibility is far from an unlikely one. During World War II, for example, subjects were asked, "Would you like to see England and the United States attempt a large-scale attack on Germany in Western Europe in the near future, or do you think they should wait until they are stronger?"* Many subjects had no comprehension at all as to what it would actually have meant to decide such a momentous issue. That is, what course of action he would have pursued in such circumstances, and the intensity of his belief in the efficiency of that action, may have had no influence whatsoever on what he said in reply to the question. This would mean that the type I error of the questionnaire was at a maximum.

Usually public opinion polls are "tested" for their adequacy by seeing what actually does happen when an individual can exhibit a certain response. Thus, if we ask how he would vote for a candidate, and he replies he would vote "for" then we can test the validity of his reply by comparing our count of "fors" with the actual election results. The same checks can be made on many consumer polls. But election and consumer polls present situations with which the subjects are familiar. They can appreciate what it means to pull down the lever on the voting machine, or put a mark on a ballot, since they are accustomed to these actions. Hence, in such situations it may be that his belief in the right action will influence what he does in reply to the question. But it is certainly poor methodology to argue from these cases to an individual's replies about a labor-management dispute, our policy toward Viet Nam, the choice of cabinet members. What we apparently need on such questionnaires (if the problem could be solved by the use of words) are items to test the reality of the situation for the individual; we need another category besides "yes," and "no," and "undecided": the category of "realism."

---

*Public Opinion Quarterly, 6, 1942, p. 658, quoted from American Institute of Public Opinion poll.

The development of opinion polling has led certain psychologists to attempt to distinguish between what in the individual's response is "content," and what is "intensity" (Guttman and Suchman, 1947). Translating this distinction into the concepts used here it appears that the content question attempts to find out which action the individual would choose in the situation, and tries to elicit the <u>degree</u> of belief. The intensity question tries to elicit an indication of either the individual's <u>intensity</u> of belief or his degree of doubt. For example, the soldiers were asked during World War II, "In general, do you think the Army is trying its best to carry out the Army score card plan as it should be carried out?" The answers from among which they were to select one were, "Yes, it is trying its best; It is trying some, but not hard enough; It is hardly trying at all." This was followed by the question, "How strongly do you feel about this?" Again, the possible answers were, "Not at all strongly, not so strongly, fairly strongly, very strongly." (Guttman and Suchman, 1947, p. 60). The latter question was designed to measure "intensity of feeling on the issue." The remarks about the nature of this intensity are enlightening:

> If a zero point were to be defined, in what respect are two people the same if one is a certain distance <u>above</u> the zero point, and the other the <u>same</u> distance <u>below</u> the zero point? They are different in that the first is higher than the second in the content scale ordering, but they are the same in their distance from the zero point. What shall we name the second variable on which they are the same? The answer proposed here is to call this second variable the <u>intensity function</u>." (Ibid, 1947, p. 60).

Within the terminology adopted here, this would mean that two individuals could hold different degrees of belief on an issue, but have the same intensity or degree of doubt. It is supposed, of course, that the individual will exhibit a verbal response in accordance with the degree and intensity of belief and degree of doubt that he would actually exhibit in the hypothetical circumstance. The usual opinion polls are designed to determine only whether an individual's belief lies in a certain range. Intensity tests divide the measure of belief into its two components. The type I error of the test (as defined above) is probably less on this account.

BIBLIOGRAPHY

Allport, F. H., "Toward a Science of Public Opinion," The Pub. Op. Quart., Vol. 1, No. 1, Jan. 1937, pp. 7-23.

Bailey, S., Essays on the Formulation and Publication of Opinions. Ticknor and ields, 1854.

Childs, H. L., An Introduction to Public Opinion, John Wiley & Sons, Inc., 1940.

Colvin, S. S., The Learning Process. The Macmillan Co., 1915.

Cowan, T. A., "Probability and Credibility," Bull. Inst. of Exper. Meth., Vol. 1, No. 4, June 1947.

Dockeray, F. C., General Psychology, Prentice-Hall, Inc., 1932.

Fairchild, H. P., Dictionary of Sociology. Philosophical Library, 1944.

Guilford, J. P., General Psychology. D. Van Nostrand Co., Inc., 1939.

Gurnee, H., Elements of Social Psychology. Farrar & Rinehart, Inc., 1936.

Guttman, L., "A Basis for Scaling Qualitative Data," Am. Soc. Rev., Vol. IX, No. 2, April, 1944, pp. 139-150.

Guttman, L. and Suchman, E. A., "Intensity and a Zero Point for Attitude," Am. Soc. Rev., Vol. XII, No. 1, Feb., 1947, pp. 57-67.

Hume, D., A Treatise on Human Nature, 1939-40.

James, W., Psychology. Henry Holt & Co., 1890.

Judd, C. H., Psychology. Chas. Scribner's Sons, 1907.

Laird, J., Knowledge, Belief, and Opinion. The Century Co. 1930.

Lund, F. H., "The Psychology of Belief," Journ. Abn. Soc. Psych., Vol. 20, 1925-26, pp. 63-81, 174-96.

Mill, J. S. An Examination of Sir William Hamilton's Philosophy. Longman, Green, Longman, Roberts and Green, London, 1865.

Miller, J.G., Unconsciousness. John Wiley & Sons, Inc., 1942.

Newman, F. W., Grammar of Assent. Doubleday & Company, Inc., Garden City, New York, 1955.

Sherif, M. and Cantril, H., "The Psychology of 'Attitudes'," Part I, Psychol. Rev., Vol. 52, No. 6, Nov. 1945, pp. 295-319.

________________________, "The Psychology of 'Attitudes'," Part II, Psychol. Rev., Vol. 53, No. 1, Jan. 1946, pp. 1-24.

Thurstone, L. L. and Chave E. J., The Measurement of Attitudes. Univ. of Chicago Press, 1929.

________________________________, A Scale for Measuring Attitudes toward the Church. Univ. of Chicago Press, 1930.

Tolson, E. B. On Belief. W. O. Stout, 1941.

CHAPTER 6

EVALUATION OF MODELS: THOUGHT AND INTUITION

MIND, n. A mysterious form of matter secreted by the brain. Its chief activity consists in the endeavor to ascertain its own nature, the futility of the attempt being due to the fact that it has nothing to know itself with (Ambrose Bierce, The Devil's Dictionary).

THE NATURE OF MODELS

The collection of beliefs that an individual has that are relevant to his choice situation constitutes his model of that situation. Beliefs consist of attributions to, or relations between, concepts and/or images. Hence a model consists of concepts and/or images and properties and interrelationships attributed to them. The model is an individual's representation of his choice situation; hence, it is a sign of that situation. ("Concepts," "images," and "signs" are defined in Chapter 9.)

Models are used in choice situations because, as representations they are less "costly" to manipulate than is reality itself. In most cases it is clearly preferable to make one's trials and errors with a model than with reality. The "economy" and relative ease of model manipulation derive in part from the fact that they are usually simpler than reality. Every situation has an unlimited number of properties but only a relatively few of these are relevant to a particular choice. Hence models of choice situations are selective. For example, in the physicist's model of a falling body he may relate its acceleration to such properties as mass, shape, and wind currents; but not to color, age, cost, and chemical composition. Only those properties are included in a model which either the individual believes has an effect on the outcomes of interest (and therefore are relevant), or he is doubtful about and wants to investigate further.

Neither the model nor the way it is used may be made explicit in a choice situation; in fact the subject may be quite unconscious of both. Nevertheless it is possible for an observer to be conscious of either or both. By uncovering a subject's relevant beliefs, his model can be revealed. It is not at all uncommon, for example, to point out to another an assumption that he has made unconsciously in reaching a conclusion. To illustrate this consider the following problem.

There is a block of cells occupied by fifteen prisoners (P's) and a warden (W) as shown in Figure 6.1. Each cell is connected by a door to the cells adjacent to it. Only the warden's cell has an exit from the block. The prisoner, P#, in the lower left-hand cell is a homicidal maniac who is compelled to kill everyone he sees but he cannot look at a person he has killed; if he does he faints. One morning he is found missing and the occupant of every other cell than his is dead in his cell. What path did he take? (Try to solve this problem before reading on.)

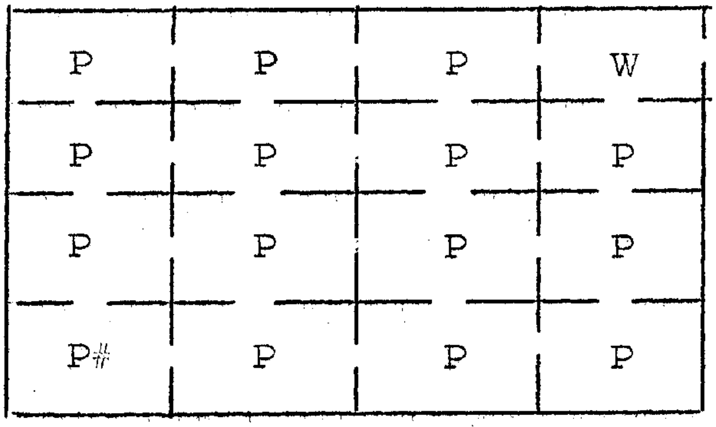


FIGURE 6.1

Most people try to solve this problem by looking for a path from the cell occupied by P# to W which goes through every cell once and only once and terminates in the Warden's cell. There is no such path. Yet the problem is solvable. Of most of those whom I have observed trying to solve this problem I can assert that *they have assumed that P# cannot return to any cell.* They frequently are not conscious of this assumption even though it restricts the alternatives that they try. As a matter of fact, P# can return to one cell; his own. Once this possibility becomes apparent, the solution is easy to obtain.

An individual's explicit formulation of a model--a representation of his beliefs and assumptions in sign-form--may not not be an accurate representation of his implicit model. Hence we sometimes say to someone, "That is not what you really believe about this situation." In scientific research the investigator tries to make his model explicit and to do so in such a way that others can evaluate it and, hopefully, use it as their own. In most of our every-day decisions, however, there is no pressure to do the same.

Models may take on one, or some combination, of three different forms: *iconic*, *analogue*, or *symbolic*. We shall consider each of these in turn.

Iconic Models

An iconic model is one made up of images or iconic signs of that which is represented. Iconic signs, which will be discussed in Chapter 9, are signs which have the same structural properties as that which they represent. Therefore, an iconic model looks like, sounds like, feels like, or in general can be observed to be like that which it represents. However, it is usually larger or smaller than what it signifies; that is, there is a transformation of scale. A photograph, for example, is an iconic model of a person's appearance. Building models, and automobile and ship models are common examples of iconic models. In general, such models are quite uniquely associated with the thing represented (e.g., a photograph of one person cannot be used to represent another person or at least not very many others). Such models are concrete, relatively easy to construct, but are usually difficult to modify or manipulate. That is, it is usually difficult to change the representation of the relevant properties; for example, the shape of a model aeroplane.

As will be indicated in Chapter 10, images are implicit iconic models; they are commonly called mental pictures of the structural properties of reality. When a person says that a photograph is not a good one, he means that it does not correspond to his image of that which it represents. For example, correspondence with his image is the basis for evaluation of different photographs of the same thing.

The economy of models is well illustrated by caricatures which attempt to minimize the number of properties required to represent a person.

It is not accidental that most toys for children are iconic models of real objects. Because they are, they provide children with an opportunity in play to practice manipulation of the real world, to gain experience in such manipulation before real opportunities and the need to do so arise for them. That which is a toy for children becomes an instrument for instruction of adults. "Dolls," for example, are used to instruct adults in anatomy, surgery, first aid, dress making, and so on.

Note that iconic models are intended to have the same relevant structure as that which they represent. To the extent that they do, they will have the same functional properties since two things with all the same structural properties necessarily have the same functional properties. However, since an iconic model has only <u>some</u> of the same structural properties as that which it represents, it will not have <u>all</u> of the same functional properties as its object. For example, a small motor-driven model aeroplane's gasoline consumption usually does not adequately reflect such consumption by the plane that it represents. Whether an iconic model lacks any of the relevant functional properties of a specific choice situation depends on how well the relevant structural properties have been selected and how well they are represented.

<u>Analogue Models</u>

In an analogue model one or more of the structural properties of the real situation is represented by a different structural property. For example, in a road map we may use different colors to represent either types of terrain, conditions of roads, or elevations. Or we may use water flowing through pipes or tubes as an analogue of electricity flowing through wire or money flowing through an economy. We can use distances on straight lines that are drawn perpendicularly to each other to represent units on a wide variety of scales, and a curve drawn between the two to represent relationships between them. Thus most graphs are simple analogue models.

The substitution of one structural property for another is frequently motivated by the desire for greater manipulatability. For example, it may be easier to control the flow of water than the flow of electricity or money.

Analogues are less specific than iconic models; that is, the unique situation that such a model is intended to represent may be difficult to identify from an examination of the model. Hence they tend to be more general and abstract than are iconic models.

Metaphors and similies are usually capsule-like analogue models. For example, to say of a search that it is "like looking for a needle in a haystack" is to say that although the needle and the haystack may differ from the current situation in many structural details, both situations have a common functional property--difficulty of search--and, hence, one can be used as a representation of the other, even if they do not "look alike."

An analogue does not have the same structural properties as that which it represents, but it does have the same relevant functional properties. Hence analogues are functional models whereas iconic models are structural. It will be recalled that things with different structures can have the same function, (e.g., sundials and clocks).

Symbolic Models

Symbolic models are ones in which linguistic signs or symbols are used to represent the structural and/or functional properties of a situation. Thus a verbal description or explanation of a situation is a symbolic model of it. Such models are clearly the easiest to modify and manipulate but are the most general and abstract. Such models may range from purely qualitative verbal descriptions to precise quantitative models expressed in terms of abstract symbols such as are commonly used in science; for example, $s = 1/2\ gt^2$.

THE STRUCTURE OF MODELS

Models may represent objects, events, total situations, and their properties. Models of choice situations must have a certain kind of structure if they are to represent the essential characteristics of such situations. Choice models must express a relationship between (1) an outcome or some property of an outcome (e.g., its value to the subject), V, and those aspects of the situation over which the subject has some control (X_i), and (2) those aspects over which he has no control but which nevertheless he believes have some effect on the outcome (Y_k). Therefore, the form of such models can be represented symbolically as

$$V = f(X_i, \ Y_k),$$

where f is the relationship between V and X_i and Y_k. This relationship defines the E_{ij}, the efficiencies of the courses of action (defined by the X_i) for the possible outcomes under certain environmental conditions (defined by the Y_j).

In many situations the subject has only limited control over one or more of the controlled variables. For example, "the amount to be paid for a service" is usually such a variable; the subject cannot pay a negative amount for it, nor an amount greater than the total that he has available to him. This may be expressed as $0 \leq X_1 \leq A$, where X_1 is the amount to be spent and A is the total amount available to him.

Therefore, the model of a choice situation usually consists of an "objective function" which can be expressed in the form: $V = f(X_i, \ Y_k)$, and a set of limits or constraints over his control which can be expressed by a set of equations or inequations.

Of course, it is almost only in science (and only occasionally there) that models are explicitly expressed as equations and inequations.

But by an analysis of explicit models, however rare they are, we can uncover a number of important characteristics of models of choice situations however different their mode of expression may be.

Before proceeding to such an analysis let us relate what has just been said about models of choice situation to what was said earlier about such situations.

The controlled variables, X_i, define the alternative courses of action which the individual perceives, C_i. For example, if X_1 is the amount to be spent to acquire a service, then C_{11}, C_{12}, ... may represent spending \$0-\$5.00, \$5.01-\$10.00, ... for that service. The probability that a course of action will produce a certain perceived outcome, O_j depends on the values of the uncontrolled (state) variables, Y_k. That is, each believed E_{ij} can be conceived of as a function (g) of the Y_k. The believed value of an outcome O_j, V_j, may also be conceived as a function (h) of the state variables, Y_k. Thus, if the two functions, g and h, were known by the subject he could determine for each course of action the probability that it would produce each possible outcome in the choice situation. The measure of performance of the choice must itself be some function of the value of the outcome and its probability of occurrence; for example, the subjects expected utility: $\Sigma_j \ E_{ij} V_j$. This is only one of many possible performance functions. If he seeks to maximize this function then such maximization is his criterion of choice, his personality function in this context

IDENTIFYING POSSIBLE COURSES OF ACTION

Identification of possible courses of action is an essential part of constructing a model of a choice situation. Most "break-throughs" in problem solving are the result of finding either a new way of accomplishing an old objective, or a new outcome obtainable by use of a familiar course of action. The "newness" of these discovered alternatives implies that a creative act has occurred. In a sense, then, we are going to examine creativity in formulating models of choice situations.

Consider the following problem. An overly generous housewife returning from a shopping trip with a bag of apples meets a friend and gives her half of the apples plus half an apple. She later meets a second friend and gives her half of the remaining apples plus half an apple. The process continues through four friends after the last of which no apples remain. How many apples did she start out with?

A fairly obvious way of solving this problem (to those who have studied algebra) is as follows. Let X represent the initial number of apples. Then the amounts she gave to each friend in succession were:

$$a_1 = 1/2\,X + 1/2$$
$$a_2 = 1/2\,(X - a_1) + 1/2$$
$$a_3 = 1/2\,X - (a_1 + a_2) + 1/2$$
$$a_4 = 1/2\,X - (a_1 + a_2 + a_3) + 1/2$$

Then, $X - a_1 - a_2 - a_3 - a_4 = 0.$

One can proceed by substitution and get a cumbersome equation in terms of X and solve it.

Most who are given this problem proceed in the way described. Some, however, "see" the problem in a different way. They start at the other end. If the woman gave her last friend half of her apples plus half an apple and had nothing left, she must have had only one apple left after meeting her third friend. Then she must have had three apples left after her second friend, giving two to the second. She must have had seven apples after her first friend, of which she gave four to her second friend; and fifteen to start with of which she gave eight to her first friend.

This second procedure is one that most people do not "see" even though "it is there."

The perception of a new potential course of action is frequently attributed to the mental function called intuition, which is defined as follows by Webster's Collegiate Dictionary, (1937):

> Immediate apprehension or cognition; the power of knowing or the knowledge obtained without recourse to inference or reasoning; insight, familiarity, a quick or ready apprehension.

Many observers of intuition have noted that the process itself is not immediate, but consciousness of its output occurs suddenly. Poincaré and others have noted that they have lived with problems for long periods before having an insight which made possible their solution. An unconscious process may well have been going on for an extended period of time.

Webster's definition also asserts that intuition is not an inferential process such as reasoning or thinking is. My own reasoning leads to the contradictory conclusion.

First let me consider thought and intuition in a discursive way. When, by intuition, one perceives a possible course of action (e.g., a possible solution to a problem) it is not necessarily a good one. The output of either intuition or thought may be "good" or "bad," "right" or "wrong," "true" or "false." Therefore, the nature or quality of the output of thought and intuition does not differentiate between them. The difference lies in the processes, not their products.

Next, observe that once a suggestion has been put forth by intuition, it can often then be extracted by thought from what one knows about the situation. For example, when a theorem is suggested to a mathematician by his intuition, he can usually go back and derive it or show that it is not derivable from his premises. Thus intuition may produce belief that a theorem follows from certain axioms and postulates, and thought may prove that it does or it does not. In this sense, intuition is a kind of eliptical thought process; it appears to jump steps and proceeds from premises to conclusion without <u>consciously</u> going through the intermediate steps that thought goes through. Intuition does not consciously relate conclusions to premises; thought does.

Intuition frequently brings with the suggestions that it yields a strong belief in their validity. This belief may persist even when the suggestion is demonstrated to be inconsistent with one's accepted premises. In such cases it may lead to re-examination of one's premises and eventually to their modification. Thought can reveal which premise must be changed to make the intuitive suggestion derivable, but intuition usually provides the motivation to do so.

It is through this process that intuition suggests new ways of thinking about a situation. By calling assumptions (particularly implicit assumptions) into question it opens up new possibilities for thought. It is for this reason that intuition is so commonly associated with creativity: it suggests new ways of representing choice situations. Its output, however, may not be superior to that which it proposes to replace.

One can also use thought to develop new ways of representing a choice situation and, therefore, of revealing previously unperceived courses of action. Thought, for example, can be used to question systematically the validity of one's premises and to determine the consequences of such denials.

We have observed, then, that intuition may draw a conclusion from a set of premises without apparently going through the steps which link the conclusion to the premises. It may also perceive a conclusion that can only be drawn if the premises are modified.

In this discussion of thought and intuition we have been talking about the process of <u>inference</u>, which, according to Webster's, is "the act of passing from one judgment to another, or from a belief or cognition to a judgment," or it is "a logical conclusion from given data or premises."

Intuition, then, appears to be a mental leap over an inferential gap, whereas thought is associated with an orderly and logical construction of a bridge across that gap. It is clear, therefore, that an understanding of the difference between thought and intuition presupposes an understanding of inference.

Inference is a process by which "new" beliefs are produced by "old" ones. For example, if I believe

B_1 : my wife is at home

B_2 : the phone at home is operating

then I believe

B_3 : I can reach my wife at home by phone.

Note that B_1 and B_2 are each necessary, but neither is sufficient for B_3. Hence, B_1 and B_2 are producers of B_3.

6.1. <u>Inference:</u> the production of one or more beliefs or assumptions by one or more other beliefs or assumptions.

An inferential process is always about something: some class of objects, events, situations, or combinations of these. An inference about choice involves the <u>elements</u> we have already identified: controlled variables, uncontrolled variables, constraints, outcomes, and so on. These are the class of things that an individual believes are relevant to his choice. (Another individual may perceive a different set of elements in the same situation.) Therefore, the first part of a formalized inferential system is a set of elements which the subject believes are relevant; that is, objects, events, or combinations of these which the subject believes are producers of his future feelings, or signs or symbols of these.

The second part of a formalized inferential process is a set of beliefs concerning the form in which relevant beliefs can be represented; that is, the relevant form of predicational and relational statements (see Chap. 4) composed only of elements of the system. These beliefs constitute a set of belief-formation rules, or representations of these: statements or propositions.

Next, there is a set of beliefs and assumptions which the subject is willing initially to accept as true. These contain only elements of the system and are expressible consistently with the formation rules. These constitute the premises of the system. In a deductive system these may be axioms or postulates; in an inductive system these may be a set of accepted facts or observations.

Finally, there is a set of beliefs concerning how acceptable beliefs (other than those contained in the premises) may be derived from those which are accepted. These can be called transformation rules. For example, "If A is included in B, and B is included in C, then A is included in C" is such a rule. Applying it to accepted beliefs of the form "Cleveland is in Ohio and Ohio is included in the United States" one can conclude "Cleveland is included in the United States."

The formation and transformation rules are regularities in the subject's behavior which an observer can attribute to him even though the subject himself may not be aware of them. These rules are, in effect, the subject's program for deriving new beliefs from old. Such a program is functionally oriented; it is part of the subject's purposeful activity.

The inferential process may be either deductive or inductive. In a deductive process the premises are believed by the subject to be more general than the consequences derived from them. In an inductive process the premises are believed to be less general than the consequences. Therefore, inferences from what one believes to be laws to facts are deductive, and inferences from what are believed to be facts to laws or theory are inductive. Since beliefs in generality may differ, what appears to be deductive to one person may appear to be inductive to another.

Now let us return to the difference between thought and intuition.

6.2. Thought is conscious inference.

That is, if an individual employs an inferential process and is conscious of its parts--the elements, the formation rules, the premises, and the transformation rules and the way they are used--he can be said to be thinking.

If any of the premises in the inferential process are false, the conclusions reached may be also. Deficiencies in the rules may also produce false conclusions.

Thought, as conceived here, is a process by which an individual can proceed from a set of beliefs and assumptions to other beliefs that he can hold which may be either more or less general than what he started with.

Intuition is a process which accomplishes the same outcome as thought, but it is not a conscious process.

6.3. Intuition is unconscious inference.

A subject may be unconscious of any part or all of the inferential process that is employed. Thus thinking-intuition represent regions on the scale of consciousness that is involved in an inferential process. Few, if any, inferences are either pure thought or pure intuition. For example, the premises or the rules in a rigorous and conscious deduction may be suggested by an intuitive process. On the other hand, intuition is based on at least some premises which are consciously held. Little wonder, then, that one man can intuit what another man has reached by thought, or that one may reach by thought a result that another has intuited.

Intuition supplies many possible beliefs--hunches, conjectures, suggestions, and so on--which thought can be (but is not necessarily) used to <u>evaluate</u> systematically. Thought is an evaluative process in which the values involved are based on the true-false scale. Intuition does not evaluate, it <u>proposes</u>. Thought <u>proves</u>.

It might seem more reasonable to conceive of intuition not only as unconscious, but also as unprogrammed. I have not done so for several reasons. First, I shall want to distinguish later between intuition and <u>guessing</u>. A guess seems to me to be obtained both unconsciously and in an unprogrammed way. Secondly, and related to the first reason, is the fact that intuition often brings with it a degree of belief (conviction) that a guess does not. It seems reasonable to me to assume that this difference is due to the fact that intuition is a programmed process. Nevertheless, there are no serious consequences for this conceptual system if intuition is taken as unconscious and unprogrammed.

6.4. <u>Rationalization.</u> If a thought process is applied to a conclusion that was arrived at intuitively, and this is done with the intention of justifying the conclusion

> rather than determining whether it is justified, the process is that of <u>rationalization</u>.

In thought one determines whether or not a specified conclusion is justified by the premises and the rules. In rationalization the validity of the conclusion is accepted and an inferential system is sought which justifies it. Hence, rationalization may involve the search for premises or for rules which yield the desired outcome.

EVALUATING MODELS

As we have seen, models of choice situations are formed by combining relevant beliefs and assumptions which are supplied either by memory or by current observation. Once a model is constructed the individual may find it is lacking in any one of several different ways:

1. He may <u>doubt</u> that he has included all the relevant variables; either choice, environmental, or outcome.
2. He may <u>doubt</u> the relevance of one or more variables that he has included in his model.
3. He may <u>doubt</u> the validity of the relationship by which he has connected these variables.
4. He may <u>doubt</u> the accuracy of his estimates of the values of the variables which are incorporated into the model.

If he has any of these doubts and he has the resources and opportunity to inquire further, he will do so. (We will consider such inquiry in the next chapter). Sooner or later, however, he reaches the stage at which he feels either that he must make a choice, whatever his doubts, or that his model is adequate.

Rather than conduct inquiry designed to remove a specific doubt of any of the four types listed above, the individual may decide to evaluate his model as a whole. It is to this process that we now turn our attention.

A model of a choice situation is a forecasting instrument. It can be used to predict that <u>if</u> something is done, <u>then</u> something specific will happen. The predicted outcome may be obtained in either of two ways: (1) it may be inferred deductively from the model; that is, obtained by thought. In such deductions the model is an aggregation of the premises of the deductive process. The variables that it contains are the elements, and the method of deduction (e. g., algebra or the calculus) provides the formation and transformation rules. (2) The consequences may be intuited.

The consequences predicted may refer to a past, present, or future state. That is, the subject can put past values of the controlled and uncontrolled variables into his model and see if it predicts what did in fact happen. He can also predict the future, using present values of the variables, and determine whether these "come true."

The subject may either believe the consequences, disbelieve them, or he may be in doubt about them. If in doubt, and if he believes that he has the resources and capability of doing so, he will inquire into their validity. If he believes them, then his belief in the validity of the model is increased; if not, his belief is decreased. When his belief in the model is increased it may become acceptable to him or he may desire further confirmation and therefore infer or intuit consequences and continue the process until he accepts or rejects the model. When he rejects a model he must change one or more of its essential characteristics and start the evaluative procedure over again.

Thought and intuition are not the only processes by which conclusions and consequences can be reached. Both are "programmable" processes, one conscious, the other not. There are two complementary unprogrammed procedures, one conscious and one not. The process of guessing, which was mentioned above, I take to be unconscious and unprogrammed. To me, intuition appears to involve an implicit logic which can frequently be raised to consciousness, but a guess carries no such implication. Randomized selection of a conclusion or consequence seems to me to be a consciously unprogrammed procedure. In this case the lack of program is deliberate and is usually motivated by a desire to avoid any implicit or explicit bias in selection. Thus randomized selection is employed when we want to give each alternative an equal chance of being selected.

CONCLUSION

In this chapter we have considered how a model or concept of a choice situation is used to select a course of action or to initate an inquiry into some aspect of that model. A model is a representation of the choice situation, a representation in which some measure of performance is related to (1) those aspects of the situation over which an individual believes he has some control, and (2) those aspects which he believes to be relevant but uncontrollable. This representation is usually accompanied by others which reflect the limits within which the individual believes control can be exercised.

Courses of action are derived (inferred) from the model, that is, the beliefs incorporated in the model produce a belief as to which of the possible courses of action will yield the best performance. The derivation may be conscious and thus be obtained by thought, or it may be unconscious and hence be obtained by intuition. On the other hand a course of action may be selected by a guess or arbitrarily (as by a random choice). In guessing and arbitrary choice, inference is not involved.

Inference is a procedure by which a set of beliefs or assumptions in the form of premises, formation rules, and transformation rules, produce another set of beliefs or assumptions. The process is deductive if it proceeds from the general to the specific, and inductive if it proceeds from the specific to the general. An individual's inferential process is his logic. The components of an individual's logic have been defined here so as to make it possible to study such a logic behaviorally.

The mental functions, thought and intuition, will be considered again in Chapter 7 where their relationship to perception and feeling is examined in some detail.

If an individual has strong doubts about any aspect of his model of a choice situation and if he believes he can and ought to try to dispel these doubts, he will inquire further before making a choice.

BIBLIOGRAPHY

Dewey, John, Logic: The Theory of Inquiry, Henry Holt, New York, 1938.

Jung, C. G., Psychological Types, (trans. by H. G. Baynes), Pantheon Books, 1923.

CHAPTER 7

EVALUATION OF SITUATIONS: FEELINGS AND ATTITUDES

> OUTCOME, n. A particular type of disappointment. By the kind of intelligence that sees in an exception a proof of the rule the wisdom of an act is judged by the outcome, the result. This is immortal nonsense; the wisdom of an act is to be judged by the light the doer had when he performed it. (Ambrose Bierce, The Devil's Dictionary).

INTRODUCTION

When part or all of a situation is observed it is susceptible to evaluation by the observer. If the situation that is observed is a product of the observer's earlier action, then the outcome of that action can be evaluated. The essential characteristic of such evaluations is the decision to change or retain the situation and/or the observer's relation to it. hence, the value placed on that which is observed is the intention to change or retain it.

Evaluations are intentions. Intentions that are produced by that which is observed--that is, intention-responses--are feelings. Therefore, when we study intentions from the point of view of what produced them, we study feelings. Feelings are about something, they are about what produced them, what they are responses to.

The purpose in this discussion is to show that feelings can be adequately treated in a system of objective teleological concepts. All possible feelings cannot be dealt with here; more than a thousand of them have been identified by Orth (in Reymert, 1928, p. 375); but several are dealt with here to show how they can be treated within the system. The particular connotations which are attributed to them here are clearly debatable because there is little agreement and precision in discussions on the nature of specific feelings. I cannot hope to resolve wide differences of opinion as to what a particular feeling connotes However, by formulating at least one possible connotation of each of several particular feelings, I do hope to show how a range of feeling connotations can be treated. To assist the reader in these exercises, I will quote definitions drawn from the fifth edition of Webster's Collegiate Dictionary.

I have already mentioned the dependence of the concept of feeling on that of intention. It will also become apparent that the concept of belief--particularly "expectation"--plays a very important role in defining particular feelings. Intention was discussed in Chapter 3 and belief in Chapter 5.

FEELING, EMOTION AND SATISFACTION

According to Webster's to disappoint is "to fail to come up to the expectation of." Using the concepts previously developed we can construct the following definition:

7.1. Disappoint. An individual (A) is disappointed, if an object, event, or situation (X) desired by A, which he believed would be present or occur at time t, does not appear or occur at t.

An individual (A) is disappointed with another entity (B) if A believed B was capable of producing and would produce the desired X by time t, and A believes that B did not do so.

Whether or not a subject is also dissatisfied by the nonoccurrence of X at t depends on what he intends to do about the failure of X to occur.

To "satisfy" according to Webster's is "to fill up the measure of a want of (a person or thing); hence to gratify fully the desire of..." Attainment of a desired outcome (i.e., an objective) brings satisfaction. To be completely satisfied is to want nothing other than what one has; to be completely dissatisfied is to want nothing that one has.

> 7.2. Satisfaction. An individual's degree of satisfaction with an object, event, property or properties of either, or a state, X, is his degree of intention to produce a non-change in X.

For example, if an individual is in a particular environment, S, and he is presented with two exclusive and exhaustive classes of courses of action, members of one of which will change the environment and members of the other will not, and the other conditions of an intention environment are met; then the probability that he will select a course of action that will not change S is his degree of satisfaction with S. The probability that he will select the course of action that will change the environment is his degree of dissatisfaction with S. If the former probability is greater than the latter he is said to be satisfied. If the latter is the greater, he is dissatisfied. If these are equal, he is indifferent to the situation and can be said to have no feelings about it.

> 7.3. Feeling. To have a feeling is to be in a state of satisfaction or dissatisfaction.

A feeling is a functional property of an individual. It is an intention to change or retain something, an intention produced by that something.

Particular feelings (e.g., fear) may frequently be accompanied by certain changes in the individual's structural properties (e.g., accelerated heart beat, perspiration, trembling, etc.). Those structural changes which occur in association with a feeling can be called emotions. It is this sense that emotions have been said by some to fall in the domain of physiology, whereas feelings fall in the domain of psychology. C. Lange put it this way more than a century ago:

> If from one terrified the accompanying bodily symptoms are removed, the pulse permitted to beat quietly, the glance to become firm, the color natural, the movements rapid and secure, the speech strong, the thoughts clear, --what is there left of his terror? (Om Sindsbevaegelser, Kobenhavn, 1855, trans. from the German translation of H. Kurella by B. Rand in The Classical Psychologists, London, 1912, p. 675).

Many concepts of the relationship between feeling and emotion different from the one suggested here have been proposed. (These have been extensively surveyed and analyzed by Hillman, 1964.) However, to me there is an attractive symmetry in the treatment of feeling as a functional response to a situation and emotion as an associated structural response. Hence, feeling and emotion are the head and tail of the same coin; two different ways of looking at the same thing.

When an individual is confronted with a situation, whether it is a product of his previous choice or not, and he is dissatisfied with it, he intends to change it. Whether he tries to do so depends on his appraisal of other aspects of the situation; for example, the availability of means for so doing. To say he intends to change the situation is to make an assertion not about what he does but about what he would do under certain idealized conditions of choice required in an "intention environment" (discussed in Chapter 3). He may not act in accordance with his intentions because of the deviations of the actual situation from the intention environment.

Some Specific Feelings

When an individual selects a course of action he may believe that a certain outcome will occur (his expectation). As indicated above, if this outcome is desired and it does not occur, he is disappointed. His feelings, however, may go beyond this.

> 7.4. Regret. An individual regrets his earlier choice of a course of action if he believes that it was a producer of an unintended outcome with which he is dissatisfied.

Put another way, regret is dissatisfaction with a previous choice. This reflects Webster's definition of regret as "To have distress of mind or misgivings concerning;... as, to regret one's past mistakes."

One can obtain a measure of an individual's regret as a function of (1) the measure of his belief that his choice produced the unintended outcome and the degree of his dissatisfaction with that outcome. If either or both of these measures (both of which range from 0 to 1) are at zero, he has no regret; if both are at their maximum value, 1, he has maximum regret. Therefore, the measure of regret can be taken to be the product of the relevant measure of belief and degree of dissatisfaction.

Curiously there does not seem to be a term uniquely applicable to the contrary of regret: belief that a previous choice was a producer of an intended outcome that brings satisfaction. This is a type of self-satisfaction, but "self-satisfaction" connotes more than this.

Whereas regret refers to dissatisfaction with things past, "hopelessness" and "despair" refer to dissatisfaction with things anticipated. "Hope," according to Webster's, is "desire with expectation of obtaining what is desired ..."

> 7.5. Hope. An individual is hopeful if he is satisfied with

what he believes will occur. If he is dissatisfied with his expectation he feels hopelessness or despair.

The measure of hope can be defined as the product of the measure of belief in the future occurrence of a desired state and the degree of satisfaction with that state. Correspondingly, the measure of hopelessness is the measure of belief that a desired state will not occur and the dissatisfaction with what is expected. Hence, either measure is one minus the other.

An individual who, in general, tends to be hopeful rather than hopeless is referred to as an optimist; one who tends to be hopeless rather than hopeful, is referred to as a pessimist. Optimism and pessimism, as we shall see, are attitudes toward the future. Attitudes are discussed below.

Confronted with the possibility of a desired or an undesired outcome an individual may have no basis for expecting one rather than the other; that is, he doesn't "know" what will happen. In such a case he may be anxious, which, according to Webster's is to be "concerned, or solicitous as to something future or unknown."

7.6. Anxiety. When an individual believes that any one of two or more outcomes, some desired and some not, are likely to occur and he is dissatisfied with that state (of uncertainty) he can be said to be anxious or to display anxiety.

The measure of anxiety, therefore, is the measure of dissatisfaction with a state in which the measures of belief associated with the possible outcomes are virtually equal (and possibly are all zero). As belief increases in any one outcome, anxiety converts to hope or despair.

Hope, despair, and anxiety are a trilogy of feelings which reflect an individual's intention-response to what he believes or does not believe about the future. If an individual believes he can prevent an undesirable outcome, he has some hope, if not he may be frustrated.

> 7.7. Frustration. When an individual has no hope of obtaining a desired outcome, and he believes it is possible to produce that outcome but that he cannot do so, he feels frustrated.

Hence frustration involves both despair and dissatisfaction with oneself; holding oneself responsible for an undesirable expectation, at least in part.

Fear has been one of the most discussed and least agreed upon feelings. This is reflected in the fact that most dictionaries define it by use of such synonyms as "dread" and "disquiet." It seems to me that fear involves dissatisfaction with expected harm to oneself, physical or psychological; that is, reduction of one's capabilities for pursuing one's objectives in the future.

> 7.8. Fear. When an individual believes that something will occur which will reduce his ability to pursue his objectives in the future, his dissatisfaction with this state is fear.

The harm anticipated may restrict either his ability to choose efficient courses of action, or his ability to desire. Expectation of harm is not sufficient for fear. Witness the masochist. Dissatisfaction is also required.

> 7.9. Inhibition. When fear of one or more expected consequences of a course of action, other expected consequences of which are desired, produce a non-choice of that course of action in an individual, he feels inhibited.

Thus inhibition is a felt constraint on choice produced by fear of undesirable consequences. The choice may or may not be made. If it is, the fear, though not the inhibition, may remain.

Now let us consider a sample of feelings which involve the relationship between two individuals.

According to Webster's, to blame is "to find fault with." However, I may find fault with a book but not blame it. I might blame its author for the book.

> 7.10. Blame. One individual (A) blames another (B) for something (X), if A believes B intentionally produced X and A is dissatisfied with X.

Note that B need not be a person, but it does have to be an entity to which A attributes intentions, and hence choice.

The contrary of blame seems to me to be gratitude:

> 7.11. Gratitude. One individual (A) is grateful to another (B) for something (X), if A believes B intentionally produced X and A is satisfied with X.

The measures of gratitude and blame are also products of measures of belief and satisfaction or dissatisfaction.

Sympathy, according to Webster's Collegiate Dictionary, is "an affinity, association, or relationship between things so that whatever affects one, similarly affects the other or others." If this were taken literally it would be possible for two persons who did not know each other to be sympathetic with each other if they responded similarly to similar stimuli. To me this does not seem consistent with common usage in which sympathy seems to connote that the response of one individual produces a similar response in another.

7.12. Sympathy. One individual (A) sympathizes with (B) relative to something (X) if B's satisfaction (or dissatisfaction) with X produces satisfaction (or dissatisfaction) with X in A.

This definition permits A to sympathize with B without B sympathizing with A, and sympathy does not seem to be symmetrical to me despite the common "sym." Note that this definition implies that if A sympathizes with B, A is aware of X and conscious of B. (See definitions of awareness and consciousness in Chapter 4.)

If B's feelings about an X fail to produce any feeling in A, A is unsympathetic with B. If B's feelings produce contrary feelings in A, A might be said to be "antisympathetic" with B.

According to Webster's to appreciate is "to approve of; to be grateful for," and to be grateful is "to be appreciative of benefits received." Gratitude, it seems to me, is directed to a responsible person for something done. Appreciation is gratitude for the person, not only for what he has done but also for what he can do.

7.13. Appreciation. One individual (A) appreciates another (B) if A believes B is capable of producing satisfaction in A (i.e., fulfilling some of A's objectives.)

Appreciation is "passive," but devotion is "active."

7.14. Devotion. A is devoted to B if A is dissatisfied with B's states of dissatisfaction and satisfied with B's states of satisfaction.

Therefore, if A is devoted to B, A intends to remove B's dissatisfactions and preserve his satisfactions. Note that devotion presupposes sympathy but also involves an intention to do something about it.

The contrary to devotion is antagonism, the desire to preserve another's states of dissatisfaction and remove his states of satisfaction. There is no convenient anyonym for appreciation, therefore, I shall use "disappreciation" to represent A's belief that B is capable of producing dissatisfaction in A.

Now let me briefly discuss two feelings on which even angels fear to tread, love and loyalty. The meanings of these concepts are much too vague and rich to hope for any agreement among those who have tried to analyze their meanings. However, following the analysis given by E. A. Singer (1923) in his essay, "Royce on Love and Loyalty" the following definitions were suggested to me.

7.15. Love. A loves B if A appreciates and is devoted to B.

7.16. Hate. A hates B if A disappreciates and is antagonistic to B.

7.17. Loyalty. A is loyal to B if A disappreciates and is devoted to B.

I cannot find a suitable term for the fourth possibility--A appreciates and is antagonistic to B--although resentment seems to come close to it.

These definitions suggest why loyalty can be "demanded" or enforced but love can't be. Appreciation cannot be legislated, but "devotion" can be. As Singer has observed, there would be no such thing as a demand for loyalty were there no call for a man to sacrifice his purpose for another's. The "other," of course, may be a group as well as an individual; for example, a nation; a school, a community, and so on.

I hope I have gone far enough and deeply enough into these few feelings to show how they can be incorporated into a system of objective teleological concepts.

Now let us turn to an aspect of feeling that I have referred to earlier in this chapter as an "attitude."

ATTITUDES

Webster's defines an attitude as a "position or bearing as indicating action, feeling, or mood; as, keep a firm attitude; the feeling or mood itself; as, a kindly attitude." In 1929 Thurstone and Chave offered the following often cited definition of an attitude: "the sum-total of a man's inclinations and feelings, prejudice or bias, preconceived notions, ideas, fears, threats, and convictions about any topic" (pp. 6-7).

In an earlier book I showed some of the consequences of taking this definition literally:

> First it would be necessary to define the relevant sets (populations) of (a) inclinations, (b) feelings, (c) prejudices or biases, (d) preconceived notions, (e) ideas (f) fears, (g) threats, and (h) convictions. Then either a complete count of each population or a probability sample would be required. An estimate of the "sum" of these would have to be made assuming the research obtained comparable measures which could be summed. As a matter of fact, none of this is done or tried, and for obvious reasons. For example, the test items are actually selected because they seem pertinent and not because they can be demonstrated to be so on the basis of the definition. Furthermore, the items provide no identifiable measure. In the test for "measuring" attitudes toward the church (Thurstone and Chave, 1930) for example, such items as the following can be found:
>
> > "I regard the church as a monument to human ignorance."
> >
> > "I feel the church is the greatest agency for uplift of the world."
>
> The subject is instructed to check those statements with which he fully agrees. Such a check or lack of it may seem to provide information concerning an attitude as defined above, but no demonstration that this is the case has been provided. The definition does not make it easy to do so. (Ackoff, 1953, pp. 305-306).

A very extensive examination and analysis of psychological definitions of attitude, including that of Thurstone and Chave, was made by Sherif and Cantril (1945). This effort yielded four properties which, they asserted, a definition of attitude should reflect:

(1) "Attitudes are always related to defined stimuli or stimulus situations" (p. 301).

(2) "Attitudes are formed" (p. 301).

(3) "Established attitudes are charged with affective or value properties in varying degrees" (p. 302).

(4) "Attitudes are more or less enduring states of readiness [for action]" (p. 303).

The definition that is developed here satisfied these conditions:

7.18. Attitude. An attitude is a feeling about something that persists over time and a variety of environments.

7.19. Mood. A mood is a feeling that is relatively short-lived that sweeps in everything or most things experienced during that period.

Thus an attitude is a directed feeling, one that is produced by its object, such as an attitude toward a particular person, organization, or event. Hence one individual can have, for example, a hostile attitude toward another and it will persist over time and manifest itself in different environments. On the other hand, a person who is in a hostile mood directs this feeling at all or most persons with whom he interacts during the life of the mood.

An attitude is an intention-set, a feeling posture toward its object. It is a characteristic intention-response pattern to a specific stimulus. Hence, attitude is to intention what trait is to familiarity.

Both are patterns of response to stimuli. One would hardly extract this relationship from an examination of the previously quoted definition of attitude given by Thurstone and Chave and that of a trait given by G. Allport (1937): "a generalized and focalized neuropsychic system (peculiar to the individual), with the capacity to render many stimuli functionally equivalent, and to initiate and guide consistent (equivalent) forms of adaptive and expressive behavior" (p. 295). Nevertheless, the simularity between traits and attitudes have been recognized often and results in their frequent association in the psychological literature.

Since an attitude is a feeling, it involves satisfaction or dissatisfaction and hence lends itself to such dichotomous characterization as favorable-unfavorable, for-against, like-dislike, and so on. These dichotomies sometimes obscure the fact that there is an underlying scale of satisfaction (i.e., of intention) that ranges from 0 to 1.

Now let us examine some of the items on the Thurstone-Chave test for attitudes toward the church in the light of this discussion. There are forty-five items in this test. The subjects are instructed to check those items with which they "fully-agree."

An examination of these items reveals that "church" is used ambiguously throughout. In some of the items the individual is asked to respond to <u>religion</u> in the most general sense and in others to a specific denomination or even a <u>specific building</u>. For example, "church" is used in a very general sense in the following items:

"4. I regard the church as a monument to human ignorance.

5. I believe that the church is losing ground as education advances.

6. I feel the church is trying to adjust itself to a scientific world and deserves support."

Some items in which "church" is used in a much narrower sense are:

"21. My church is the primary guiding influence of my life.

31. There is much wrong in my church, but I feel it is so important that it is my duty to help improve it.

36. In the church I find my best companions and express my best self."

Because of the ambiguous treatment of "church" in the test it is not at all clear what it measures an attitude toward: a specific congregation; a denomination; a religion, or religion in the general sense.

There are other difficulties. Consider item 39: "It seems absurd to me for a thinking man to be interested in the church" is a statement which specifies only a certain aspect of intention relating to the church. Considered as an agent of emotional uplift, or as a center of social activity, rather than an agent of thinking, preservation of the church might be an end of high intention. A religious man may deny this statement because he does not find the church interfering with his thought and it provides him with "religious uplift." On the other hand, a sociologist who is an ardent atheist might agree with the statement because he considers the church as a social institution, rather than a religious one, for he finds it cannot be ignored by a thinking man who would completely understand a culture. Is "interested in" meant to imply "participate in"?

Many of the items of the test do not seem to be designed to elicit the same expression of belief in intention over a variety of people. For example, consider 34: "I feel that church attendance is a good index of the natinn's morality." If a person felt that the nation was immoral and church attendance was low, then he might very well agree with Thurstone. However, a person who feels the church is immoral (as

Lenin did) and that the nation's morality is low, would also agree with this item. Clearly, the attitudes of the two persons are not the same.

It should be noted that a verbal test of an attitude does not tell us what a person wants, only what he says he wants. These are not necessarily the same thing. Hence, unless a verbal test of an attitude is validated experimentally against relevant behavior its basic assumption of the equivalence of what a person feels and what he says he feels, is not justified.

Finally, it is not at all apparent, even if we assume the identity of what a person says and feels, that answers to these items are evidence from which intentions can be inferred. No explicit criterion of relevance of these items was used in selecting them; all that was required was agreement among independent judges who were given no criterion to use in their judgments.

MENTAL FUNCTIONS AND CHOICE

Once a model is accepted a choice of a course of action can be made. The process of selection was considered in Chapter 3, but now we can focus more clearly on one aspect of it by use of the concepts that have been developed since then. Intuition suggests possible courses of action which can be evaluated by use of the choice model and the process of thought. The model itself is the product of past and present observations, or sensations. The consequences predicted are evaluated by feeling. The course of action which is predicted to yield the most desirable outcome is selected.

It is apparent that thought, intuition, sensation, and feeling are all involved in choice. These are what C. G. Jung (1923) considered to be "the four basic psychological functions." The correspondence of the meanings that Jung attributed to these concepts with the meanings attributed here is not accidental since my thinking has been considerably influenced by his.

Jung's Psychological Functions

For Jung,

> Thinking is that psychological function which, in accordance with its own laws, brings given presentations into conceptual connection... The term 'thinking' should in my view be confined to the linking up of representations by means of a concept... (p. 611).

Hence, for Jung, thinking relates representations. However, he did not discuss the process of relating them as I have tried to do in the discussion of inference in Chapter 6.

Intuitition for Jung,

> is that psychological function which transmits perceptions in an unconscious way... Through intuition any one content is presented as a complete whole, with or without our being able to explain or discover in what way this content has been arrived at.. Its contents... have the character of being given, in contrast to the 'derived' or 'deduced' character of feeling and thinking contents (pp. 567-568).

"Sensation is sense-perception, i.e., perception transmitted via the sense organs and 'bodily senses'... (p. 586)." With this much my treatment of observation in Chapter 4 agrees. Jung went on, however, to assert that he regards "sensation as conscious, and intuition as unconscious perception (p. 587)." My treatment of sensation does not require that it be conscious. Unfortunately, Jung did not define "perception" but if one can say that possible courses of action can be perceived, then my treatment of intuition is in essential agreement with his. Finally,

> Feeling is primarily a process that takes place between the ego and a given content, a process, moreover, that imparts to the content a definite value in the sense of acceptance or rejection ('like' or 'dislike'); but it can also appear, as it were, isolated in the form of 'mood', quite apart from the momentary contents of consciousness or momentary sensations (p. 543).

> ...feeling is also a kind of judging, differing, however, from an intellectual judgment, in that it does not aim at establishing an intellectual connection (as thought does) but is solely concerned with the setting up of a subjective criterion of acceptance or rejection (p. 544).

Thus he contrasted feeling with thought, both of which are judgmental or evaluative, on the basis of the criteria that they employ. Feeling uses such a "subjective" criterion as "like-dislike" and thought uses such an "objective" criterion as "true-false." Because these functions are both judgmental, Jung calls them "rational," whereas he considered sensation and intuition to be "irrational" because they involve perceptions rather than judgments. Unfortunately again, Jung did not define "judgment." If it means the act of acceptance or rejection by use of a criterion, then my treatment of thought and feeling is consistent with his.

Jung did not consider either the exclusiveness or exhaustiveness of these four functions. It is apparent in his writings that he considered them to be at least exhaustive (which I have not), but he did not argue to this effect. One might set up a table of his concepts as is done in Table 8.1.

The difficulty with this table lies in the distinction between columns. Although Jung differentiated between sensation and intuition by use of consciousness and unconsciousness, he did not so differentiate between thinking and feeling. Furthermore, although he wrote that feeling is "entirely subjective," he did not characterize thinking and "entirely objective," and did not seem to imply that it is. Therefore, one cannot extract an argument for the exhaustiveness of these functions from his writings.

<u>Dewey's Pattern of Inquiry</u>

The role of the four mental functions in making a choice is

TABLE 8.1

JUNG'S PSYCHOLOGICAL FUNCTIONS

	CONSCIOUS	UNCONSCIOUS
IRRATIONAL (Perception)	Sensation	Intuition
RATIONAL (judgmental)	Thinking	Feeling
	OBJECTIVE	SUBJECTIVE

greatly illuminated in the work of John Dewey (1938) even though he made no explicit reference to Jung or his thought. A review of Dewey's concept of "the pattern of inquiry" not only illuminates these functions but it provides an opportunity for restating in a different way some of the critical aspects of the conceptual system that I have been constructing here.

Dewey considered five aspects of inquiry. I shall let him describe them for himself and comment around his discourse.

> I. The Antecedent Conditions of Inquiry: The Indeterminate Situation. ...it is of the very nature of the indeterminate situation which involves inquiry to be questionable... The peculiar quality of what pervades the given materials, constituting them a situation, is not just uncertainty at large; it is a unique doubtfulness which makes that situation to be just and only the situation it is (p. 105).

The position taken here by Dewey is equivalent to mine: that a choice situation is a necessary antecedent of a problem. In Chapter 2 I tried to make explicit what such a situation consists of. Like Dewey, I pointed out that a choice situation becomes a problem situation only if the situation produces a state of dissatisfaction in the subject (a feeling) and he is doubtful about what to do. I also pointed out that the "existential situation"--to use Dewey's term--is never known in all its detail by either the subject or the one who observes him. Different subjects and different observers may--and usually do--"see" the situation differently.

> II. Institution of a Problem. ...The indeterminate situation comes into existence from existential causes...There is nothing intellectual or cognitive in the existence of such situations, although they are the necessary conditions of cognitive operations or inquiry. In themselves they are precognitive. The first result of evocation of inquiry is that the situation is taken, adjudged, to be problematic. To see that a situation requires inquiry is the initial step in inquiry (p. 107).

Unless the subject responds to the possibility of choice--and hence is aware of it--a problem cannot arise. For Dewey and for me this awareness and the state of doubt produced by it are necessary before the individual can be said to have a problem.

> III. The Determination of a Problem Situation. ...The first step...is to search out the constituents of a given situation which, as constituents, are settled...All of these observed conditions taken together constitute "the facts of the case"... they are conditions that must be reckoned with or taken account of in any relevant solution that is proposed (pp. 108-109).

The role of observation, and hence sensation, is made explicit here. It provides information and hence affects possible choices (see Chapter 8). The contributions of the senses, present and past, when believed or assumed become the raw material out of which a model of the choice situation is constructed.

> A possible relevant solution is then suggested by the determination of factual conditions which are secured by observation. The possible solution presents itself, therefore, as an idea, just as terms of the problem (which are facts) are instituted by observation. Observation of facts and suggested meanings or ideas arise and develop in correspondence with each other.
>
> ...suggestions just spring up, flash upon us, occur to us... Every idea originates as a suggestion but not every suggestion is an idea. The suggestion becomes an idea when it is examined

> with reference to its functional fitness; its capacity as a means of resolving the given situation (pp. 109-110).

A solution is a course of action and a course of action can be defined by a set of values of the controlled variables. The perception of a possible course of action, when it just "pops up," is a product of intuition. Not all suggestions, however, are intuitive; they can also be the result of thinking the situation over, of deriving them from what is known or believed about the situation.

Note also that for Dewey a suggestion becomes an idea only when it is evaluated by a thought process employing the inputs of observation and feeling. Evaluation here means predicting whether or not a suggested course of action will produce a desired outcome in the situation involved.

> This examination takes the form of reasoning... But the final test of its possession of these properties is determined when it (the suggestion) actually functions--that is, when it is put into operation so as to institute by means of observations facts not previously observed, and is then used to organize them with other facts into a coherent whole (p. 110).

Possible courses of action can be evaluated either by predicting their consequences using what is believed about the situation (a thought process), or by trying them and observing the consequences and evaluating them (feeling). These are not exclusive processes. Every evaluation of a possible course of action involves all the psychological functions, but the one that dominates may differ from evaluation to evaluation, or from evaluator to evaluator depending on what, according to Jung, is his psychological type.

> Because suggestions and ideas are that which are not present in given existence, the meanings which they involve must be embodied in some symbols. Without some kind of symbol no idea; a meaning that is completely disembodied can not be entertained or used... To "look at an idea" is not a mere literary figure of speech (p. 110).

Thinking is an operation and what it operates on are concepts, images, and signs of these. Meanings are contained in beliefs about consequences of courses of action that can be taken, and beliefs are expressible in sign complexes that are statements or propositions. Thus, thought presupposes at least a private (if not a public) language. Thought involves the manipulation of signs. Hence, it is only by communication with himself that a subject can think, let along solve a problem. Thinking involves a conversation with oneself. Therefore, <u>not only do we need to understand choice to understand communication, but we also need to understand communication in order to understand choice and the psychological processes of which it is composed.</u>

> IV. <u>Reasoning</u>. ...developing the meaning of ideas in their relation to one another...operating with symbols (constituting propositions) is reasoning... This examination (of meaning) consists in noting what the meaning in question implies in relation to other meanings in the system of which it is a member, the formulated relation constituting a proposition. If such and such a relation of meanings is accepted, then we are commited to such and such relations of meanings because of their membership in the same system. Through a series of intermediate meanings, a meaning is finally reached which is more clearly <u>relevant</u> to the problem in hand than the originally suggested idea. It indicates operations which can be performed to test its applicability, whereas the original idea is usually too vague to determine crucial operations (pp. 111-112).

Thought relates concepts, images, and beliefs. In my earlier discussion of thought I tried to identify the components of this process and to describe the process itself. The product of the process is either a conclusion that is believed and hence becomes a basis for selecting a course of action, or a conclusion that can be tested (i.e., tried, observed, and evaluated).

> V. <u>The Operational Character of Facts-Meanings</u>. ...Ideas are operational in that they instigate and direct further operations of observation; they are proposals and plans for acting upon existing conditions to bring new facts to light and

to organize all the selected facts into a coherent whole.

> What is meant by calling facts operational? Upon the negative side what is meant is that they are not self-sufficient and complete in themselves. They are selected and described... for a purpose, namely, statement of the problem involved in such a way that its material both indicates a meaning relevant to resolution of the difficulty and serves to tests its worth and validity (pp. 112-113).

Here Dewey emphasizes the interaction and cycling of the various phases of the choice process, a characteristic of the process that was discussed in Chapter 2. The process has no fixed sequence of a fixed number of steps. One choice situation (and hence problem) arises out of another in a continuing stream. Several problems may co-exist and interact. Hence, the process of choice is very rich; it can be infinitely varied. It is a process in which each step can feed back to every other. Little wonder then that it is so seldom carried out in a completely efficient manner or that we do not know what the "most efficient manner" is. Nevertheless, the combined efforts of science and philosophy have made it possible for us to become <u>more</u> efficient in the making of choices. The compilation ~~of our~~ knowledge on this subject constitutes the field we have come to call <u>methodology</u>. Methodology is inquiry into the process of inquiry, the process of making choice, the purposeful pursuit of objectives.

CONCLUSION

Every problem begins and ends with an evaluation of a situation. Without dissatisfaction there can be no problem; but dissatisfaction is only a necessary, not a sufficient, condition for a problem. Unless choice is possible and there is a state of doubt about what to choose, a problem does not exist.

A problem does not cease to exist or is not solved until satisfaction replaces dissatisfaction. Hence feelings initiate and terminate every problem. Furthermore, the decision maker's values influence every step in the process of selecting a course of action. For discussion of how the inquirers' values enter into every decision made in the inquiry as well as in his ultimate choice, see Churchman (1961) and Ackoff (1962).

Feeling is one of the four mental functions out of which the choice process emerges; the others being sensation, intuition, and thought. The discussion of Dewey's pattern of inquiry revealed that thought is a type of communication with oneself. From this the following conclusion was drawn: not only do we need to understand choice to understand communication, but we also need to understand communication in order to understand choice and the psychological processes of which it is composed. Hence the next part of this book takes up the subject of communication in the context of choice.

BIBLIOGRAPHY

Ackoff, R. L., The Design of Social Research, University of Chicago Press, Chicago, 1953.

______, Scientific Method: Optimizing Applied Research Decisions, John Wiley and Sons, New York, 1962.

Allport, G. W., Personality: A Psychological Interpretation, Henry Holt & Co., New York, 1937.

Churchman, C. W., Prediction and Optimal Decisions, Prentice-Hall, Englewood Cliffs, New Jersey, 1961.

Dewey, John, Logic: The Theory of Inquiry, Henry Holt, New York, 1938.

Hillman, James, Emotion, Northwestern University Press, Evanston, Illinois, 1964.

Jung, C. G., Psychological Types, (trans. by H. G. Baynes), Pantheon Books, 1923.

Reymert, M. (ed.), The Wittenburg Symposium, Clark University, Worcester, Massachusetts, 1928.

Singer, E. A., Jr., On the Contented Life, Henry Holt & Co., New York, 1923.

Thurstone, L. L. and Chave, E. J., The Measurement of Attitudes, University of Chicago Press, Chicago, 1929.

______, A Scale for Measuring Attitudes Toward the Church, University of Chicago Press, Chicago, 1930.

CHAPTER 8

MODES AND MEASURES OF COMMUNICATION

TALK, _v_. _t_. To commit an indiscretion without temptation, from an impulse without purpose (Ambrose Bierce, _The Devil's Dictionary_).

INTRODUCTION

The significance of Claude Shannon's work in communication theory is such that anyone presuming to contribute to this theory is obliged to relate his work to Shannon's. In exploring this relationship it will be helpful to refer to Warren Weaver's masterful non-technical exposition of Shannon's work (Shannon and Weaver, 1949).

According to Weaver, "Relative to the broad subject of communication, there seems to be problems at three levels." These are

Level A. How accurately can the symbols of communication be transmitted? [The technical problem.]

Level B. How precisely do the transmitted symbols convey the desired meaning? [The semantic problem.]

Level C. How effectively does the received meaning affect conduct in the desired way? [The effectiveness problem.] (pp. 95-96)

Weaver classifies Shannon's work as follows:

The mathematical theory of the engineering aspects of communication, as developed chiefly by Claude Shannon at the

> Bell Telephone Laboratories, admittedly applies in the first instance only to problem A, namely the technical problem of accuracy of transference of various types of signals from sender to receiver (p. 97).

He goes on to note, however, that "the theory of Level A is, at least to a significant degree, also a theory of levels B and C" (p. 98). He does not make clear, however, exactly how this is so.

The effort in this chapter is primarily concerned with level C, the effectiveness problem. In the next chapter, we shall consider level B as well as C. But the effectiveness problem is conceived here in more general terms than those in which it appears to have been conceived in Weaver's formulation. My effort has the following objectives:

1. To identify the ways in which a receiver's behavior can be affected by a sender.
2. To construct measures of these effects.
3. To define and construct measures of the value of these effects for the receiver and for the sender and third parties, as well.

The question, "What is communication?" is treated in more detail here than it is by Shannon and Weaver. A related question, "How does one measure the amount of information transmitted?" is as critical here as it is in Shannon's theory. But I give "information" a considerably different meaning than Shannon did. According to Weaver (1949),

> The word <u>information</u>, in this [Shannon's] theory, is used in a special sense that must not be confused with its ordinary usage. In particular, <u>information</u> must not be confused with meaning (p. 99).

In my treatment, information and meaning will be closely related, and information will be conceived in a way that comes close to the way it is ordinarily used.

The meaningfulness and value of information is central in this discussion. "Information," according to Weaver, "is a measure of one's freedom of choice when one selects a message" (p. 100). Here I shall develop a concept of information in which the concept of "choice" is also fundamental, but here the choice is not related to messages but to courses of action. For reasons which shall be made apparent in the next chapter, Shannon's concept of information can be referred to as syntactic, whereas the one developed here is pragmatic.

Weaver defines communication as "all of the procedures by which one mind may affect another" (p. 95). His and Shannon's discussion, however, is restricted to only one such type of procedure: the transmission of messages. Their use of the term "communication" conforms better with common usage than does their definition. For example, the man who produced the slide rule I use may affect my mental processes without communicating to me. In general, many who have shaped my environment or the instruments which I use have affected my mental processes without communication in the ordinary sense.

If communication is to be restricted to the transmission of messages, the concept "message" must be clarified. This will be done in Chapter 9. First, however, "the effect of one mind on another" must be translated into behavioral terms. This can be done by use of the concept of a purposeful state, and its parameters, which were discussed in Chapter 2.

8.1. Communication. One purposeful individual (B) communicates to another (A) when a message produced by B produces a change in one or more of the parameters (P_i, E_{ij}, V_j) of A's purposeful state. B can be referred to as the sender and A as the receiver.

Several aspects of this definition of communication should be noted. First, A and B may be the same individual; that is, a person may communicate to himself as in writing a "reminder" to himself. Secondly, the sender of the message need not intend or desire to communicate to the receiver in order to do so. An interceptor of a message, for example, may be communicated to, although unintentionally. Thirdly, the sender and receiver may be widely separated in time and space. Through their writings both Aristotle and Nehru have communicated to, though not with, me.

Finally, note that both parties in communication must be purposeful. If we push a button to start a machine and the machine has no choice, communication has not taken place. On the other hand, if we push a button at the front door of a house, though we do not communicate with the bell, we do so with the occupants of the house; both they and we have alternative ways of pursuing our objectives.

Now we want to concentrate on the communication received and the receiver.

THE VALUE OF A COMMUNICATION

It will be recalled that a purposeful state of an individual ($\underline{A}$) is described by

(1) the set of available coursed of action, $\underline{C}_i$,

(2) the set of possible outcomes, $\underline{O}_j$,

(3) the environment, $\underline{S}$,

(4) the probabilities of $\underline{A}$ selecting each course of action, $\underline{P}_i$,

(5) the efficiencies of the courses of action for each objective, $\underline{E}_{ij}$, and

(6) the value of the outcomes to $\underline{A}$, $\underline{V}_j$.

Then, given the available courses of action and possible outcomes, the value of a purposeful state, $\underline{V}*$, must be some function of $\underline{P}_i$, $\underline{E}_{ij}$, and $\underline{V}_j$; that is,

$$V* = f(P_i,\ E_{ij},\ V_j). \quad (1)$$

The nature of the function, $\underline{f}$, depends on the definition of the state's value. This value may be defined in several different ways; for example, in terms of expected return, expected gain, or expected loss. The discussion and measures of state value that follow are independent of the function that is used. But for illustrative purposes, I shall use "expected relative value" as the state value, that is,

$$V* = \sum_{i=1}^{m} \sum_{j=1}^{n} P_i\, E_{ij}\, V_j\,. \quad (2)$$

Since $\underline{P}_i \leq 1.0$, $\underline{E}_{ij} \leq 1.0$, then, if a measure of relative value is used in which $0 \leq \underline{V}_j \leq 1$ and $\Sigma\ \underline{V}_j = 1.0$, it follows that the minimum and maximum values which the state value ($\underline{V}*$) can assume are zero and one, respectively.

Receipt of a communication involves a change in the receiver's purposeful state. Let $\underline{V}_1*$ represent the value of the initial state (just prior to receipt of the communication) and $\underline{V}_2*$ represent the value of the terminal or changed state where the change is the receiver's response to a message. Then the changes must be in one or more of his P_i's, E_{ij}'s, or V_j's, or some combination of these. Therefore, the value of the communication to the receiver is $V_2* - V_1*$. Even if only positive absolute values of V_j are used, the value of a communication may be negative: where $V_1* > V_2*$. For example, an oral prohibition from a parent may reduce the value of a situation to a child by precluding behavior which is a source of pleasure to him. Incorrect information can, as we shall see, also reduce the value of a purposeful state.

The value of a communication to its sender can be obtained by determining the message-produced change in his expected relative value from his initial to terminal state. There need be no correlation between the values of a message to the sender and the receiver. One may benefit, the other may not, or both may benefit or lose (unequally). The parent's communication to his child may increase the value of the parent's state (e.g., by the elimination of noise) but decrease the value of the child's state.

The value of the communication to third parties can similarly be determined: by finding the message-produced changes in their expected relative value from their initial to terminal states. One who overhears another's communication may benefit or lose for having done so.

MODES OF COMMUNICATION

A particular communication may change the receiver's probabilities of choice (P_i), the efficiencies of his choices (E_{ij}), the relative value of the possible outcomes (V_j), or some combination of these. Even where a communication produces a combination of changes in the receiver, each type of change can be studied separately. Each of the three types of change produced by a message can be identified and defined as follows.

8.2. Information. A communication which produces a change in any of the receiver's probabilities of choice informs him, and hence transmits information.

8.3. Instruction. A communication which produces a change in the efficiencies of any of a receiver's courses of action instructs him, and hence transmits instruction.

8.4. Motivation. A communication which produces a change in any of the relative values which the receiver places on possible outcomes of his choice motivates him, and hence transmits motivation.

There appears to be one other way in which a purposeful state can be changed: some of the available courses of action which were not potential choices of the receiver before a communication may become potential as a result of the communication. However, such a possibility is covered. Since, in a purposeful state, the available courses of action are formulated as an exhaustive and exclusive set, every possible choice is included. Therefore, if any choice which was not potential becomes so, this must be reflected in a change in a probability of selecting one of the alternatives.

Now let us examine each of the three modes of communication in more detail.

Information

Because of the pervasiveness of the use of "information" in Shannon's restricted (technical) sense, it might seem preferable to use another term here. But since the way that I use "informatinn" here conforms more closely to common usage than does Shannon's, if a change is required it would seem preferable to change Shannon's term. Shannon's usage is based on that of Hartley (1928). Cherry (1957) seems to reflect my opinion:

> In a sense, it is a pity that the mathematical concepts stemming from Hartley have been called 'information' at all. The formula for $\underline{H}_n$ is realy a measure of one facet only of the concept of information; it is the statistical rarity or 'surprise value' of a source of signs (p. 50).

Despite his terminology, Shannon was concerned with what might better be called the amount of message that is transmitted rather than with the amount of information that is communicated. He was primarily involved with systems in which each possible message can be coded into a combination of two symbols. For example, if there are four possible messages and two symbols (0 and 1), the messages can be represented as 00, 01, 10, and 11. Then, to select one message out of the four, two choices from among the two symbols (i.e., binary choices) may be made. One binary choice allows two messages (0 and 1) and three binary choices allows eight messages (000, 001, 010, 100, 110, 101, 011, and 111). In general, x binary choices allows 2^x possible messages.

For Shannon, the amount of "information" contained in a message is the amount of freedom of choice involved in the selection of the message.* A unit of choice is defined as the selection of one out of two equally available symbols. Thus, in selecting one of two equally available symbols, one choice-unit is involved and the resulting one-symbol message contains one unit of "information."

In general, if there are M equally available messages in a state, the selection of one contains x units of information where

$$x = \log_2 M.$$

Equal availability of the symbols means equal likelihood of choice by the sender. That is, if there are M possible messages and the probability of each being selected is $1/M$, complete freedom of choice exists. If the probability of selecting a particular message (p_i) deviates from $1/M$, the choice is not completely free. In the extreme case, if the probability of selecting any one of a set of messages is 1.0, then there is no freedom of choice and no "information" can be communicated by the one message which is always selected.

In order to cover cases in which choices are not equally likely (as well as where they are), Shannon derived the following

*An alternative approach to the measurement of syntactic information has been proposed by D. M. MacKay (1950 and 1955). A recent discussion of its application can be found in Payne (1966).

general measure of the amount of "information" (symbolized by $\underline{H}$ in his system) contained in a state:

$$H = \Sigma\, p_i \log p_i\ ,$$

where $\underline{p_i}$ is the probability of choice of the $\underline{i}^{th}$ message. If $\underline{\log_2}$ is used, then $\underline{H}$ is expressed in binary units which are called bits. Thus, a state which contains two equally likely messages contains one bit of "information."

The measure of information* to be developed here will also be related to freedom of choice; that is, it will be a function of the probabilities of choice associated with alternative courses of action. It will be a different function, however, because of the difference in selecting between messages and courses of action. The measure developed here is a function of the number of alternative potential courses of action, $\underline{m}$.

In Shannon's use of "information," we cannot speak of how much information a person has, only how much a message has. Clearly, from the behavioral scientists point of view, the former is much more important.**

When we talk of the amount of information that a person has in a specified situation (purposeful state), we do so in two different but related senses. First, we refer to the number of available courses of action of which he is aware; that is, to the number of potential courses of action. For example, a person who is aware of four exits from a

*Unless I indicate to the contrary "information" will henceforth be used as "pragmatic information."

**Attempts to use Shannon's theory of communication in the behavioral sciences has hardly met with success. See Hardy and Kurtz (1963) for an evaluation of these efforts. See also Schramm (1966) who observed, " ... we must admit frankly the difficulty of bridging the gap between the [H] formula's concept of information (which is concerned only with the number of binary choices necessary to specify an event in a system) and our concept of information in human communication..." (p. 534).

particular building has more information than the person who is aware of only two when there are four. The act of informing, then, can consist of converting available but not-potential choices into potential choices. For example, a statement such as "There are exits at either end of this hall" may convey information in this sense. The person who has this information (i. e., who has these potential choices) may or may not exercise it depending on his appraisal of the relative efficiencies of the alternative exits. In one sense, then, the amount of information in a state is a monotonically increasing function of the amount of potential choice of courses of action which an individual has in that state.

The second sense in which we talk of information involves the <u>basis</u> of choice from among the alternative potential courses of action. For example, an individual who knows which exit is nearest to him has a basis for choice and hence has information about the exits. Information in this sense pertains to the efficiencies of the alternatives relative to desired outcomes (e. g., a rapid exodus). Suppose, for example, that there are two exits and one is nearer to a person (A) than is the other. If A knows this and his objective (valued outcome) is to leave the building quickly, the choice is <u>determined</u> in the sense that A will always select the nearest exit. If he always selects the most distant exit then he is obviously misinformed (i. e., he has information, but it is incorrect). If he selects each exit with equal frequency then he apparently has no basis for choice; that is, no information. In this sense, then, information is the amount of choice which <u>has been</u> made. Now let us make this concept more precise.

Consider the case of an individual ($\underline{A}$) who is confronted by two potential courses of action, $\underline{C}_1$ and $\underline{C}_2$. If the probabilities of selecting the courses of action are equal, $\underline{P}_1 = \underline{P}_2 = \frac{1}{2}$, the situation may be said to be <u>indeterminate</u> for A.

8. 5. Indeterminate Choice Situation. A purposeful state in which a subject's probability of choice of each of the $\underline{m}$ available courses of action (defined so as to be exclusive and exhaustive) is equal to $1/\underline{m}$.

A person in an indeterminate state has no basis for choice and hence can be said to have no information about the alternatives. This is clearly the case when one of the alternatives is more efficient than the other. But if the two courses of action are equally efficient, the individual may have information to this effect and select each with equal frequency. Strictly speaking, however, he has no real choice in this situation since the alternatives are equally efficient. In a situation in which all alternative choices are equally efficient, information has no operational meaning. Such a situation does not constitute a purposeful state (see definition 3. 33). Consequently this discussion has relevance to only those situations in which the alternative courses of action are not necessarily equally efficient.

If $P_1 = 1.0$ and $P_2 = 0$, then the situation is determinate for the person involved. All the choice that can be made has been made. The maximum possible amount of information is contained in the state. It may not be correct information but this is another matter which will be considered below.

8. 6. Determinate Choice Situation. A purposeful state in which a subject's probability of selecting one of the available courses of action is equal to 1. 0.

Now we can define a unit of information as follows:

8. 7. Unit of Information. The amount of information which changes an indeterminate two-choice situation into a determinate choice situation.

Let us consider the general case involving m available courses of action. In order to select one from this set, a minimum of m - 1 choices from pairs of alternatives (i.e., paired comparisons) is required. Table 8.1 illustrates this fact.

TABLE 8.1

m =	2	3	4	5
	C_1	C_1	C_1	C_1
	1	1	1	1
	C_2	C_2	C_2	C_2
		2	2	2
		C_3	C_3	C_3
			3	3
			C_4	C_4
				4
				C_5

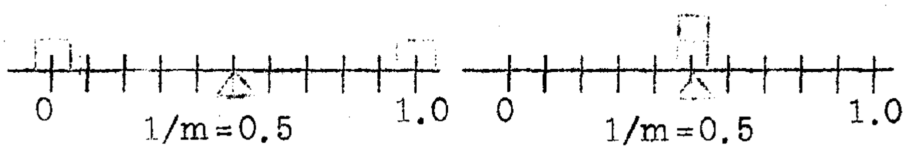
Implicit in Shannon's bit-measure of information is the assumption that an ultimate choice is the result of a series of choices from contracting dichotomous sets. For example, if there are four possible messages it is assumed these are grouped into two sets of two each, say (M_1 and M_2) and (M_3 and M_4). The first choice then consists of selecting one of these sets. The second consists of selecting one of the messages in the selected set. Hence two choices of different type are involved. The procedure of choosing among courses of action that I assume differs from the one just described; it involves three paired comparisons each of the same type. I do not assume choices are necessarily made in this way, although they may be, but I use this concept because it involves the maximum possible number of non-redundant choices.

The maximum amount (number of units) of information that a state can contain, then, is m - 1; that is, the amount of information required to choose completely from m - 1 pairs of alternatives.

We can conceive of the amount of information contained in a purposeful state as a point on a scale bounded at the lower end by no information in a state of indeterminism (i.e., no choice has been made), and at the upper end by complete information in a state of determinism (i.e., complete choice has been made). Location on this scale will depend on the values of the probabilities of choice, P_i.

Understanding these concepts is facilitated by visualizing a weightless platform that is scaled from 0 to 1.0 and is balanced on a fulcrum located at the value 1/m. A unit weight represents each course of action. Then two-choice determinate and indeterminate states can be represented as shown in Figure 8.1. Note that since $\Sigma P_i = 1.0$ these platforms will be in balance for every possible combination of P_i's. I shall use this analogy again as new concepts and measures are introduced.

Probabilities of Choice



(a) Determinate State (b) Indeterminate State

Figure 8.1. Physical analogy of information in a purposeful state.

In an indeterminate state each $\underline{P}_i = 1/\underline{m}$. Therefore, the amount of deviation of a state from indeterminism is

$$\sum_{i=1}^{m} \left| P_i - \frac{1}{m} \right| .$$

For an indeterminate state this sum is equal to zero. In a determinate state one $\underline{P}_i$ is equal to 1.0 and the remaining ($\underline{m}$-1) $\underline{P}_i$'s are equal to zero. Therefore, in such a state

$$\sum_{i=1}^{m} \left| P_i - \frac{1}{m} \right| = \left| 1 - \frac{1}{m} \right| + (m-1) \left| 0 - \frac{1}{m} \right| = 1 - \frac{1}{m} + (m-1)\frac{1}{m} = 1 - \frac{1}{m} + 1 - \frac{1}{m} = 2 - \frac{2}{m} .$$

Given a state with m possible courses of action. The fraction of the maximum possible amount of information that it contains is the ratio of (a) its deviation from the corresponding indeterminate state to (b) the deviation of the corresponding determinate state from that indeterminate state:

$$\frac{\sum_{i=1}^{m} \left| P_i - \frac{1}{m} \right|}{2 - \frac{2}{m}}$$

This ratio has a minimum value of zero and a maximum value of one.

The product of this fraction and the maximum amount of information that such a state can contain (i. e., m - 1) provides a measure of the amount of information (here symbolized by α) in that state:

8. 8. Amount of Information in an Individual's State (α) :

$$\alpha = (m-1) \frac{\sum_{i=1}^{m} \left| P_i - \frac{1}{m} \right|}{2 - \frac{2}{m}} = \frac{(m-1)\left(\frac{m}{2}\right) \sum_{i=1}^{m} \left| P_i - \frac{1}{m} \right|}{m-1}$$

$$= \frac{m}{2} \sum_{i=1}^{m} \left| P_i - \frac{1}{m} \right|$$

where m is the number of (exclusively and exhaustively defined) available courses of action and the P_i's are the probabilities of the subject's selecting the i^{th} course of action.

The net amount of information communicated is the amount of information contained in the state of the receiver immediately following the communication (the terminal state) minus the amount contained in his state immediately preceding the communication.

8.9. Net Amount of Information Communicated to a Receiver (α_N):

$$\alpha_N = \alpha_2 - \alpha_1 = \frac{m}{2} \sum_{i=1}^{m} \left| P_i' - \frac{1}{m} \right| - \frac{m}{2} \sum_{i=1}^{m} \left| P_i - \frac{1}{m} \right| ,$$

where α_2 and α_1 are the amounts of information contained in the terminal and initial states, respectively; and P_i' and P_i are the probabilities of choice in the terminal land initial states, respectively.

This measure can take on values from $-(m-1)$ to $(m-1)$. Negative values represent a loss of information (e.g., as in going from a determinate to an indeterminate state).

Suppose that in an initial state involving two courses of action, C_1 and C_2, $P_1 = 1.0$ and $P_2 = 0$. This state contains one unit of information. If as a result of communication P_1 is changed to 0 and P_2 to 1.0, the terminal state also contains one unit of information. Hence, the net amount of information communicated (α_N) is equal to zero. This results, so to speak, because the information in the initial state was removed and replaced by an equal amount of different information. Clearly, the value of the terminal and initial states to the receiver may differ, and this will be reflected in the measure of the value of information to be developed below. It does seem peculiar, however, to say that no information has been transmitted; one should more properly say that although the net amount of information transmitted was zero, there was an exchange of information. Therefore, if we develop a measure of the gross amount of information transmitted, substraction of the amount transmitted from this provides a measure of the amount of information exchanged.

In measuring the net amount transmitted we determine the amount by which the initial and terminal states differed from an indeterminate state. Now let us measure the amount by which the terminal state differs from the initial state: $\Sigma | P'_i - P_i |$. As before, let us take the ratio of this deviation to the maximum distance deviation ($2 - \frac{2}{m}$), and multiply it by the maximum amount of information that the state can contain (m-1).

8.10. Gross Amount of Information Communicated to Receiver (α_G):

$$\alpha_G = m-1 \frac{\Sigma | P'_i - P_i |}{2 - \frac{2}{m}} = \frac{m}{2} \Sigma | P'_i - P_i |.$$

This quantity has a minimum value of zero and (since max $\Sigma | P'_i - P_i |$ = 2.0) a maximum value of m.

8.11. Amount of Information Exchanged (α_E) :

$$\alpha_E = \alpha_G - | \alpha_N |.$$

Since $\alpha_G \geq \alpha_N$, this measure has maximum and minimum values of m and 0, respectively.

Returning to the previous example in which P_1 changed from 1.0 to 0 and P_2 from 0 to 1.0, since the amount of information in both states was 1.0; α_N, the net amount transmitted was 0. The gross amount of information transmitted in this case is

$$\alpha_G = \frac{2}{2} [(1.0) + (1.0)] = 2.0.$$

Hence, the amount of information exchanged is

$$\alpha_E = 2.0 - 0 = 2.0,$$

the maximum amount possible.

Returning to the physical analogy (see Figure 8.2) it is apparent that the sums of the distances from the fulcrum (1/m) in the terminal and initial states are both equal to 1.0. Hence, the amounts of information in these states are equal and the net amount of information communicated is equal to zero. However, the total distance travelled by C_1 and C_2 over the P_i- scale is 2.0 (the gross amount of information communicated). The difference between the gross and net amounts of information communicated (2.0-0 = 2.0) is the amount exchanged. The amount exchanged can be interpreted as the amount of movement from the initial state less the minimal amount required to obtain the same amount of information contained in the terminal state.

These measures can be illuminated by considering the slightly more complex examples shown in Table 8.2.

TABLE 8.2.

i	Initial State P_i	Terminal States P_i'	$\|P_i' - P_i\|$	P_i'	$\|P_i' - P_i\|$
1	0	0	0	0.6	0.6
2	0.1	0.1	0	0.1	0
3	0.1	0.1	0	0.1	0
4	0.1	0.2	0.1	0.2	0.1
5	0.7	0.6	0.1	0	0.7
α =	2.5	2.0		2.0	
$\Sigma\|P_i' - P_i\|$=			0.2		1.4

The net amount of information communicated in both cases is 2.0 - 2.5 = -0.5 units. For the first terminal state the gross amount of information communicated is $\frac{5}{2}(0.2) = 0.5$. Therefore, the amount of information exchanged in this case is 0.5 - 0.5 = 0. For the second terminal state, however, the gross amount of information communicated is $\frac{5}{2}(1.4) = 3.5$, and hence the amount exchanged is 3.5 - 0.5 = 3.0.

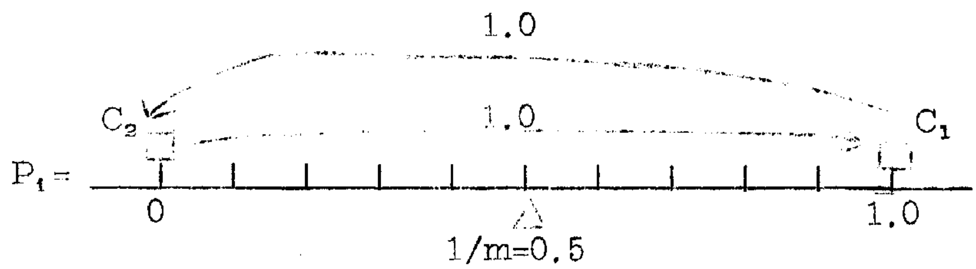
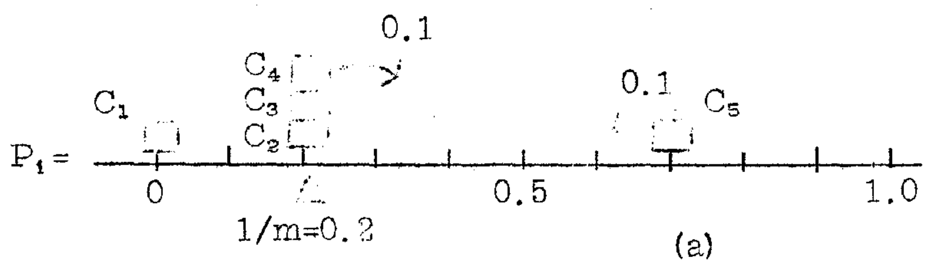


Figure 8.2.

Returning to the physical analogy (see Figure 8.3) note in (A) that the sum of the distances from the fulcrum is decreased and hence a negative net amount of information is transmitted. The gross amount transmitted is proportional to the sum of the distances traveled (0.2). Since this sum is the minimal amount required to reach a terminal state with the distribution of P_t's indicated, no information has been exchanged. In the second case (b) of the total movement ($0.1 + 0.6 + 0.7 = 1.4$) it is clear that two moves of distance 0.1 each would have produced the same distribution of P_t's. Since $\frac{m}{2}(0.2) = \frac{5}{2}(0.2) = 0.5$, then $2.5 - 0.5 = 2.0$ is the amount of information exchanged.



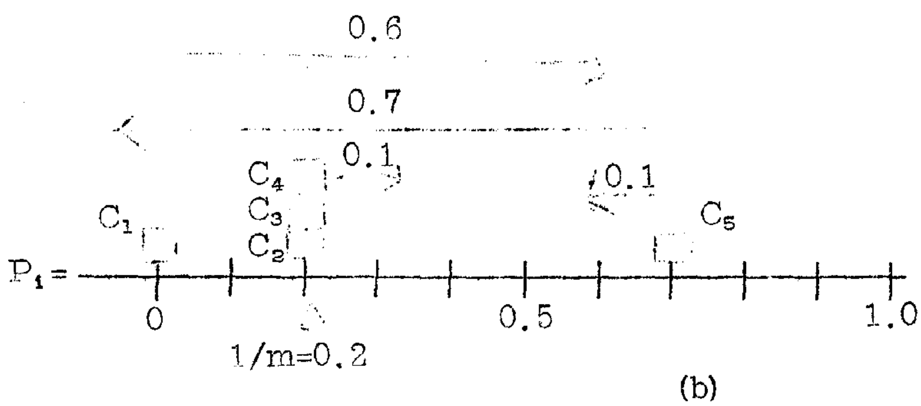


Figure 8.3

The measure of information which has been developed here depends on how the alternative courses of action are formulated by the investigator. For example, suppose one investigator formulates two exclusive and exhaustive courses of action:

C_1: use of an automobile

C_2: use of any other mode of transportation

and another investigator formulates

C_1: use of an automobile

C_2: use of bus

C_3: walking

C_4: use of any other mode of transportation

If the subject always uses an automobile ($P_1 = 1.0$ in both cases), then the first investigator would find one unit of information, and the second would find three. Hence, the measure depends on the investigator as well as the subject.

There are two aspects of this "relativity" of the measure of information which should be noted. First, it is possible to adjust the measures obtained by the two investigators so that they are in agreement. The definitions have been constructed so as to make this possible. Secondly, the same "relativity" is present in Shannon's measure of syntactic information. In applying his measure, one can use a letter of the alphabet, a phoneme, a word, or even a message as a unit for which the probabilities of choice are to be determined. The use of different units may yield different (Shannon) amounts of syntactic information in a message.

As long as we can make comparable the results of different investigations of the same thing, the fact that they may yield apparently different results presents no serious methodological problem.

It is also important to observe that the measures of information developed here contain no implication concerning the correctness or incorrectness of the information received. Further, it should be noted that this measure is relative to a specific receiver in a specific state. The same message may convey different amounts of information to different individuals in the same choice environment or to the same individual in different choice environments. Consequently, to specify the amount of information transmitted by a message it is necessary to specify the set of individuals and states relative to which the measure is to be made. If more than one individual or state is involved it is also necessary to specify what statistic (e.g., an average) is to be used. Generality of information may be defined in terms of the range of individuals and/or states over which it operates.

It should also be noted that messages are not the only possible source of information; one may also obtain information by perception. The measures of information given here are equally applicable to states before and after perception. The measures of instruction and motivation to be developed below are also applicable to perception. This generality is an important property of these measures.

Finally, what a message that informs does is either (1) change the subject's conception of the choice situation (what choices he believes are possible) and, through such changes, modifies his beliefs in the efficiencies of the alternatives that he perceives; or (2) changes his beliefs in efficiencies without changing his beliefs about available choices.

Instruction

To inform is to provide a basis for choice; that is, a belief in the greater efficiency of one choice compared to another. Hence information modifies objective probabilities of choice by modifying believed(subjective) probabilities of success. Instruction is concerned with modification of the objective probabilities of success: efficiency. The amount of instruction that a subject has in a particular state is equivalent to the amount of control that he can exert over possible outcomes in that state. He has maximum control if he is capable of bringing about any of the possible outcomes by any of the means available to him. To instruct is to impart such a capability where it is lacking.

Consider a course of action C_1 and two (exclusively and exhaustively defined) outcomes, O_1 and O_2. A person has complete control over C_1 if he can use it to make either outcome occur with certainty when he desires that outcome. For example, if he can "use an automobile" (C_1) to go somewhere (O_1) or not (O_2), he has complete control over the course of action and the outcome. If he cannot change the probability of an outcome occuring by changing the way he pursues a course of action, then he does not control that course of action and the outcome. Suppose, for example, that the efficiency of C_1 for O_1, E_{11}, is equal to 1.0 no matter what the subject desires, and hence the efficiency of C_1 for O_2 must be equal to 0. Then his choice is like pushing a button that releases a mechanically defined course of action over the outcome of which he has no control.

A measure of the amount of control that a person has in a purposeful state can be developed as follows. Consider one course of action, C_1, and two outcomes, O_1 and O_2. If (a) when the relative value of O_1 to the subject is maximum (say, 1.0; i.e., $V_1 = 1.0$) and hence $V_2 = 0$, the efficiency of his use of C_1 for O_1 is 1.0 (i.e., $E_{11} = 1.0$); and (b) when the relative values are reversed (i.e., $V_1 = 0$ and $V_2 = 1.0$), the efficiency of his use of C_1 for O_2 is 1.0 (i.e., $E_{12} = 1.0$); then he has maximum control over C_1. Therefore, the amount of control that a person has over a course of action is reflected in the range of its efficiency as a function of the value he places on possible outcomes.

8.12. <u>Amount of Control (β_{ij}) that an Individual Has over a Particular Course of Action (C_i) relative to a Particular Outcome (O_j)</u>:

$$\beta_{ij} = (E_{ij} \mid V_j = 1.0) - (E_{ij} \mid V_j = 0).$$

This quantity has a maximum value of 1.0 and a minimum value of -1.0. For example, suppose the course of action is "use of a desk calculator" and the two outcomes are "correct computation" and "incorrect computation." If a subject can always use the calculator correctly when he wants to and always incorrectly when he so desires, then he has complete control over use of the desk calculator with respect to the relevant computations.

When we consider a course of action (C_i) over a set of (exclusively and exhaustively defined) outcomes ($O_1, O_2, \ldots, O_n$) then, for every pair of outcomes, O_j and O_k, the following equality holds:

$$[(E_{ij} \mid V_j = 1.0) - (E_{ij} \mid V_j = 0)] = [(E_{ik} \mid V_k = 1.0) - (E_{ik} \mid V_k = 0)].$$

This follows from the fact that

$$(E_{11} \mid V_1 = 1.0) + (E_{12} \mid V_2 = 0) + \ldots + (E_{1n} \mid V_n = 0) = 1.0$$

$$(E_{11} \mid V_1 = 0) + (E_{12} \mid V_2 = 1.0) + \ldots + (E_{1n} \mid V_n = 0) = 1.0$$

$$\vdots$$

If we substract the second equation from the first we obtain

$$(E_{11} \mid V_1 = 1.0) - (E_{11} \mid V_1 = 0) + (E_{12} \mid V_2 = 0) - (E_{12} \mid V_2 = 1.0) = 0.$$

Therefore,

$$(E_{11} \mid V_1 = 1.0) - (E_{11} \mid V_1 = 0) = (E_{12} \mid V_2 = 1.0) - (E_{12} \mid V_2 = 0).$$

This result can be obtained for each pair of outcomes. From this we can obtain the following measure.

8.13. Amount of Control (β_i) that an Individual Has over a Particular Course of Action (C_i) relative to as Set of n Outcomes:

$$\beta_i = n\beta_{ij}.$$

This measure has maximum and minimum values of n and -n, respectively.

Now one can generalize over a set of courses of action.

8.14. Amount of Control (β) that an Individual Has in a Purposeful State relative to a Set of m Courses of Action and a Set of n Outcomes:

$$\beta = \Sigma_{i=1}^{m} \beta_i = \Sigma_{i=1}^{m} n\beta_{ij}.$$

This measure has maximum and minimum values of mn and -mn, respectively.

The amount of control an individual has in a state is the amount of instruction he has in that state.

8.15. The Net Amount of Instruction Communicated to a Receiver (β_N):

$$\beta_N = \beta_2 - \beta_1,$$

where β_2 and β_1 are the amounts of instruction in the terminal and initial states, respectively.

This measure has maximum and minimum values of 2mn and -2mn, respectively.

Communication can result in "unlearning" as well as learning, that is, the loss of control. The gain or the loss of control may be either good or bad for the subject; the amount and value of control must be measured separately. A measure of its value will be developed below.

To obtain a measure of the gross amount of instruction transmitted we sum over the absolute values of the changes that occur relative to each C_i-O_j combination.

8.16. The Gross Amount of Instruction Communicated to a Receiver (β_G):

$$\beta_G = \Sigma_{i=1}^{m} n|\beta_{ij}' - \beta_{ij}|,$$

where β_{ij}' refers to the terminal state and β_{ij} refers to the initial state.

Since $|\beta_{ij}' - \beta_{ij}|$ has maximum and minimum values of 2 and 0, respectively, β_G has maximum and minimum values of 2mn and 0, respectively.

8.17. The Amount of Instruction Exchanged (β_E) :

$$\beta_E = \beta_G - |\beta_N|.$$

Since $\beta_G \geq \beta_N$, this measure has maximum and minimum values of 2mn and 0, respectively.

Information and instruction are also relative concepts; one can be converted into the other by redefinition of courses of action. For example, consider the course of action "using a computer" where the outcome of interest is a correct solution to an equation. The subject involved may not be aware of the availability of a "packaged program" for solving the equation and hence he does his own programming with, say, a probability of success (efficiency) equal to 0.5. If a message makes him aware of the packaged program his efficiency in use of the computer may go up to 1.0. Then we can measure the amount of instruction he has received.

On the other hand, we could define two courses of action: C_1 which is "use of the computer with his own program" and C_2 which is "use of the computer with a packaged program." Then, before the communication, his probability of selecting the first course of action (C_1) may have been 1.0, but after communication the probability of selecting the second course of action (C_2) may change from 0 to 1.0. In this interpretation, the message would be said to inform rather than instruct.

Again this relativity of concepts presents no problem as long as we can adjust the two different interpretations of the same objective state of a subject so that they become comparable. The definitions provided here clearly make such adjustment possible.

Motivation

If an individual equally values all possible outcomes in a purposeful state, then he has no basis for selecting one from among them to pursue. He can be said to be unmotivated in that state. It should be recalled that the outcomes used to define a purposeful state are defined so as to be exclusive and exhaustive. Then, since the maximum relative value of an outcome is 1.0, the sum of the relative values over a set of outcomes is also 1.0. Therefore, if relative value is added to one outcome, an equal amount must be subtracted from others.

A state containing no motivation is described by the condition: $V_1 = V_2 = \ldots = V_n = \frac{1}{n}$. A state containing complete motivation is one in which one outcome has a relative value of 1.0 and all the others have none. These observations correspond exactly to those made in the discussion of information and probability of choice. Therefore, measures of motivation communicated can be formulated in a way that is completely analogous to the way used to develop measures of information.

8.18. Amount of Motivation in an Individual's State (γ):

$$\gamma = \frac{n}{2} \Sigma^{n}_{j=1} | V_j - \frac{1}{n} |.$$

8.19. <u>Net Amount of Motivation Communicated to a Receiver</u> (γ_N):

$$\gamma_N = \gamma_2 - \gamma_1,$$

where γ_2 and γ_1 are the amounts of motivation contained in the terminal and initial states, respectively.

8.20. <u>Gross Amount of Motivation Communicated to a Receiver</u> (γ_G):

$$\gamma_G = \frac{n}{2} \Sigma |V_j' - V_j|,$$

where V_j' and V_j are the relative values of outcomes in the terminal and initial states, respectively.

8.21. <u>Amount of Motivation Exchanged</u> (γ_E):

$$\gamma_E = \gamma_G - |\gamma_N|.$$

As was observed in Chapter 2, courses of action and outcomes (means and ends) are relative concepts. That is, by reconceptualizing a subject's purposeful state an investigator can convert courses of action into outcomes, or outcomes into courses of action. Therefore, by using such transformations it is possible to convert what appears as information in one formulation of another's purposeful state into motivation in another formulation; or, conversely, to convert motivation into information. Finally, since we noted in the last section that instruction and information could be converted into each other, it follows that each of the three measures has a transformation into each of the others.

THE VALUE OF THE COMPONENTS OF COMMUNICATION

It will be recalled that the value of a communication to the receiver is given by $V_2^* - V_1^*$, where these are the values to him of his terminal and initial states, respectively. Using expected relative value for the measure of value of a state, this difference may be rewritten as follows:

$$V_2^* - V_1^* = \Delta V^* = \Sigma_{j=1}^{n} \Sigma_{i=1}^{m} (P_i + \Delta P_i)(E_{ij} + \Delta E_{ij})(V_j + \Delta V_j) -$$

$$\Sigma_{j=1}^{n} \Sigma_{i=1}^{m} P_i E_{ij} V_j .$$

By expansion this equation may be converted into the following:

$$\Delta V^* = \Sigma\Sigma \Delta P_i E_{ij} V_j + \Sigma\Sigma P_i \Delta E_{ij} V_j + \Sigma\Sigma P_i E_{ij} \Delta V_j +$$
$$\Sigma\Sigma \Delta P_i \Delta E_{ij} V_j + \Sigma\Sigma \Delta P_i E_{ij} \Delta V_j + \Sigma\Sigma P_i \Delta E_{ij} \Delta V_j +$$
$$\Sigma\Sigma \Delta P_i \Delta E_{ij} \Delta V_j .$$

The first three terms represent the value added to the initial state by the communicated information, instruction, and motivation, respectively.

8.22. <u>Value of Information Communicated</u> (ΔV_α^*):
$\Delta V_\alpha^* = \Sigma\Sigma \Delta P_i E_{ij} V_j$.

8.23. <u>Value of Instruction Communicated</u> (ΔV_β^*):
$\Delta V_\beta^* = \Sigma\Sigma P_i \Delta E_{ij} V_j$.

8.24. <u>Value of Motivation Communicated</u> (ΔV_γ^*):
$\Delta V_\gamma^* = \Sigma\Sigma P_i E_{ij} \Delta V_j$.

Any of these expressions may be either positive or negative. If ΔV_α^* is negative, the receiver has been <u>misinformed</u>; if positive, he has been informed. If ΔV_β^* is positive, he has been instructed; if negative, he has been "misinstructed." Unfortunately we have no commonly used negative of the verb "to instruct." The same remarks apply to ΔV_γ^*.

The remaining four terms in the equation for ΔV^* represent $\Delta V_{\alpha\beta}^*$, $\Delta V_{\alpha\gamma}^*$, $\Delta V_{\beta\gamma}^*$, and $\Delta V_{\alpha\beta\gamma}^*$. For example, $\Delta V_{\alpha\beta}^*$ is the joint contribution (<u>not</u> the sum of the independent contributions) to value of the information and instruction communicated. The other terms may be interpreted similarly. It is convenient, then, to think of the value of a communication as the sum of the independent and dependent contributions of information, instruction, and motivation. That is

$$\Delta V^* = \Delta V_{\alpha}^* + \Delta V_{\beta}^* + \Delta V_{\gamma}^* + \Delta V_{\alpha\beta}^* + \Delta V_{\alpha\gamma}^* + \Delta V_{\beta\gamma}^* + \Delta V_{\alpha\beta\gamma}^*.$$

CONCLUSION

Some attempts to apply the measures which have been developed here are described in Appendix II. Such applications are not easy. They are time-consuming and costly, and may require a degree of control over subjects that is difficult, if not impossible, to obtain. The situation in which we find ourselves with respect to these measures is similar to the one a physicist would be in if the only way of measuring the temperature of a body were to determine the mean-squared velocity of its point particles. We have yet to develop "thermometers" to facilitate measures of human communication; but measures such as have been developed here can take us a giant step toward easy and <u>relevant</u> measurement.

Apparently easy measurement may not be measurement at all and may not even be relevant. Good measures have usually evolved through four stages. In the first stage, subjective judgment is used. For example, we "estimated" the intelligence of people or, at one time in history, the temperature of an object. In the second stage easier-to-apply <u>indices</u> are sought which correlate highly with "expert" judgment. For example, the procedure described by Thurstone and Chave (1929) for the construction of attitude tests--a procedure still followed widely--is based on correlation of test scores with "expert" judgment. Such objective indices of subjective judgment, however useful they may be, do not yield measures in any strict sense because they involve no unit of measurement and, more important, no idealized operational definition of, and hence standard for, the property being quantified. At the present time, for example, citation counts provide such a subjective index of the value of a scientific article because they are not based on an operational definition and measure of the "value of a scientific article."

The third stage of the evolution is the development of idealized operational definitions and measures of the property involved, such as we have tried to develop here, or as in the development of a definition of temperature as "mean-squared velocity of point particles." The existence of such measures, even when not practical or easy to apply, as in the above definition of temperature, provides an objective standard for which indices can be sought. Development of such indices--ones which correlate with, or are structurally related to, the standard (e.g., use of thermometers)--constitute the fourth stage of the evolution.

Very few of the so-called measures in the behaviorial sciences have gone beyond the second stage of this evolution. Even most of the standard psychological tests provide, at best, indices of human judgment, not indices of objective measures.

The analysis which yielded the measures defined here show the dangers of indiscriminately applying Shannon's measures to human communication. They do not deal with most of the important characteristics of such communication. The measures proposed here will certainly be modified and replaced in time, but what replaces them should be at least as rich as they are.

I turn now to an analysis of the meaning of "message" on which the definition of "communication" offered here is based.

BIBLIOGRAPHY

Attneave, Fred, Applications of Information Theory to Psychology. Holt, Rinehart and Winston, New York, 1959.

Berlo, D. K., The Process of Communication. Holt, Rinehart and Winston, New York, 1960.

Cherry, Colin, On Human Communication. John Wiley and Sons, New York, 1957.

Hardy, R., and P. Kurtz, "A Current Appraisal of the Behavioral Sciences," Research Council Bulletin, Sec. 6, (1963), 99-105.

Hartley, R. V. L., "Transmission of Information," Bell System Technical Journal, 7(1928), 535f.

MacKay, D. M., "Quantal Aspects of Scientific Information," Philosophical Magazine, 4(1950), 289-311.

________, "Quantal Aspects of Scientific Information," IRE Representation of Processes, Symposium on Information Theory, PGIT-1 (1955), 60-80.

Martin, M. W., Jr., "The Measurement of Value of Scientific Information," in Operations Research in Research and Development, ed. by B. V. Dean. John Wiley and Sons, New York, 1963, pp. 97-123.

Payne, Buryl, "A Descriptive Theory of Information," Behavioral Science, 11(1966), 295-305.

Pierce, J. R., Symbols, Signals and Noise. Harper & Brothers, New York, 1961.

Rapoport, Anatol, "What is Information?" in Communication and Culture, ed. by A. G. Smith. Holt, Rinehart and Winston, New York, 1966, pp. 41-55.

Schramm, Wilbur, "Information Theory and Mass Communication," in Communication and Culture, ed. by A. G. Smith. Holt, Rinehart and Winston, New York, 1966, pp. 521-534.

_______, (ed.), The Science of Human Communication. Basic Books, New York, 1963.

Shannon, C. E., and W. Weaver, The Mathematical Theory of Communication. University of Illinois Press, Urbana, 1949.

Thurstone, L. L., and E. J. Chave, The Measurement of Attitudes. University of Chicago Press, Chicago, 1929.

Watzlawick, P., J. H. Beavin, and D. D. Jackson, Pragmatics of Human Communication. W. W. Norton, New York, 1967.

Chapter 9

SIGNS, MESSAGES, AND LANGUAGE

LANGUAGE, n. The music with which we charm the serpents guarding another's treasure (Ambrose Bierce, The Devil's Dictionary).

INTRODUCTION

The definition of 'communication' given in Chapter 8 used the concept 'message' which is as yet undefined. Since a message consists of one or more signs, it is first necessary to define 'sign.' I do so in this chapter and also develop a set of measures to characterize sign performance. Using these concepts, 'message' and 'language' are then defined.

The conceptual development in this chapter is similar in many respects to that provided by Charles Morris (1946 and 1964). The names of the concepts in my treatment are similar to his, but the kind of definitions given are quite different. Although Morris's work is behaviorally oriented he does not provide operational definitions of the concepts he treats and only infrequently do his definitions specify measures of the variables involved in them. Finally, his effort does not involve placing his treatment of signs within the general context of purposeful behavior even though his approach is teleologically oriented. Nevertheless, as will be apparent to those familiar with Morris's work, my debt to him is considerable.

Morris is probably more responsible than any other single person for what attention has been given to the pragmatic study of signs. He popularized the term 'semiotic' about which he wrote (1964):

> Semiotic has for its goal a general theory of signs in all their forms and manifestations, whether in animals or men, whether normal or pathological, whether linguistic or non-linguistic, whether personal or social. Semiotic is thus an interdisciplinary enterprise.
>
> Part of the widespread interest in this area is motivated by the belief that higher-level sign processes (often called symbols) are of central importance in understanding man and his works. Ernst Cassirer called man "the symbolic animal" ..., instead of the "rational animal" ..., and much contemporary work has shown the aptness of this conception.
>
> The term 'semiotic' was adopted by John Locke from the Greek Stoics, who in turn were influenced by the Greek medical tradition that interpreted diagnosis and prognosis as sign processes. Charles S. Peirce (1839-1914), who followed John Locke's usage, is responsible for the present wide-spread employment of the term 'semiotic'...
>
> Philosophers and linguists made the main historical contributions to the general theory of signs, but today extensive work in this area is also being done by psychologists, psychiatrists, aestheticians, sociologists, and anthropologists (p. 1).

Morris himself did not produce a theory of signs in the usual sense of theory, but rather a conceptual framework within which such a theory could be developed. To a large extent this chapter is devoted to modifying his conceptual framework and imbedding it in the more general conceptual system being constructed here. This, I hope, will increase its usefulness in both constructing a theory of signs and executing the experimentation on which such a theory must be based.

SIGNS

We can divide the task of analyzing the meaning of 'sign' into two questions: "What can be called signs?" and "By virtue of what properties can they be called signs?" The first of these questions is the easier to answer.

It is apparent that objects can be signs; for example, billboards, posters, and, in general, those physical objects we commonly call signs. But behavior patterns can also be signs; for example, gestures and speech. Sometimes it is fruitful to consider the properties of objects and behavior as signs rather than objects and behavior themselves. For example, a red light is frequently a sign of danger but we do not respond to all the properties of the object which throws the light. We may not respond to the material of which the lamp is made, but we do respond to its redness and location. This distinction between objects, behavior, and their properties is only a matter of emphasis since only objects or events (including behavior) have properties, hence a response to a property is also always a response to that which has the property. It will be important, however, to identify the properties of an object or event which makes it serve as a sign.

At the level of common sense it is apparent that an object, event, or property which is a sign is a sign of something. This suggests that something, *X*, is a sign of something else, *Y*, if it can in some sense substitute for *Y*. It is in an analysis of the nature of this substitution that the nature of a sign is to be found. Recognition of this fact was at the base of Morris's work. Morris' analysis, however, goes well beyond what can appropriately be called common sense.

Morris (1946) began his analysis with the following preliminary definition:

> If something, *X**, controls behavior towards a goal in a way similar to (but not necessarily identical with) the way something else, *Y**, would control behavior with respect to that goal in a situation in which it was observed, then *X* is a sign (p. 7).

Morris then defined a series of concepts in terms of which he revised his preliminary definition of "sign". The concepts and definitions

---

*He used '*A*' where I use '*X*' and '*B*' where we use '*Y*', but I replace them for consistency with previously used symbols.

are as follows:

1. Preparatory-stimulus: "any stimulus which influences a response to some other stimulus."

2. Stimulus: "any physical energy which acts upon a receptor of a living organism; the source of this energy will be called the stimulus-object."

3. Response: "any action of a muscle or gland."

4. "Disposition to respond in a certain way is a state of an organism at a given time which is such that under certain additional conditions the response in question takes place."

5. Response-sequence: "any sequence of consecutive responses whose first member is initiated by a stimulus object and whose last member is a response to this stimulus-object as a goal object, that is, to an object which partially or completely removes the state of the organism (the 'need') which motivates the sequence of responses."

6. Behavior-family: "any set of response-sequences which are initiated by similar stimulus-objects and which terminate in these objects as similar goal-objects for similar needs."

Then, according to Morris, "the set of conditions sufficient for something to be a sign' is "if anything, X, is a preparatory-stimulus which in the absence of stimulus-objects initiating response sequences of a certain behavior-family causes a disposition in some organism to respond under certain conditions by response-sequences of this behavior family, then X is a sign" (pp. 8-10).

I use Morris' definition as a point of departure. My departure from it is intended to eliminate its bio-physical orientation and recast it in functional terms related to a purposeful state. It will be recalled (from definition 4.1) that a stimulus is anything which produces a change in the functional properties of a subject in a purposeful state, and that a response is the change in the functional properties of a subject that is produced by a stimulus. Hence, a stimulus produces a change in either the subject's probabilities of choice, efficiencies of choice, relative values of outcomes, or some combination of these; that is, it informs, instructs, or motivates him.

Consequently, for me, a 'preparatory stimulus' is anything which produces a response to something other than itself. A sign is such a stimulus; it produces responses to other stimuli, but I do not restrict these other stimuli to objects or events. These other stimuli may, for example, be either concepts or images (both of which I will define later), or signs themselves.

Everything that produces a response produces a response to itself in a trivial sense. Therefore, we do not want to call every stimulus a sign. A closed door produces a turning of its knob, but we do not want to call the door a sign.

According to Morris (1964), a sign produces a disposition to respond:

> ... a disposition to react in a certain way because of the sign (food-seeking behavior or site-probing behavior in the case of bees), has no necessarily "subjective" connotation. Such a disposition can, if one wishes, be interpreted in probabilistic terms, as the probability of reacting in a certain way under certain conditions because of the appearance of the sign (p. 3).

Hence, for Morris, a sign produces a potentiality for response. I prefer, however, to place the potentiality in the sign rather than in the respondent because, for Morris, an X is a sign only if it produces a disposition to respond; when it does not do so it is not a sign.

It seems to me that X should be a sign if it can produce the required type of response, even though it may not be doing so in a particular situation.

9.1. Sign: anything which is a potential producer of a response to something other than itself.

9.2. Signification of a Sign: that to which a sign potentially produces a response.

This permits an X to be a sign to a potential respondent even though he is not responding or is not disposed to respond to it at the moment. For example, we can say a book or a letter contains signs even though no one is reading it at the moment. Yet we can determine experimentally if the marks in the book have the required potentiality.

Note that there is no requirement that a sign and that to which it produces a response be in the same environment or even exist at the same time. The name of a person in another environment, or who has died, can produce a response to him. Furthermore, since a purposeful state has been so defined that machines (e.g., computers) can be placed in such states (by appropriate programming) an X may be a sign of something to a machine as well as to a person. I want the definitions of "communication," "signs," "message," and "language" to permit communication to and with machines. This is one of the reasons for eliminating the biological orientation of Morris's definition.

The definition of 'sign' presented here is very similar to one which Morris rejected. He based his rejection on the case of a drug which produces a sensitivity in an individual to something he would not otherwise respond to. Administration of such a drug appears to satisfy the sign-requirements, but, Morris argued, this conflicts with common sense. Note, however, that the drug leaves no choice to the responder; it _imposes_ the increased sensitivity on him. This is critical. If we were to use physical force on a person to make him look at something, the applied force would not be a sign of what he perceives because he was not free to do otherwise. A stimulus is a producer, not a deterministic cause, of a response. It is a necessary but _not_ a sufficient condition of the relevant response. Therefore, something like a drug which is sufficient to produce a response to something else in a given set of circumstances is _not_ a stimulus, and hence is not a sign.

It is apparent at the common-sense level that many signs operate in the way we have described. For example, when someone yells "Fire" in a burning building it may produce a wide variety of purposeful responses to that fire: escape, attempts to subdue the fire, to save contents of the building, and so on. Signs of fire may be spoken words, written, gestures, or objects or events (e.g., a screaming siren or blinking red lights). Note that smoke is a different kind of a sign of fire than is the word "fire;" it is a _natural_, not a man-made or _artificial_, sign. But both operate in the same way: producing responses to the fire. It is not equally obvious that such signs as "and" or "plus" satisfy these conditions, but we shall consider such less obvious cases below.

The meaning of a sign can be shown schematically as is done in Figure 9.1.

The way that signs can be studied is conditioned by the fact that their essential property is functional in character. Before turning to a more detailed analysis of how they function it should first be noted that the structural relationships between different signs may be the subject of study. Such studies form the branch of semiotic called _syntactics_. Morris (1946) defined syntactics as "that branch of semiotic that studies the way in which signs of various classes are combined

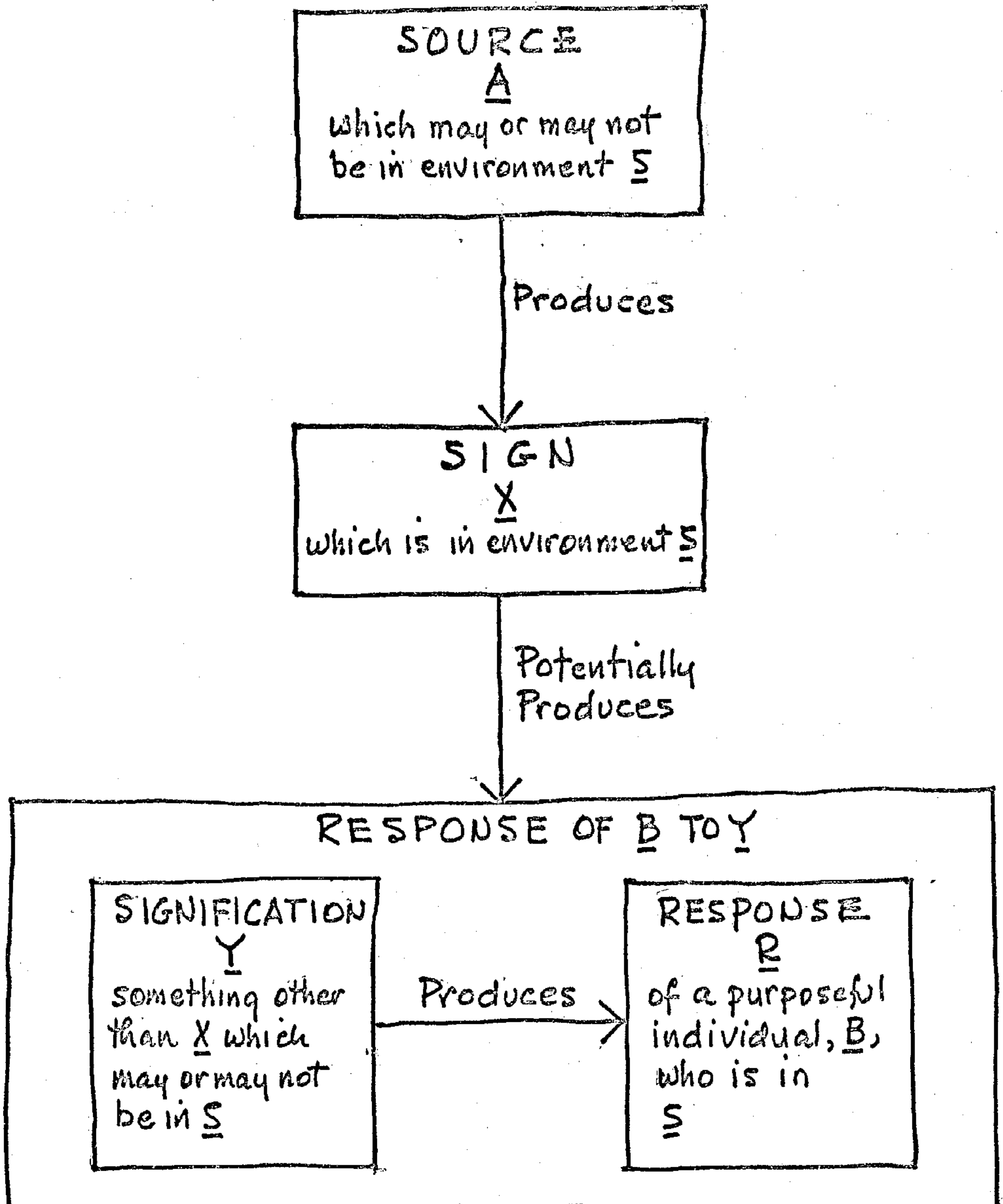


Fig. 9.1. Diagram of a sign's operation.

to form compound signs" (p. 355). Hence the study of grammar is part of syntactics. Much of logic can also be looked at as a part of this branch of semiotic. Since my concern here is with the way signs function little reference will be made to syntactics. It will come up, however, when I discuss language later in this chapter.

Figure 9.1 provides a basis for analysis of the functioning of signs. We may concentrate our attention on the relationship between the sign, X, and that which it signifies, Y; or we may consider the relationship between the sign and the response, R, or respondent, B, and/or the source, A. Analyses of the first type are called semantic; analyses of the second type are called pragmatic. Semantics, therefore, is the study of what signs refer to, their signification; pragmatics is the study of their effects, the characteristics of the responses that they bring about. According to Morris (1964), "Pragmatics is the aspect of semiotic concerned with the origin, uses, and effects of signs" (p. 44). He used 'origin' in the same sense in which I use 'source'.

SEMANTIC PROPERTIES OF SIGNS: DENOTATION AND CONNOTATION

The secondary stimulus, Y, to which a sign produces a response (i.e., what it signifies) can be considered in two ways: denotatively and connotatively.

9.3. Denotation of a Sign: the set of objects or events which are signified by a sign.

Hence the denotation in the shout "Fire" in a building is the fire in that building, but when we speak of fire as in 'fire prevention' we denote a wide range of fires. The denotation of a sign may range from particular to general, may change with circumstances, and vary for different individuals. For example, "the television program I watched last" may denote different programs to different persons at the same time and to the same person at different times.

9.4. Connotation of a Sign: the set of properties (of the objects or events responded to) which produce the response.

Thus the connotation of 'matches' may be their ability to light a cigarette or to produce damage to persons and property. Hence the connotation of a sign may also change with circumstances and vary for different individuals. For example, 'the last television I watched' may connote different properties (e.g., humor, drama, news, etc.) to different people, and to the same person at different times.

Two different signs may have the same denotation but different connotations: for example, 'mate' and 'housewife'. Conversely, two different signs may have the same connotation but different denotations; for example, 'matches' and 'lighter'.

'Denotation' and 'connotation' are sometimes used synonymously with 'meaning'. If it is so used it should be borne in mind that this is semantic meaning, not pragmatic. I prefer to use 'meaning' in its pragmatic sense, as will be apparent when I discuss this concept below.

This discussion of denotation and connotation may seem appropriate for signs which signify observable things or properties of such things. But how do these concepts apply to signs such as 'centaur', 'James Bond', 'snark', and 'angel'? Furthermore, many of the signs we use refer to things which may be or have been observable but which we have never observed; for example, names of historic figures or places that we have never visited. Both 'centaur' and 'Abraham Lincoln' fail to produce a response to a relevant observable object. They do, however, produce a response to an image or concept.* Even signs

*It will be recalled from the model of choice discussed in Chapter 1 and the discussion of memory in Chapter 4 that images and concepts could be called forth from memory and could be produced or modified by the observations of, or communications received by, the subject.

which signify observable things may produce responses to images or concepts brought forth from our memory. In some cases these were produced by the signified things when we observed them in the past. In others, they were produced by communication. A description can produce an image of even a nonexistent thing and an explanation can produce a concept of it.

We might argue that 'centaur' denotes the top half of a man joined to the body of a horse, or a drawing of a centaur. In a sense this is so. To one who has never seen a horse or a picture of a horse or a centaur, 'centaur' is unlikely to denote anything, but common sense indicates that the respondent to 'centaur' is not responding to the parts of a horse and man, but to an image or concept. The behaviorist or operationally oriented are disinclined to accept such mentallistic concepts and, indeed, it would defeat our purposes here if they were treated as such. But an operational definition of these notions is not impossible.

Images

Note that in common parlance an individual can have the following kinds of image: an image of (1) an object in the same environment (e.g., the chair behind me); (2) an object in a different environment (e.g., the chair in my bedroom); and (3) a non-existent object such as a centaur or James Bond. Hence our definition must be broad enough to cover all these possibilities.

The notion of an image has been a very fruitful one in the development of psychology. It has been used to explain our ability to use past experience in the present or to explain why different people react differently to the same stimulus. It was noted earlier that different observers may describe the same thing differently. This is frequently explained by saying that their images--mental pictures--differed, and it is their images, not the stimulating object, which they describe.

Such images were called <u>impressions</u> and <u>ideas</u> in early psychology.

A mental picture was once <u>assumed</u> to be one that no one could "see" but its possessor; hence, the earlier prevalence of introspective or subjective psychology. But this assumption is not necessary.

We call the capability of image-construction <u>imagination</u>, and we feel quite free to say of someone that he has or lacks imagination without asking him. On what evidence are such statements based? Somehow we base such statements on what we have observed. What observations were relevant?

Let us follow a common-sense account of how an image is formed. Suppose you are reading a story to a child and 'horse' appears. The child wants to know what a horse is. You may draw one, show a picture of one, or describe it. In so doing you create an image of a horse in the child. A verbal description of a horse may provide an image of a horse or reveal one, but it is not itself an image of a horse. It is not an image of the horse, as a picture is, because the signs used in the verbal description do not themselves have any of the relevant properties of a horse. The picture does. The picture is a sign of a horse which has some of the same (geometric) properties as that which it signifies. It <u>looks like</u> a horse.

9.5. <u>Iconic Sign</u>: a sign which has some of the same structural properties as the thing(s) which it signifies.

Structural properties, as noted earlier, include geometric, kinematic, physical, and morphological properties. Hence, iconic signs look, taste, feel, sound, or smell like what they signify, but they need not, and usually do not, function in the same way as that which they signify. Therefore, iconic signs not only signify, but they also <u>represent</u> what they signify and hence may substitute for it under some circumstances. A photograph which is a common type of iconic sign can frequently substitute for the person that it represents.

Iconic signs that sound like what they represent are called onomatopoeic; for example, 'bow-wow', 'meow', and 'cock-a-doodle-do'.

Iconic signs individuate; that is, they represent things or events taken as individuals, differentiated from other things. It is for this reason that we can have an image of a horse but not an image of animal. There is no set of structural properties which individuate animals; functional properties are necessary to do so. Or again, we can have an image of a pistol but not of weapon, because the individuating property of weapon is functional, not structural.

A physical image is an iconic sign. If it is an image of something that we have experienced (say, horse), then it facilitates our response to the verbal sign 'horse'. The image of a horse increases our probability of responding to the object horse. It is for this reason that books and lectures are so frequently illustrated.

Note that we can have a picture of a picture, and hence an image of an image.

Up to this point I have considered only images which can be seen, heard, felt, smelled, or tasted: physical images, and images of things which exist. Let us now return to the centaur and images of things which do not exist.

The image of a centaur combines physical properties of man and horse into an individual thing. We have experienced each of the properties involved but not their combination. The image, then, is a combination of properties. If this combination is represented by iconic signs these signs are a physical image. But the combination of properties is itself an image whether or not it is represented physically.

Images which are not physically represented are called _mental_ and, as indicated earlier, are assumed to be inacessible to all but him who has it. We can now see why this assumption is false. The combination of properties that form an individual's image (is "in his mind") is the same combination that coproduces his responses to non-iconic signs. By observation and analysis of his response we can determine what his "mental" image is. Therefore, a mental image is the collection of structural properties and the relationships between them to which an individual responds. Such images intervene between the sign and that which is signified, even when it exists. When it is an image of the real thing, that thing is at least a coproducer of the image. When it does not exist, the image is produced by signs.

> 9.6. _Image_: an individuated set of structural properties and the relationships between them to which a subject responds.

Concepts

The difference in common usage between 'image' and 'concept' suggests how to define the latter. First we note that concepts are not iconic; they do not look like, sound like, . . . , what they signify. Secondly, whereas images help us _describe_, concepts help us _explain_. Herein lies the critical difference. _Images connote structural properties but concepts connote functional properties._

Explanations are of two sorts: (1) we explain how something comes to be; that is, we identify that which produced it. For example, we explain the presence of a strange piece of furniture in our home to a friend by saying, "It was a gift." (2) We explain a thing by identifying what it can do; that is, what its function is. For example, we explain a _Clipit_ by "It is used to cut clippings from a newspaper without damaging the sheets below the one being cut."

Hence, to say that we cannot conceive of a particular thing is to say that we can't explain it: either we do not know what could have produced it or we can't determine what it can do, or both.

The definition of a sign developed earlier signifies the author's conception of a sign: that combination of functional properties of objects or events which explain a particular phenomenon of communication.

> 9.7. Concept: an individuated set of functional properties and the relationships between them to which a subject responds.

To have an image of _Y_ and to have a concept of _Y_ are not the same thing. We can, for example, have an image of something but not a conception of it. A child may have an image of God but no conception of Him. Conversely, we may have a concept of something of which we have no image. An adult may have a concept of God but no image of Him. Models of reality are either images, conceptions, or some combination of these. A model is a representation of those structural and/or functional properties of reality which the subject believes to exist and to be relevant to his purposes.

Symbols and Signals

Both 'symbol' and 'signal' are types of signs, but semioticians seem to agree on little more than this with respect to them. There are several different meanings associated with these terms each of which seems to be justified by common usage. One of these meanings of 'symbol' is put forth in the following quotation from Suzanne Langer (1948):

> Instead of announcers of things, they [symbols] are reminders. They have been called "substitute signs," for in our present experience they take the place of things we have perceived in the past, over even things that we can merely imagine by

combining memories, things that <u>might</u> be in past and present experience (p. 24).

... it is the conceptions, not the things, that symbols directly mean (p. 49).

Hence, according to Langer, a symbol is a sign that signifies a concept. This is certainly one way in which 'symbol' is commonly used. For example, it is in this sense that the American flag is a symbol of our concept of our nation, and a skull and cross bones are a symbol of our concept of death. Symbols, in this sense, are frequently, but not necessarily, natural or non-linguistic signs. Proper names can also be symbols in this sense; for example, 'Abraham Lincoln' is a symbol of honesty. But 'honesty' itself signifies a concept. It too would be a symbol in Langer's sense. This seems to me to be too general an applicability.

'Symbol' is also commonly used in another sense, particularly in logic and mathematics, but also in more commonplace activities. For example, '+', '=', and '>' are commonly called symbols in arithmetic, and '$', '%', and '&' are commonplace symbols. In what sense is '+' different from 'plus'? Most would answer that it is just a convenient "short-hand" for 'plus'. It is this sense of 'symbol' which Morris (1946) used when he defined a symbol as a sign "that is produced by its interpreter and that acts as a substitute for some other sign with which it is synonymous" (p. 355). Hence, for Morris, a symbol is a sign of another sign that is produced by the same person who responds to it. "Where an organism provides itself with a sign which is a substitute in the control of its behavior for another sign, signifying what the sign for which it is a substitute signifies, then this sign is a <u>symbol</u>..." (p. 25).

It does not seem to me that a sign can serve as a symbol only to the one who produced it. In some sense '+' is as much a symbol to you in an equation that I write as is one which you write. When you read "Let $\underline{P_i}$ represent the probability of selecting a course of action $\underline{C_i}$," '$\underline{P_i}$' and '$\underline{C_i}$' become symbols for you as well as for me. Nevertheless, it is clear that we use 'symbol' in the sense of a substitute for other signs.

Langer used 'symbol' as a sign of a concept; Morris as a sign of a sign. It seems reasonable to ask whether some have not used 'symbol' as a sign of an image. Obviously they have. For example, characatures are frequently used symbolically.

'Symbol', therefore, seems to be used as a sign of an image, concept, or another sign. Now images, concepts, and signs all have a common property: each represents something other than itself; that is, they can produce responses to something other than themselves. This suggests a definition of 'symbol' which synthesizes at least several of its common uses:

> 9.8. Symbol: a sign which is a potential producer of a response to something which in turn is a potential producer of a response to something other than itself.

Signal. Morris (1946) defined a 'signal' as "a sign that is not a symbol" (p. 354). This definition, it seems to me, completely misses the usual sense in which 'signal' is used. For example, in Webster's Seventh New Collegiate Dictionary (G. and C. Merriam, Springfield, Mass., 1963), 'signal' is defined as "an act, event, or watchword that has been agreed upon as the occasion of concerted action" or "a sound or gesture made to give warning or command."

Signals, I believe, are intended to initiate or terminate action. This is certainly true, for example, of a traffic signal which "starts" and "stops" us. But a traffic sign (e.g., a stop sign) may also stop us. The difference between a traffic sign and a traffic signal, I believe, holds the clue to the essential difference between 'sign' and 'signal'. A signal is always the behavior of an object, an act or event; a sign need not be. For example, a constantly red light is a "sign" of danger, but a traffic light that changes its color is called a "signal."

Putting these observations together yields the following definition:

> 9.9. Signal: an act of a purposeful individual (or individuals) or of an object whose behavior is produced by such an individual (or individuals), which serves as a sign of that individual's intention that other purposeful individuals or himself (themselves) respond by behaving in a specified way at the time of the act.

Note that an individual can signal himself as in setting an alarm clock or in arranging to be called in a hotel at a certain time in the morning. A traffic signal may be manually operated by a policeman or set by him so that it operates itself in a desired way. Even in the latter case its behavior is produced by the one who set it.

PRAGMATICS

Pragmatics, it will be recalled, is concerned with the relationship between a sign, its source, and/or its respondent. My concern here is with the respondent. In Chapter 10 I will consider the interaction of the source and the respondent.

Meaning

'Meaning' has been used in so many different senses that some, like Morris (1946), exclude the concept from consideration. Cherry (1957) observed, "There is a move today to avoid 'meaning' so far as can possibly be done, in communication studies" (p. 111). He goes on to cite ten different meanings of 'meaning' (pp. 112-113). Despite the caution of Morris and Cherry, the analysis of meaning, largely stimulated by the work of Ogden and Richards (1947, originally published in 1923), continues into the present. See, for example, Loundsbury (1966) and Osgood _et al_ (1957). A recent review of the literature on meaning by Marjorie B. Creelman (1966) reaches the following conclusion:

> ... meaning, the elusive Cinderella, is still at large, evading identification and capture. Perhaps one of the difficulties lies in the various images that her various suitors have of her--images that have led them to seek her different ways. Some see her as a simple-minded creature, some as complex, subtle, and sophisticated. Some have focused on her intellectual qualities, and some imagine her to be sensitive and emotional. Some, overwhelmed by her mystery, have from the first contented themselves with living with her only in fantasy, concluding that she is essentially unknown and unknowable (p. 207).

I have already pointed out that 'meaning' is sometimes applied in the semantic context to refer to what I have called the _denotation_ and _connotation_, or the _signification_, of a sign. I see no good reason for using 'meaning' in this context where we already have complete and adequate terminology; it would only introduce unnecessary redundancy. This is not so in pragmatics where there would be a conceptual and terminological gap if meaning were not considered.

Pragmatic philosophers from Peirce to Dewey have pointed out that in practice the meaning of a term does not lie in what it comes from, but in what it leads to; or, as they put it, in the difference it makes in the respondent's behavior. A sign which does not affect behavior has no meaning, no matter what it signifies. Thus the meaning of a sign lies in what it can make one do. For example, when one cries "Fire!" in a crowded theater, the meaning of the cry is not to be found in the flames denoted or the heat connoted, but in the effort to escape harm or avoid destruction that it produces. In effect, meaning, though a function of what a sign signifies, is separate from it; it lies not in the signification of a sign, but in its _significance_.

Furthermore, 'meaning' is not only applicable to signs but also to any experience or thing that is experienced. All things which act as signs have meaning, but not everything with meaning is a sign. For example, one asks of an event, "What does it mean?" This is equivalent to asking, "What will it lead to?" or "What significance does it have?" In this sense one can, and has, asked about the meaning of life itself. When one is asked, for example, what television means to them, they are likely to refer to entertainment, keeping informed about world affairs, and perhaps even education. They do not define television but reveal its significance to them. Meaning is not captured in definitions; signification is.

Cherry (1957) has observed that

> ... the meaning of the utterance to the listener, B,' is the selection of the particular response he actually makes; and that, 'the meaning of the utterance to the speaker, A, ' is that selection of a response in B which A intends his utterance to evoke (p. 114).

This concept of meaning was also presented by Ogden and Richards (1947), and much earlier by Gardiner (1921-22).

Meaning is a property of a purposeful response to a stimulus. It is quite naturally attributed to the stimulus because the stimulus produces it.

> 9.10. Meaning. The meaning of a stimulus (sign or otherwise) is the set of functional properties of the response which it produces.

Therefore, a sign may have different meanings for different individuals, or different meanings for the same individual at different times. For example, Paul Revere's cry, "The British are coming!" had one meaning for the American Revolutionists, and another for the Tories. Or again, "No rain is expected today" may mean one thing to a farmer at work, but another when he is vacationing. Its meaning may also change with the seasons. As Cherry (1957) noted, "A 'meaning' is not a label tied around the neck of a spoken work or phrase. It is more like the beauty of a complexion, which lies 'altogether in the eye of its beholder' (but changes with the light!)" (p. 115).

Since the functional properties of different responses to the same sign may differ, the only meaning that 'the meaning' of a sign can have, lies in a common functional property of these different responses. That is, we may find a more general function which persists among responses that are functionally different at a lower level of generality. For example, in the wide variety of responses to "It will not rain today" we are likely to find a common functional property such as the shedding of protective cover or increased outdoor activity. But even in this sense it may be unlikely that we can find any one meaning for any sign.

For those who prefer to use 'meaning' in a different way than I have, I am willing to qualify my use by referring to it as <u>pragmatic meaning</u>.

Morris (1964) discusses three types of signification (i. e., semantic properties) of signs and three corresponding "dispositions to react in a certain way" (i. e., pragmatic properties). Semantically speaking a sign, for Morris, is

(1) <u>Designative</u>, "insofar as it signifies <u>observable</u> properties of the environment or the actor" (e. g., 'black').

(2) <u>Prescriptive</u>, "insofar as it signifies how the object or situation is to be reacted to so as to satisfy the governing impulse" (e. g., 'ought').

(3) <u>Appraisive</u>, "insofar as it signifies the consummatory properties of some object or situation" (e. g., 'good') (p. 4).

The corresponding (pragmatic) functions are to produce

(1) "a disposition to react to the designated object as if it had certain observable properties,"

(2) "a disposition to act in a certain kind of way to the designated object or situation," and/or

(3) "a disposition to act toward a designated object as if it would be satisfying or unsatisfying" (p. 6).

There is a considerable correspondence between these three functions that Morris identified and the three discussed in the preceding chapter: (1) information, (2) instruction, and (3) motivation. Furthermore, it is apparent that signs which are designative in Morris's sense, inform in my sense; those which are prescriptive, instruct; and those which appraise, motivate.

One could pursue such an analysis of the signification and functions of signs considerably further, but since individual signs seldom function independently of other signs, it seems more fruitful to discuss the properties of sign-combinations; that is, messages. The signification and meaning of a message is never the simple sum of these properties of the component signs; it is a resultant of a considerable interaction between the individual sign-properties. For example, consider the difficulty of translating a message in an unfamiliar language with only the help of a dictionary.

In sum, the signification and significance of a sign depend on the sign environment and the situation in which it is used. This is why a dictionary must give so many different definitions of most signs.

SIGN MEASURES

Up to this point I have dealt with only the qualitative aspects of signs. Now I consider some of their quantitative aspects. Here too it is convenient to distinguish between measures that are semantic and pragmatic.

Semantic Efficiency and Related Measures

9.11. Semantic Efficiency of a Sign in an Environment (*S*): the probability that the sign produces a response in that environment by the receiver to that which the sender intended him to respond.

If the sign represents objects or events, then the semantic efficiency is equivalent to denotative efficiency; if the sign represents properties, then its semantic efficiency is equivalent to connotative efficiency.

Ambiguous Signs. Suppose an individual is told to get 'the book' off a table on which two books are located. He may get either or both, but he is not likely to pick up a pen rather than a book. In this context 'the book' is ambiguous because it has denotative efficiency for more than the item intended. If the instruction had been to get 'one or the other of the books' or 'the larger book' the ambiguity would be removed. The receiver in the first situation may seek to remove the ambiguity by asking, "Which one?"

The nature of ambiguity, then, lies in the discrepancy between the intended response to a sign and the actual response. The ambiguity exists for the receiver relative to the sender. A receiver may deliberately misinterpret the sender's intention; for example, he may bring a pen in order to annoy the sender. This, however, is not a case of ambiguity. Ambiguity implies that the receiver desires to cooperate with the sender.

9.12. Ambiguity. A sign (*X*) is denotatively or connotatively ambiguous if (a) the sender intends *X* to denote or connote something (*Y*), (b) *X* is an efficient denoter or connoter of something other than Y for the receiver, and (c) the receiver intends to respond to the denotation or connotation that is intended by the sender.

That ambiguity is not always undesirable becomes apparent when we realize that it is one of the most important instruments of the verbal artist. The richness of poetry, for example, lies in the fact that it has many different denotations and connotations. The ability of ambiguity to stimulate imagination was exploited by James Joyce in Finnegan's Wake. In this work Joyce invented words which deliberately have several denotations and connotations; for example, 'Wellingdome Museyroom' has many more connotations than 'Wellington Museum Room'. Puns, of course, are a deliberate manipulation of signs to obtain ambiguity, to give signs more than one signification.

The measure of semantic efficiency given above is clearly relative to the intended signification (Y), the environment in which the sign operates (S), and the respondent (A). Now we can determine how the efficiency of a sign depends on these three variables: Y, S, and A. This dependence reflects on the semantic generality of a sign. Consideration of its sensitivity to (1) the Y which is signified leads to a definition of signification-generality; the environment, S, to a definition of environmental generality; and (3) the respondent, A, to a definition of social generality.

Signification Generality. The word 'chair' usually produces a response to only a few objects in a normal room. The word 'furniture' usually produces a response to a wider range of objects than does 'chair'. Hence, 'furniture' has a more general denotation than does 'chair'. It also has a more general connotation because the properties of furniture include, but are not exhausted by, the properties of chairs.

9.13. Signification-Generality. If the things signified by one sign, $\underline{X}_1$, include all the things signified by a second sign, $\underline{X}_2$, $\underline{X}_1$ is more general (denotively, connotatively, or both) than is $\underline{X}_2$.

We can have a hierarchy of signs relative to the generality of their signification. 'Furnishings' is more general than 'furniture', and 'furniture' is more general than 'chair'.

If their is no overlap of the classes of things denoted by two signs, then the only basis for comparison is the number of things signified. This criterion by itself, however, is not very useful. For example, it serves no useful purpose to assert that 'horse' is more general than 'buffalo' because there are more horses than buffalo.

The signification-generality and ambiguity of a sign are not to be confused. A general sign may denote a large number of different things, but it is intended to do so. An ambiguous sign denotes more than it is intended to. Where the intention is that an individual respond to many objects and he does, the sign is not ambiguous, though general. Therefore, 'books' is a more general sign than 'novel', but it may be less ambiguous.

Environmental-Generality. As we have already indicated, a sign may have different denotations or connotations in different environments. For example, the denotation of 'the man on my right' changes from time to time; it therefore has less denotative reliability than does the name of the man. Yet 'the man on my right' will usually produce responses to the man in the same relative position and hence is connotatively reliable, at least with respect to the property position.

9.14. Environmental-Generality of a sign relative to a particular signification ($\underline{Y}$), one or more receivers, and an exclusive and exhaustive classification of environments, is the fraction of this set of environments,

in which the sign produces responses to Y
in the relevant receivers.

Social-Generality. Finally there is the measure that reflects the number of people for whom a sign signifies the same thing under the same set of conditions.

9.15. Social-Generality of a sign relative to a particular signification (Y), a set of receivers, and a specified set of environments, is the fraction of the set of receivers in which the sign produces responses to Y in the relevant environments.

Using the concept of social-generality of a sign, two other important sign characteristics can be defined.

9.16. Obscure Signs: ones which have a low denotative or connotative efficiency relative to any possible denotation or connotation for most but not all of the members of a social group.

The degree of obscurity is simply the fraction of the group's members for whom the sign is semantically inefficient. Thus archaic words (e.g., 'ere' and 'perchance') are usually called obscure because few people know what they are intended to signify.

9.17. Esoteric Signs: ones which are obscure to members of one subgroup of a population but efficient when used on members of another, and the second group has a common set of objectives not shared by members of the first subgroup.

Thus 'homoscedastic', which is an efficient signifier among mathematical statisticians but not among others, is an esoteric sign. The jargon of special interest groups usually consists of esoteric signs.

<u>Pragmatic Efficiency and Related Measures</u>

9.18. <u>Pragmatic Efficiency of a Sign in an Environment</u> (S). is the probability that the sign produces a response in that environment by the receiver that was intended by its source.

It is apparent that by an extension of the discussion of semantic efficiency we can define three types of pragmatic generality: <u>response</u>, <u>environmental</u>, and <u>social</u>. Since the extension is straightforward it is omitted here.

SIGNS WHICH AFFECT OTHER SIGNS

As noted earlier, signs are normally used in sign-complexes. In such complexes the signs interact. Some signs have a particular role to play in unifying the signs in the complex. These signs have the function of affecting other signs either by <u>modifying</u> them, <u>relating</u> them, <u>connecting</u> them, or <u>emphasizing</u> them. It is to these special sign-roles that I now turn. (The discussion which follows relates to that of forms of statements which appeared in Chapter 4.)

<u>Modifiers</u>

9.19. <u>Qualifier</u>: a sign which produces a change in the connotation of another sign.

Hence a qualifier attributes a property to that which is denoted by another sign and puts what is denoted into a class of things having the attributed property. For example, in 'red book', 'red' qualifies 'book' and directs the response to the book to its redness. Note that in 'The book is red', 'is red' serves the same function. Adjectives, of course, normally qualify nouns. Adverbs similarly qualify verbs.

Qualification may individuate that which is modified; that is make the denotation more specific and remove ambiguity. This follows from the fact that a sign which changes the connotation of another sign may also change its denotation. 'Red book' and 'blue book' have different denotations. A qualifier may change the connotation of a sign, however, without affecting its denotation. For example, 'the room in which I am working' and 'the reading room in which I am working' have the same denotations but may have different connotations.

> 9.20. Quantifiers: signs which affect the number of things denoted by other signs.

Some examples are 'all', 'few', 'four', 'many', and so on. Note that whereas 'four' in 'four books' quantifies, 'fourth' in 'fourth book' qualifies since it signifies a locational property.

Relators

> 9.21. Relators: signs which relate the signification of one sign to that of another.

They may do this by attributing a property to the signification taken collectively. For example, in 'John is the brother of Tom', 'is the brother of' relates John and Tom. It attributes a property to the pair, a property that cannot be attributed to either member taken separately. Relators may also signify the similarity or difference between the signification of two signs; for example, 'John is younger than Tom'. The difference can be quantified as in 'John is five years younger than Tom'. Of course more than two things can be related as in 'John is the brother of Tom and Mary'.

Connectors and Disconnectors

> 9.22. Connectors (Disconnectors): signs which combine (separate) the signification of two or more other signs.

In 'John and Mary are at home' the 'and' is used to produce a response to the joint presence of John and Mary, rather than a response to either taken separately. This expression may have a different connotation than 'John is at home. Mary is at home.' It is this difference in connotation which 'and' signifies.

The role of connectors and disconnectors is most apparent in mathematical expressions. For example, we readily recognize the difference between '4 + 2' and '4 - 2' and between '(3x2) + 2' and '3x(2+2)'. Verbally we get the same results by using 'and', 'or', 'plus', and so on. Punctuation marks such as the comma, colon, semicolon, and hyphen serve the same purpose.

Sometimes proximity of signs is sufficient to connect them. For example, we may either say 'nice and big toy' or 'nice big toy'.

Emphasizors

9.23. Emphasizors (De-emphasizors): signs which produce an increase (decrease) in the probability that an individual will respond to a signifier.

In writing, for example, a word or passage may be called to one's attention by italicizing it, or by changing the type in which it is printed, or the color of the type, and so on. In speaking changes in intonation or repetition have the same effect. On the other hand, smaller type or a drop in one's voice can be used to de-emphasize a sign or a sign-complex, as in a footnote or an aside.

It should be noted that things which modify, relate, connect, and emphasize other signs are themselves signs. They either produce responses to other signs or affect their signification. Hence, they signify either the change in signification that they produce or the intention of the source that more or less attention be given to other signs.

MESSAGES

9.24. Message: a set of one or more signs intended by its producer to produce a response either in another or himself.

One can, of course, send a message to oneself; for example, a reminder entered on a calendar. Further, a message can be sent without the use of words, by gestures. However, messages are normally formed out of linguistic signs. It is necessary, therefore, to understand the nature of language if one wants to understand fully the nature of messages.

LANGUAGE

Not all signs are part of a language. For example, smoke may be a sign of fire but it is not an element of a language. The signs which form a language are ones which can be produced by purposeful individuals. Hence the word 'smoke' is an element of our language.

Linguistic signs must satisfy other conditions than that of being produced by purposeful individuals. They must be semantically and pragmatically efficient for a significant portion of the people who use them. Otherwise they could not be used in communication. This efficiency must pertain over a wide range of environments. Hence linguistic signs must be environmentally and socially general in both the semantic and pragmatic sense.

The set of individuals relative to which linguistic signs must have these properties is the set for which they are to serve as a language. Languages are "relative" in the sense that what constitutes a language for one set of individuals may not be so for another.

Finally, there must be more than a set of signs to form a language; there must also be a set of rules for combining signs into groups in such a way that the resulting sign-complexes have the same properties required of linguistic signs. These rules specify the form that linguistic expressions should take and how the resulting expressions should be interpreted. The rules of our language, for example, allow us not only to form 'dog bites man' and 'man bites dog', but also to interpret these same combinations of words differently. We do not know how to interpret sign-complexes which do not satisfy these rules. For example, if I rearrange the words in the last (italicized) sentence in a randomly selected sequence, I get 'complexes satisfy we do these which not do interpret rules to how sign not know'.

Linguistic rules are what Morris would call prescriptions and what I call instructions; that is, they are messages which increase the efficiency with which we can communicate. Of such rules Cherry (1957) observed:

> Human languages have an excess of rules, so that some can be broken without serious harm. The rules we call grammar and syntax are not inviolate, but the more we break them, the lower are our chances of successful communication (p. 19).

The rules of a language have two sources: common usage and experts. In The American Language, H. L. Mencken described the way Americans actually do use and combine signs. The experts--those who propare dictionaries, write "grammars," and teach the language professionally--prescribe what signs ought to be used and how. The 'ought' derives from their beliefs about the communicative efficiency of alternative ways of using linguistic signs. The experts and common usage frequently do not agree. They "battle" in the classroom and the streets; sometimes one wins, sometimes the other.

Summarizing, then, the following definition can be formulated:

9.25. Language: a set of signs and instructions for their use such that (1) the signs can be produced by purposeful individuals, (2) they are semantically and pragmatically efficient for a significant portion of those who use them, (3) they are environmentally and socially general in the semantic and pragmatic sense, and (4) the instructions signify ways of permuting and combining signs in the set to form sign-complexes which also satisfy conditions (2) and (4).

It is not possible to specify how efficient and general the signs must be over what portion of the population before a language can be said to exist. By complex social processes languages grow, evolve, and change in many ways; less efficient and less general signs are dropped (e.g., archaic ones) or modified, and new ones are added (e.g., 'turbo jet' and 'transistor'). Languages can be created _de novo_ as Esperanto was in the last century and as such computer languages as FORTRAN, COBOL, and ALGOL have been only recently. One person can create and use a language for his own purposes. Languages need not be social instruments, but they usually are. Clearly communication between people is greatly facilitated when they share a language, but it is not precluded when they do not share one, as many who have travelled to foreign countries know.

CONCLUSION

Up to this point I have considered only the elements of communication: the _material_ out of which communications are made. In the next chapter, I take up the process of communication; that is, how signs, messages, and language are used.

BIBLIOGRAPHY

Cherry, Colin, On Human Communication. John Wiley and Sons, New York, 1957.

Creelman, M. B., The Experimental Investigation of Meaning. Springer, New York, 1966.

Gardiner, A., "The Definition of the Word and the Sentence," British Journal of Psychology, 12, Part 4 (1921-1922), 352-361.

Langer, Suzanne, Philosophy in a New Key. Penguin Books, New York, 1948.

Loundsbury, F. G., "The Varieties of Meaning," in Communication and Culture, ed. by A. G. Smith. Holt, Rinehart and Winston, New York, 1966.

Morris, C. W., Foundations of the Theory of Signs. International Encyclopedia of Unified Science, 1, No. 2. University of Chicago Press, Chicago, 1938.

______, Signs, Language, and Behavior. Prentice-Hall, New York, 1946.

______, Signification and Significance. The M. I. T. Press, Cambridge, 1964.

Ogden, C. K., and L. A. Richards, The Meaning of Meaning. Harcourt, Brace and Co., New York, 1947.

Osgood, C. E., G. J. Suci, and P. H. Tannenbaum, The Measurement of Meaning. University of Illinois Press, Urbana, 1957.

CHAPTER 10

MODELS OF COMMUNICATION

NOISE, _n_. A stench in the ear. Undomesticated music. The chief product and authenticating sign of civilization (Ambrose Bierce, _The Devil's Dictionary_).

INTRODUCTION

One of the simplest acts of communication occurs when one individual, A, communicates to another, B, about something, _X_, and receives no reply. This is _one-way_ communication. Following T. M. Newcomb (1966) I represent such communication by "AtoBreX." (The discussion that follows was greatly stimulated by Newcomb's work.) If B replies, we have _two-way_ communication.

ONE-WAY COMMUNICATION

A and B may be the same person; for example, when one writes a memorandum to oneself. None of the discussion that follows requires that A and B be different. Instances in which they are the same party are obviously special cases of the more general two-party case that I will consider in detail.

A and B need not be in the same physical environment (e.g., A may phone or write to B in another city), nor do they have to exist at the same time (e.g., Plato communicates with me when I read him today). Of course, I cannot communicate with Plato; hence, our communication is one-way. One-way communication can take place between two contemporaries, as when I read a living author or listen to a lecture or broadcast.

The Subject Communicated

A state of communication can be divided into three parts: the sender (A), the receiver (B), and all other things that affect the communication (Z). Therefore, the subject (X) about which A intends to communicate to B may be himself (A), the receiver (B), something else (Z), or some combination of these. The possibilities are shown in Table 10. 1.

TABLE 10. 1 POSSIBLE SUBJECTS OF COMMUNICATION

The subject of the message (X) is about:	Example
1. A alone : AB'Z'	I'm tired today.
2. B alone : A'BZ'	You're tired today.
3. Z alone : A'B'Z	He is tired today.
4. A and B : ABZ'	You and I are tired today.
5. A and Z : AB'Z	He and I are tired today.
6. B and Z : A'BZ	You and he are tired today.
7. A, B, and Z: ABZ	He, you, and I are tired today.

The production of a message (M) by A is itself a product of something (Y) to which A responds. The producer of the message (Y) may or may not be the same as its subject (X). For example, seeing a friend (Y) may lead me to ask about his health (X).

Some messages do not seem to be about anything. For example, in passing an acquaintance whom one passes often there is usually a ritualistic exchange such as:

"Hello."

"Hello. How are you?"

"Fine. And you?"

By this time the two are too far apart to hear each other, but it does not matter because the exchange has served its purpose. It is clear that such an exchange is purposeful; what is not clear is just what that purpose is.

In such exchanges each party observes the presence of the other and so indicates by his remarks or gesture. In addition such communication usually signifies recognition of the other. If a passing stranger says "Hello" we may respond to avoid hurting his feelings, but we wonder why he addressed us. However, if a person that we know well fails to say "Hello" to us in passing we may either think that he did not see us, or that he did and is snubbing us. Hence, the remark made in passing and much of what we call "small talk" signifies recognition of the other and his significance to the sender.

The failure to so communicate under certain circumstances established by custom may produce a change in one's attitude toward another. Hence such communication, as a minimum, produces a non-change in the attitude of the receiver toward the sender. Under some circumstances two strangers who do not communicate do not offend each other; for example, on a subway train. Under other circumstances offense might be taken; for example, at a party.

The subject of such communications, then, is the <u>relationship between the communicators.</u> Gregory Bateson (1966) called such an exchange <u>metacommunication.</u> He commented on it as follows:

> When A communicates with B, the mere act of communicating can carry the implicit statement "we are communicating." In fact, this may be the most important message that is sent and received. The wisecracks of American adolescents and the smoother but no less stylized conversation of adults are only occasionally concerned with the giving and receiving of objective information; mostly, the conversations of leisure hours exist because people need to know that they are in touch with one another.
>
> Similarly, every courtesy term between persons, every

> inflection of voice denoting respect or contempt, condescension or dependency, is a statement about the relationship between the two persons (pp. 425-426).

Bateson identifies another form of metacommunication: communications about communication. For example,

> Such a statement as "The word 'cat' stands for a certain small mammal" is neither true or false. Its truth depends upon agreement between the speakers that it be true. In terms of such agreement they understand each other: or where disagreement occurs they will meet with misunderstanding. And this statement about the word 'cat' is only one of a vast category of statements about codification, which category ranges all the way from the conventions of local phonetics up through the conventions of vocabulary to the conventions of syntax... (p. 425).

Communication that is about the relationship between the communicators is at least motivational since it affects the attitudes and feelings--and, hence, intentions--of the parties involved. Communication which is about communication may be either informational or instructive since it may effect the choice of signs and messages or the way they are used.

The professional or amateur entertainer or performing artist is not usually concerned with informing or instructing members of his audience. Such communication may be directed toward making the receivers "forget their troubles"; that is, to produce greater satisfaction in the receiver with his present state. Hence, such cathartic or recreational communication is motivational.

Some estheticians have argued that the great tragic dramas do not produce satisfaction with one's current state, but rather produce dissatisfaction and move one to action. Hence, such communication is not recreational or cathartic, but is intended to stimulate if not inspire; nevertheless, it is also motivational. Aristotle emphasized the cathartic function of art, Plato the stimulative. Some estheticians find both functions in art.

Although estheticians may disagree on the function of art, they generally agree that it affects feelings, and hence is motivational. Not only does art affect feelings, but more often than not its subject matter is feeling itself.

Producers of Messages

A message from A to B about X is very likely to be affected by the following properties of A, some of which are in turn produced by B:

1. A's beliefs about X: AbX.

   What structural and functional properties A believes X to have; that is, his image and concept of X.

2. A's attitude toward X: AaX.

   What A feels about X: the value he places on it.

3. A's belief about B: AbB.

   In particular, how A believes B will respond to possible messages from A about X. This, in turn, probably depends on what A believes the following properties of B to be:

   a. B's beliefs about X, BbX: Ab(BbX).
   b. B's attitude toward X, BaX: Ab(BaX).
   c. B's beliefs about A, BbA: Ab(Bba).
   d. B's attitude toward A, BaA: Ab(BaA).

4. A's attitude toward B: AaB.

These properties are, in all likelihood, interdependent. Any or all of them may co-produce the message that A sends to B about X. It follows from the definitions that the messages which A sends to B about X are ones that A believes will produce or maintain the beliefs and attitudes in B toward A and X that A intends B to have.

Newcomb implied that knowledge of these attitudes and beliefs and the environment in which communication takes place is sufficient to predict and/or explain the communication that takes place between two parties. Research designed to test this implication is described in detail in Appendix III. The test consists of an effort to predict and explain behavior in two-person conflict games.

Noise

The message that A sends to B about X may differ from the message B receives from A about X. These may differ structurally or functionally. For example, a vocal message over the telephone may be distorted, cut-off, or obscured by noise. A printed message may be smeared or torn. A television picture may be obscured by "snow." In each of these cases the message received is structurally different from the message sent. Anything which alters the structure of the message produces syntactic noise.

> 10.1. Syntactic Noise: any structural difference between a message that is sent and the message that is received.

Even if a message is not changed structurally it may not be received (i.e., interpreted or decoded) as it was sent. For example, what is intended as a compliment by A may be interpreted as an insult by B: "You look so much younger than you are."

> 10.2. Semantic Noise: ambiguity in the denotation or connotation of a message.

A message may be misinterpreted--that is, B responds to the "wrong" thing--and still produce the type of response intended. For example, A may be annoyed by "noise" he believes is caused by a radio and tell B, "Shut that thing off." B may turn off the television set which is actually causing the noise.

10.3. Pragmatic Noise: anything which appears in a message or its environment that was not produced by the sender and which decreases the probability that the receiver will respond in the way intended by the sender.

Hence, syntactic noise may not produce pragmatic noise; however syntactically noisy a message may be it may be received correctly and responded to as intended. On the other hand, a syntactically noise-free message may fail to produce the desired response because something diverts the attention of the receiver. Furthermore, as mentioned above, a receiver may respond to a message as intended even if it is ambiguous and hence full of semantic noise. Syntactic and semantic noise may produce pragmatic noise, but need not necessarily do so. (For an experimental situation in which it does, see Heise and Miller, 1966.)

10.4 The Amount of Pragmatic Noise in a message received is the difference between the probability that the message sent will produce the sender's intended response by the receiver and the probability that the message received will produce that response.

This measure can vary from +1 to -1. A negative measure indicates that the "interference" has enhanced the sender's chances of success. For example, this may occur when a message that is sent in a language not understood by the receiver is translated into a language that he does understand. Unfortunately we do not have a term which signifies negative noise.

The Receiver's Effect on a Message

The response to a message that B receives from A about X is a product not only of the message that A sent, but also of some of B's properties:

1. B's beliefs about X, BbX.
2. B's attitude toward X, BaX.
3. B's beliefs about A, BbA.
4. B's attitude toward A, BaA.

Note the similarity to A's relevant beliefs and attitudes.

If A combines his relevant beliefs about B, the environment, and the medium through which he communicates to B, into a model which predicts what message B will receive and how he will respond to it, given the message A has sent, then A can use this model to formulate his message effectively. To take a simple case, if A knows that B will only receive every other word of a message, he can obviously formulate the message so that when received it is what he intends. In more complex cases A can use his knowledge of how B usually responds to various types of messages to frame a message whose chances of producing the intended response are high; for example, knowing what form of request a person is most likely to respond to. Parents frequently tell children not to do what they want the youngsters to do because they believe a "negative" response is more likely than one that is "positive."

Redundancy

If A has doubts about either the message that B will receive or how it will be interpreted, he may repeat the message or send it in several different forms which he indicates are intended to be equivalent. This allows the receiver to select the alternative that is least ambiguous to him. Expressions starting with "that is," "i.e.," "in other words," and "put another way" have this function. They provide deliberate redundancy in the message.

Like other concepts in communication theory (e.g., noise and information) redundancy can be dealt with at either the syntactic, semantic, or pragmatic level, at each of which it has a different meaning.

Syntactic redundancy reflects the lack of randomness in the selection of signs, symbols, or messages. For example, most persons can correctly supply the missing letter in "Q- ICK:" U.". The U is therefore redundant because there is relatively little, if any, free choice involved in its selection. Similarly, a message that begins with "A stitch in time" does not have to be completed for many because they know what follows. Warren Weaver (1966) has put it as follows:

> Having calculated the entropy (or the [syntactic] information or the freedom of choice) of a certain information source, one can compare it to the maximum value this entropy could have, subject only to the condition that the source continue to employ the same symbols. The ratio of the actual to the maximum entropy is called the relative entropy of the source. If the relative entropy of a certain source is, say, eight-tenths, this means roughly that the source is, in its choice of symbols to form a message, about 80 percent as free as it could possibly be with these same symbols. One minus the relative entropy is called redundancy. That is to say, this fraction of the message is unnecessary in the sense that if it were missing the message would still be essentially complete, or at least could be completed (p. 21).

Syntactical redundancy can overcome the effects of syntactical noise. A. G. Smith (1966) points this out as follows:

> Redundancy ... improves the accuracy with which signals are transmitted ... Redundancy is the repetition of a signal that ... helps overcome noise.
>
> If the same signal is simply repeated over and over again, the redundancy is 100 percent. There is no variability or indeterminacy at this high degree of redundancy. The receiver can predict with confidence what the next signal will be. This means ... that the signal has no surprise

> and carries no new information. There is too much redundancy for communication. Zero percent redundancy leaves the receiver with sheer unpredictability--the next signal can be anything. At this low degree of redundancy the receiver cannot tell what is noise and what is information. The fact is that communication requires a balance between the predictable and the unpredictable (p. 365).

Semantic noise and redundancy have not been treated as extensively as have their syntactic counterparts. Macy, Christie, and Luce (1966) provide one of the few discussions of these concepts that I have seen. They treat semantic (or coding) noise much as I have: as ambiguity (which, of course, has been discussed extensively, but not as it relates to noise). Semantic redundancy, then, arises from the use of synonyms. The more "extra names" for the same thing that are used or remembered, the greater the semantic redundancy. The experiment reported by Macy _et al_ (1966) "supports the hypothesis that [semantic] redundancy is used to overcome the errors due to semantic noise" (p. 291).

To the best of my knowledge, pragmatic redundancy has not been dealt with in the literature. It is a difficult concept because it appears to be unrelated to other types of redundancy. Note first, that a necessary (but not a sufficient) condition for pragmatic redundancy is that it produces no functional response. It if produces such a response then it is necessary for that response and hence is not redundant. But now we observe that messages which are completely redundant in the syntactic sense may not be redundant in the pragmatic sense. For example, seeing or hearing a play that one "knows by heart" or hearing a memorized musical composition may affect the receiver: produce a response in him. A message, however well it is known, may still "do something" to the receiver. This is obviously the assumption if not the fact, behind repetition of commercial messages and pledges of allegiance to the flag.

The same message may produce the same or different response at different times. "Close the door" when addressed to one of my children to whom it is a highly redundant message syntactically, is nevertheless effective pragmatically since it produces a behavior that would not otherwise occur. In fact, repeating the message several times in a row often increases the probability that my son will respond as I intend and hence even the repeated messages are not pragmatically redundant. This too is a "basis" for repeated advertising messages.

I noted that failure to elicit a response--a change in the functional properties of the receiver--is only a necessary, not a sufficient, condition for pragmatic redundancy. That it is not a sufficient condition is apparent from a situation in which a person is told something that he either does not believe to be so or does not feel to be right, and he does not respond even though what he is told is completely unfamiliar to him.

Therefore, a message is completely redundant in the pragmatic sense if the response intended by the sender has already occurred and is not reproducable. It is <u>ineffective</u> if it fails to produce an intended response when the receiver has <u>not</u> so responded previously. For example, if after I have instructed my son to close the door, and he has already done so without my observing it, and I repeat the order, it is pragmatically redundant. He has already responded and cannot do so again. If however, I tell him to pick up the papers on the floor and he does so but drops some in the process, then a repetition of the message is not pragmatically redundant even if he is aware of having dropped some papers and knows what I am going to say. Even if he intended to pick up the dropped papers <u>later</u>, and my remark produces a response <u>now</u>, it produces a change in his behavior, and is not pragmatically redundant.

As much as I have said only classifies messages as completely redundant or not; it does not provide a measure of such redundancy.

> 10.5. The Amount of Pragmatic Redundancy in a message relative to a receiver is the percent of elements of the message (letters, words, sentences, or any message unit that is appropriate to the inquiry) that can be eliminated without changing the receiver's response to it.

To illustrate how this measure can be applied let me describe an exploratory experiment that several of my colleagues and I conducted to determine the effect of condensation on articles appearing in scientific journals. Since the experiment was conducted for exploratory purposes only, small samples of articles, journals, and subjects were used. This work was not intended to be reported in the literature, but only to indicate whether or not a certain line of inquiry was worth pursuing.

A number of experts in the field of operations research were asked to classify articles which had appeared in recent issues of several journals dealing with operations research. The classes used were "above average," "average," and "below average." Eight articles were selected on whose quality all of the experts agreed, four above and four below average. Letters were sent to the authors of the selected papers requesting that they prepare an "objective" examination on the content of their papers, an examination that was to be given to graduate students to whom the papers were to be assigned for reading. They were also asked to provide the answers. All did so.

Other experts who were knowledgeable in the subject matter of the papers were asked to use a red pencil and reduce each paper first to two-thirds and then to one-third of its original length. They did so only by eliminating words, sentences, or paragraphs; not by rewriting. In addition the abstracts of the articles which had appeared in the journals with them were also used. Therefore, each article was available in four versions: 100%, 67%, 33%, and abstract.

A group of graduate students who had not previously read the papers were given one version of each paper. Each version of each paper was assigned at random to an equal number of students. After reading the papers each student took the examinations prepared by the authors of the original articles.

There was no significant difference (at the 0. 5 significance level) between the average performance on the examination obtained by those who read the papers in their 100%, 67%, or 33% form. This was true for both the above- and below-average papers. These results indicate, using the measure of pragmatic redundancy constructed here, that each paper was at least 67% redundant.

Those who read only the abstracts of the above-average papers obtained a significantly lower average grade on the examination than that obtained by those who read the paper in any of its longer forms. Those who read abstracts of the below-average papers obtained an average grade that was not significantly lower than that obtained by those who had read these papers in one of their longer forms. The redundancy of the poorer papers was therefore significantly greater than than that of the better papers, but the amount of redundancy in each of them was surprisingly large. Unfortunately we did not give the examinations to students who had not read the papers in any form.

If results such as these are reproducible in a large enough and properly designed experiment, they would indicate that a considerable amount of condensation of scientific literature is possible without any significant loss of effectiveness. The amount of condensation justified by study of pragmatic redundancy would probably be much larger than that justified by study of syntactic or semantic redundancy.

The results obtained in this exploratory study help to explain the following observation by Martin and Ackoff (1963): "The fact that Digests, or Abstracts, are read twice as much [by physicists and chemists] during browsing [as compared with directed reading] might not be expected by some. It is consistent with the findings of the earlier study in which it was found that abstracts are used more as a substitute for articles than as a guide to them" (pp. 330-331).

An article that lacks any pragmatic redundancy may also lack readability. The optimal amount of redundancy, however, remains to be determined. It is likely to be dependent on other aspects of the communication situation; for example, the attitudes and beliefs of the participants.

I have already noted that a message that contains syntactically or semantically redundant parts may not be pragmatically redundant. A part of a message, or a message that is pragmatically redundant, however, must be either syntactically or sematically redundant, or both. Hence, a message (or part of one) may be redundant in all three senses.

A message that is pragmatically redundant in the absence of pragmatic noise may not be redundant when such noise is present. For example, a lecturer may repeat important points to be sure he catches some members of the audience during one of their intermittent moments of attention. Furthermore, sheer repetition can often penetrate inattention.

Redundancy is not the only way of overcoming noise, <u>feedback</u> is another.

Feedback

If A can observe B receiving his message while he is sending it, he may obtain information from the behavior of B that is usable either in formulating the as-yet-unsent part of his message, or in reformulating the message already sent. Teachers and lecturers, of course, constantly make use of such feedback in formulating their messages to their audience. *

10.6. Feedback: information received by the sender of a message about the receipt of or response to his message.

Therefore, feedback is a stimulus which produces a response in the sender of a message. More generally, feedback is information obtained by any functional entity about the product of its behavior. The product need not be a message; it may be any type of behavior. The feedback that a message-sender receives may itself be a message from the receiver of his message. This observation leads us into consideration of two-way communication.

TWO-PARTY TWO-WAY COMMUNICATION

Two-way communication between A and B involves a minimal sequence of messages:

$$[(AtoBreX) \rightarrow (BfromAreX)] \rightarrow [(BtoAreY) \rightarrow (AfromBreY)]$$

where " → " represents "produces" and X and Y may be either the same or different subjects. The sequence of messages may, of course, be extended to a larger number than two.

The conceptualization of A's communicating to B given in the first part of this chapter can also be applied to B's communicating to A, and hence the model of a two-way communication emerges out of

---

*See Chapter VIII of Smith (1966) for discussions of the effect of feedback on communication and performance of tasks.

that for one-way communication. The new ingredient is that each message after the first may be (but is not necessarily) a response to any of the preceding messages, the sender's or the receiver's.

In Chapter 8 it was shown that a message may inform, instruct, and/or motivate its receiver, whatever the intention of the sender. The sender, of course, may intend to inform, instruct, or motivate either the receiver or himself.

10.7. Question. Any message which is sent by A with the intention of producing a responsive message that will inform, instruct, or motivate A, whatever its structure (syntax).

When A sends a question to B, he asks him something. On the other hand, if the intent of A's message to B is to inform, instruct, or motivate B, A tells B something.

10.8. Statement: any message which is sent with the intention of informing, instructing, or motivating the receiver.

A question and a request are related but are not identical.

10.9. Request. If A sends a message to B which A intends to produce a choice of any type of course of action (including, but not necessarily, communication) by B which A desires, then A makes a request of B.

Every question is a request for further communication, but not every request is a question; for example, "Please, close the door."

Some other important types of messages which are related to those just considered require the concepts of reward and punishment.

10.10. Reward. An individual is rewarded for doing (or not doing) something if his action (or lack of action) produces behavior in another (or himself) which increases his probability of obtaining something that he desires.

10.11. Punishment. An individual is punished for doing (or not doing) something if his action (or lack of action) produces behavior in another (or himself) which decreases his probability of obtaining something that he desires.

10.12. Threat: a message which signifies both an intention by the sender that the receiver does (or does not do) something and an intention by the sender to punish the receiver if he does not do (or does) that something.

10.13. Promise: a message which signifies both the intention of the sender to do something of value to the receiver, and the intention of the sender to receive punishment if the intended act is not carried out.

10.14. Order: a request which carries with it a threat of punishment to the receiver if he does not respond as the sender intends he should.

Requests and orders do not require two-way communication; but questions do. Questions require answers.

Minimal two-way communication may consist of either

(a) tell - tell
(b) ask - tell
(c) ask - ask
(d) tell - ask

A communication that terminates with a "tell" may or may not be complete; one that ends with a question is necessarily incomplete: it leaves a request unfilled.

An important class of communications between an A and a B about an X consists of those that can be viewed as attempts to produce agreement or disagreement between A's and B's beliefs and/or attitudes toward X. Newcomb (1966) has examined this process and formulated several postulates about such communication. I would like to analyze one of these postulates and by so doing show how the conceptual system provided here can enrich Newcomb's assertions, make them more precise, and provide the basis for designing effective tests of their validity.

Newcomb's Hypothesis on Two-Way Communication

If, in this examination, I do injustice to Newcomb's intentions, it is not intentional. I try to get at what _he_ means but if I fail to do so it is not because the type of operational translation into an objective teleology that I attempt is of no value, but because I do not understand him. To some, what I am about to do may appear like nit-picking. However, it is intended to support, by example, several fundamental criticisms of much of contemporary behavioral science: (1) that the psychology and social psychology of communication is rife with imprecise definitions and inconsistent use of concepts, (2) that a systematic way of assigning numbers to a phenomenon is not sufficient to produce measurements, and (3) that the use of quantitative relationships in assertions about communication does not necessarily produce a quantitative theory of communication.

Newcomb's first postulate is as follows:

> The stronger the forces toward A's co-orientation in respect to B and X, (a) the greater A's strain toward symmetry with B in respect to X; and (b) the greater the likelihood of increased symmetry as a consequence of one or more communicative acts (p. 69).

He defined the key terms in this postulate as follows:

> "Co-orientation" ... represents an assumption; namely, that A's orientation toward B and toward X are interdependent (pp. 66-67).
>
> A's orientation toward X, including both attitude toward X as an object to be approached or avoided (characterized by sign and intensity) and cognitive attributes (beliefs and cognitive structuring).
>
> A's orientation toward B, in exactly the same sense.(For purposes of avoiding confusing terms, we shall speak of positive and negative attraction toward A and B as persons, and as favorable and unfavorable attitudes toward B.) We shall refer to lateral similarities of A's and B's orientation to X as symmetrical relationships (p. 67).

This last definition is illuminated by the discussion preceding it:

> In order to examine the possible relationships of similarity and difference between A and B, we shall make use of simple dichotomies in regard to these four relationships [A's orientation toward X and A, and B's orientation toward X and A]. That is, with respect to a given X at a given time, A and B will be regarded as cathectically [i. e., with respect to feeling] alike (++ or --) or different (+ - or - +) in attitude and in attraction; and as cognitively alike or different. We shall also make use of simple dichotomies of degree--i. e., more alike, less alike (p. 67).

First consider Newcomb's condition: "the stronger the forces toward A's co-orientation in respect to B and X." A's co-orientation according to Newcomb is characterized by four variables:

(1) A's attitude toward X

(2) A's cognitive attributes (beliefs and cognitive structuring) of X

(3) A's attraction toward B

(4) A's cognitive attributes of B

Although I can see how A's attraction toward B and attitude toward X can each be represented on a single scale and hence treated dichotomously (alike or different), it is not clear to me how to so represent "beliefs and cognitive structuring." The number of relevant beliefs that A can have about either X or B may be very large. Under what conditions are sets of measures of beliefs to be taken to be alike or different?

What of "the stronger the forces toward..."? I would translate this to refer to the strength of the interdependence of the variables listed above. Let us assume we can find one measure to represent beliefs, let alone beliefs and cognitive structuring (I do not understand the latter term and hence conveniently ignore it henceforth).

First, what interdependencies are to be measured? Between A's and B's attitudes, and between A's and B's beliefs; or between A's beliefs and attitudes, and B's beliefs and attitudes? If the former, then there will be two measures of interdependency. How are these to be aggregated? If the latter, it is even more difficult to see how interdependency is to be represented because four relationships are involved: (1) A's attitude and B's belief, (2) A's attitude and B's attitude, (3) A's belief and B's attitude, and (4) A's belief and B's belief. This assumes, of course, that only one belief is involved.

Further, what does "interdependency" mean? Is a correlation implied? Positive, or negative, or both? Or is interdependency the probability that a change in one of the related measures will produce a change in the other? Of the same magnitude? In the same direction?

Unless "interdependency" is defined operationally in measurable terms, and unless the variables involved are identified and similarly defined, the postulate itself has no operational significance.

Continuing with the first consequence of the premise we have been examining--"the greater A's strain toward symmetry in respect to X"--we must clarify "strain" and "symmetry." It seems to me that by "strain" Newcomb intended to connote something very much like what I have called "intention." A measure of symmetry involves the same difficulties discussed above with respect to interdependencies. Newcomb refers to "lateral similarities;" therefore, several comparisons are involved. If each produces a judgment of "like" or "different" how are these to be aggregated? Are we to take the ratio of "likes" to the total number of comparisons? If we do, we would be assuming that each comparison is equally important. Is this what Newcomb meant to imply?

The second conclusion--"the greater the likelihood of increased symmetry as a consequence of one or more communicative acts"--appears to be translatable into "the greater the probability that a specified number of communicative acts will produce an increase in symmetry." But this translation and the original both require a definition of a "communicative act." Is the voicing of one word one act? Or is it the production of one continuous uninterrupted message? Is it independent of the length of the message or its duration, and so on?

Now let me try to use what I have done here to formulate a less general hypothesis than Newcomb's, but one of the same type, and to make it less ambiguous than his. First, I shall restrict attention to attitudes and again use "AaB" to represent A's attitude toward B. Following the discussion in Chapter 7, by A's attitude toward B, I mean A's intention to retain B in his environment (hence satisfaction with B's presence). The degree of this intention can range between 0 and 1. If this measure is greater than 0.5, A can be said to have a favorable attitude toward B; if it is less than 0.5, his attitude is unfavorable; and if equal to 0.5, A is indifferent to B.

Now I want to make precise the following statement: A's attitude toward X depends on both his attitude toward B and B's attitude toward X; that is, AaX depends on AaB and BaX.

With Newcomb let us treat attitudes dichotomously and let $(AaB)^+$ represent a favorable attitude, and $(AaB)^-$ an unfavorable one. Then we can say that AaX depends on AaB and BaX if the probability that AaX is favorable (or unfavorable) is greater if AaB and BaX are favorable. Now an interesting point arises: the probability that AaX is favorable (or unfavorable) may be greater when both AaB and BaX are unfavorable (favorable) than if only one is. That is, if A's attitude toward B is unfavorable and B's attitude toward X is unfavorable, A's attitude toward X may very likely be favorable. One may like something because his enemy doesn't.

Now let us define "strain toward symmetry" as A's intention to minimize the difference between his attitude toward X and B's; that is, to minimize (AaX - BaX). If this intention is greater than 0.5, A strains toward symmetry; if it is less than 0.5, A strains toward asymmetry. Let P_A [min(AaX - BaX)] represent the probability that A strains toward symmetry, and P_A [max(AaX - BaX)] represent the probability that A strains toward asymmetry.

We can now formulate the following hypotheses:

(1) As

$$[P(AaX)^+ \mid (AaB)^+ \& (BaX)^+] - \left[P(AaX)^+ \,\middle|\, \begin{matrix} (AaB)^+ \& (BaX)^- \\ \text{or} \\ (AaB)^- \& (BaX)^+ \end{matrix} \right]$$

increases

P_A [min (AaX - BaX)]

also increases.

(2) As

$$[P(AaX)^+ \mid (AaB)^- \& (BaX)^-] - \left[P(AaX)^+ \,\middle|\, \begin{matrix} (AaB)^+ \& (BaX)^- \\ \text{or} \\ (AaB)^- \& (BaX)^+ \end{matrix} \right]$$

increases,

$P_A[\max(AaX - BaX)]$

also increases.

Complementary hypotheses can be obtained by changing all the plus superscripts to minuses, and the minus superscripts to pluses.

The advantage of a symbolic statement of this hypothesis over a statement of it in words becomes apparent when I try to express the first one in words:

> As the difference between (1) the probability that A's attitude toward X is favorable (given that his attitude toward B and B's toward X are favorable), and (2) the probability that his attitude toward X is favorable (given that either his attitude toward B or B's toward X is unfavorable), increases; then A's intention to minimize the difference between his and B's attitude toward X also increases.

The second hypothesis covers a possibility not considered by Newcomb: if A's attitude toward B is unfavorable and his attitude toward X depends on his attitude toward B, and B's attitude toward X is unfavorable, A may strain for asymmetry with B with respect to X.

Now consider Newcomb's second conclusion: "the greater the likelihood of increased symmetry as a consequence of one or more communicative acts." Let us define a communicative act as the sending and receipt of a message containing a specified amount of syntactic information in the absence of pragmatic noise. Then we can formulate the following hypothesis:

(3) As

$$[P(AaX)^+ \mid (AaB)^+ \& (BaX)^+] - \left[P(AaX)^+ \,\middle|\, \begin{matrix}(AaB)^+ \& (BaX)^- \\ \text{or} \\ (AaB)^- \& (BaX)^+\end{matrix}\right]$$

increases, then the greater is the probability that a communicative act between A and B will reduce $[(AaX) - (BaX) \mid (AaX) \neq (BaX)]$; and as

$$[P(AaX)^{+} \mid (AaB)^{-} \; \& \; (BaX)^{-}] - \left[P(AaX)^{+} \;\middle|\; \begin{matrix} (AaB)^{+} \; \& \; (BaX)^{-} \\ \text{or} \\ (AaB)^{-} \; \& \; (BaX)^{+} \end{matrix}\right]$$

increases, then the greater is the probability that a communicative act between A and B will increase $[(AaX) - (BaX)]$, given that this difference is not maximum.

Similar hypotheses can be formulated about beliefs but, as I have indicated, there is no summary belief as there is a summary attitude (e.g., favorable or unfavorable). Hence the content of the beliefs taken to be relevant must be specified. For example, whether or not an object is believed to be hard may be relevant in some situations but not in others.

I hope I have shown (1) how loosely formulated hypotheses can be tightened up, and (2) how a conceptual system assists in doing so. In this latter connection it should be recalled that measures of belief and attitude, so central to this discussion, were developed in earlier chapters. Without these measures the hypotheses formulated here would be empty, no matter how precise their formulation.

Rapoport's Hypotheses

Consider the following relatively simple hypothesis:
If $(AaB)^{+}$, $(BaA)^{+}$, and $(AaX) \neq (BaX)$, then two-way communication between A and B about X will produce a decrease in $[(AaX) - (BaX)]$.

That is, if A and B have favorable attitudes toward each other but their attitudes toward X differ, communication between them will decrease this difference. This hypothesis suggests the question:

If $(AaB)^-$, $(BaA)^-$, and $[(AaX) - (BaX)] \neq \max$, then will communication between A and B reduce the difference $[(AaX) - (BaX)]$?

In his discussion of the effectiveness of debates in resolving conflicts, Rapoport (1960) suggests several ways of increasing this effectiveness. These suggestions can be translated into hypotheses formulated within the conceptual system developed here. For example, Rapoport suggests that if two hostile persons must debate on a subject on which they disagree, they are more likely to reach agreement if each is required to formulate the other's point of view in a way that the other accepts. Thic can be translated as follows:

If

(1) $(AaB)^-$, $(BaA)^-$, and $[(AaX) \neq (BaX)]$,

(2) A sends a message to B connoting what A believes to be B's attitude toward and beliefs about X, and B accepts these connotations, and

(3) B sends a corresponding message to A which A similarly accepts,

then the probability that subsequent communication between A and B will reduce the difference, $(AaX) - (BaX)$, increases, as compared with what would happen if either condition (1) or (2) were not satisfied.

Now let us consider how this hypothesis could be tested. First, we must be able to measure four attitudes: (AaB), (AaX) (BaA), and (BaX). We have already considered how this can be done in Chapter 7. Next we require a sample of pairs of people who satisfy condition (1) above with respect to an X. (At the time of this writing, for example, if X were "U.S. policy in Viet Nam," they would be easy to find.) The attitude of each person toward the other and X would also be determined.

We would then randomly divide these hostile pairs into two groups of equal size. Pairs in one group would be told to try to reach agreement on X within a specified time. Pairs in the other group would be told to do the same thing only after they had satisfied conditions (2) and (3) above. At the end of the designated time, the attitudes toward X of each member of each pair would again be measured and the differences obtained. A comparison of the "before" and "after" differences would confirm or disconfirm the hypothesis.

Rapoport's second hypothesis involves the effect of each party "deliniating the region of validity of the opponent's stand." He explains as follows:

> It is not unusual in debate to point out grounds for considering the position of the opponent invalid. It is argued, for example, that some or all of the premises assumed by the opponent do not hold. In the approach where the removal of threat is a major consideration, this procedure must be reversed. The logical implications remain formally the same: by deliniating the conditions under which the opponent's point of view is valid, we imply the residual conditions, under which it is not valid. But the emphasis is on the former, not on the latter. Showing examples which support the opponents' point of view is a continuation of our message to him that he has been heard and understood (p. 287).

This hypothesis involves a message or messages from each party of the conflict to the other which states the conditions under which he believes (1) the other's beliefs about X to be valid and (2) his attitudes toward X to be justified. It asserts that if there is such an interchange that differences between attitudes toward X will be reduced by subsequent communication. These assertions can also be translated into the conceptual system being developed here.

> Once A and B have each produced a statement of the other's beliefs and attitudes toward X, which the other has accepted if A sends a message to B which connotes the conditions under

which A's beliefs and attitudes toward X would be the same as B's are under current conditions, and B does the same; then the probability that subsequent communication between A and B will reduce the difference, (AaX) - (BaX), increases.

Communication of almost any form between conflicting parties does seem to reduce the tendency to conflict. In several laboratory experiments on conflict and cooperation in which the interaction takes place under conditions that remain the same except for the presence or absence of communication; a significantly greater tendency to co-operative behavior has been found where communication is possible. (See, for example, Ackoff et al, 1966.)

Up to this point I have only considered communication between two parties. I turn now to communication between more than two-parties.

MORE THAN TWO-PARTY COMMUNICATION

Westley and MacLean (1966) have produced a very provocative conceptual model for research on communications which involve more than two parties.

Their concern is with mass communications but, I believe, their concepts can be fused with mine, to produce a more general model of what might be called (following Bavelas, 1966) chain communication; that is, situations in which A communicates to B through C, I shall refer to C as an intermediary in this context. The model is extendable to any number of intermediaries and hence to a chain of any length. Furthermore, by reversal of roles (say between A and B) types of communication networks other than the chain result. Now I let Westley and MacLean speak for themselves:

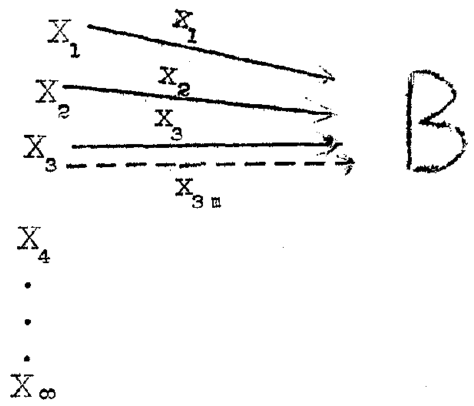


FIGURE 1.

FIGURE 1. Objects of orientation ($X_1 \ldots X_\infty$) in the sensory field of the receiver (B) are transmitted directly to him in abstracted form ($X_1 \ldots X_3$) after a process of selection from among all Xs such selection being based at least in part on the needs and problems of B. Some or all are transmitted in more than one sense (X_{3m}, for example).

From the standpoint of B, the world consists of a confusion of Xs. And these Xs may include As. B has within his field an infinity of potential Xs. He has learned that in order to maximize satisfactions and solve security problems he must orient toward Xs selectively. But the mature B... does not orient toward X alone, but tends, in the presence of an A, to orient simultaneously toward both A and X...

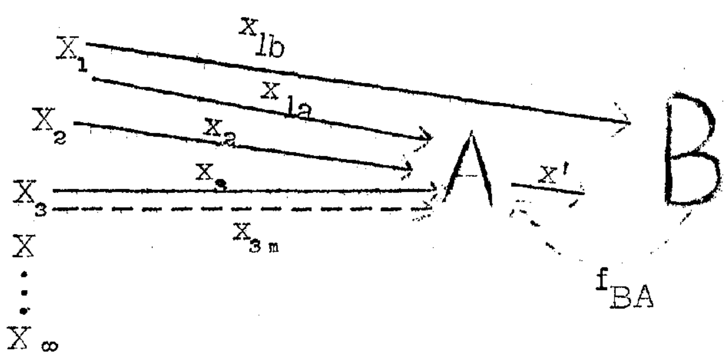


FIGURE 2.

FIGURE 2. The same Xs are selected and abstracted by communicator (A) and transmitted as a message (X') to B, who may or may not have part or all of the Xs in his own sensory field (X_{1b}). Either purposively or non-purposively B transmits feedback (f_{BA}) to A.

With respect to the As and Xs in his own immediate sensory field, B is capable of receiving and acting upon information thus transmitted to him and must do so if he is to maintain an adequate orientation to his immediate environment.

But what of As and Xs relevant to such orientation but lying outside his immediate reach? If these are to impinge on him, there is need for another role, which we will call C.

C is conceived of as one who can (a) select the abstractions of object X appropriate to B's need satisfactions or problem solutions, (b) transform them into some form of symbol containing meanings shared with B, and finally (c) transmit such symbols by means of some channel or medium to B. . . .

It may be asked why C would choose Xs "appropriate" to the requirements of B. The answer would appear to be that the C role can survive only to the extent that this is true. For B is still a selector among the offerings of various Cs and this means that Cs are in effect competitors for the attention of Bs (and for that matter competitors with As and Xs in B's immediate field). Cs therefore survive as Cs to the extent that they satisfy needs for Bs. And Bs, on the basis of the most obvious propositions of learning theory, will tend to return to those Cs which have provided past need satisfactions and problem solutions.

C, then, is capable of serving as an agent for B in selecting and transmitting information about an X (or an A - X relationship). He does so by means of symbols expressing shared meanings about Xs through channels that provide connection between X and B. And he does so in circumstances where such a connection is otherwise impossible for B. Thus B has a basis for increasing his security in the larger environment and for gaining increased need satisfactions. In other words, the effect of the additon of the C role is to provide B with a more extended environment.

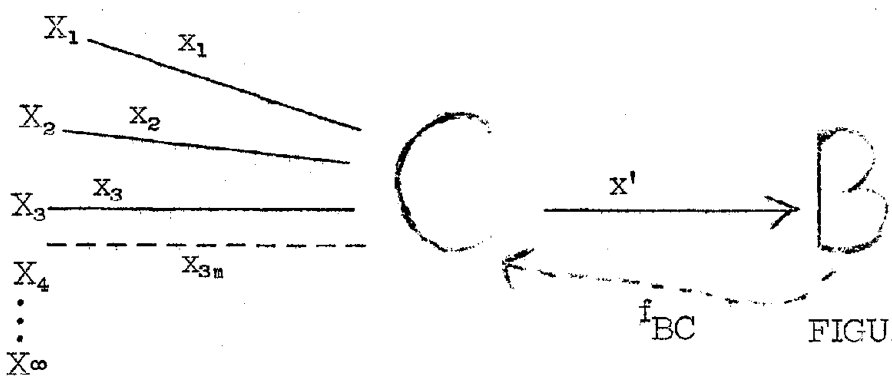


FIGURE 3

FIGURE 3. What Xs B receives may be owing to selected abstractions transmitted by a non-purposive encoder (C), acting for B and thus extending B's environment. C's selections are necessarily based in part on feedback (f_{BC}) from B.

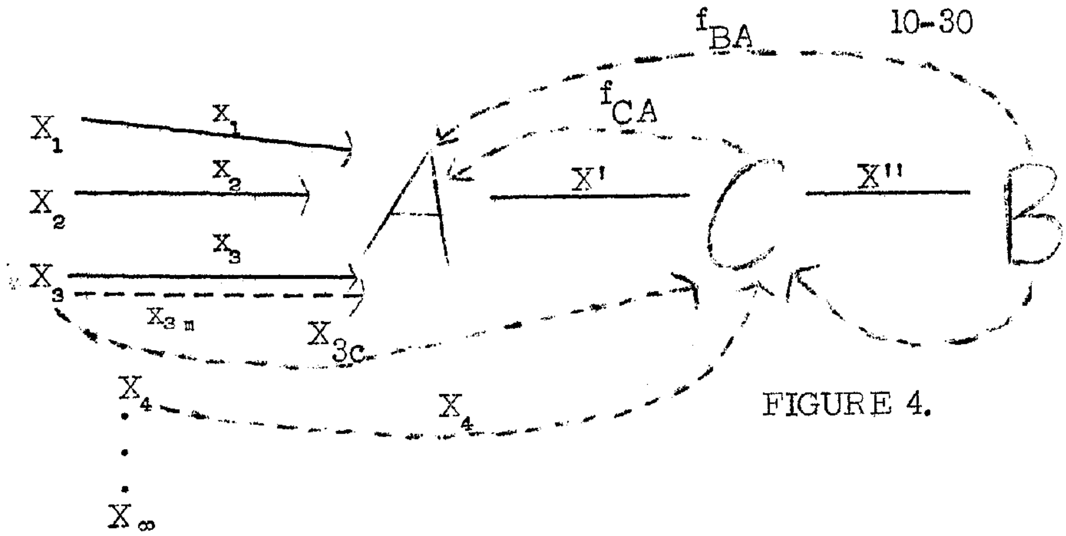


FIGURE 4.

FIGURE 4. The message C transmits to B (X'') represent his selections from both messages to him from A's (X') and C's selections and abstractions from Xs in his own sensory field (X_{3c}, X_4), which may or may not be Xs in A's field. Feedback not only moves from B to A (f_{BA}) and from B to C (f_{BC}) but also from C to A (f_{BA}). Clearly, in the mass communication situation, a large number of Cs receive from a very large number of As and transmit to a vastly larger number of Bs, who simultaneously receive from other Cs.

For Newcomb, A and B can only be persons. While we have tended to imply persons in these roles, it should now be made clear that we do not intend to confine the model to the level of the individual personality. The role of B, for instance, may be that of a person, or a primary group, or a total social system. At the social system level, a national state requires and maintains an elaborate network of Cs performing such special information functions as that of the diplomatic service...

"PURPOSIVE" OR "NON-PURPOSIVE"?

A purposive [what I have called "intended"] message is one A originates for the purpose of modifying B's perception of an X. A non-purposive [unintended] message is one which is transmitted to B directly or by means of a C and in the absence of any communicator's intent to influence him. The absence of a communicator's intent to influence B transforms his act into an X. When a person says something he hopes will reach another person's ears, he is an A; but if he says it without such intent and it nevertheless is transmitted to B, his act must be conceived of as an X, the selection and transmission having been performed by a C...

Messages are transmitted in codes (symbol systems). But this model is by no means limited to the most obvious ones--linguistic systems. In fact...the crucial characteristic is the shared meanings associated with symbols. Such symbols can take virtually any form, so long as and to the extent that there exist shared meanings and that they are transmissible. Such shared meanings surrounding symbols can be either affective or cognitive...

Our Bs vary in the degree to which they share common problems. Common problems imply the necessity of attaining communication with common Xs. Media serving to bring such Xs to such Bs arise out of the perceptions by Cs of the existence of just such a need. Special symbol systems are developed to maximize transmission...

FEEDBACK

Another concept crucial to the model is that of "feedback." In the first place it should be clear from the foregoing that it is feedback that assures the system character of the ABX (or ABCX) relationship. If A is to utilize his experience in influencing B, he must have information about any changes in the condition of B attributable to his communications. C is equally concerned with effects on B if he is to make realistic adjustments in his role as B's "agent." Such As as advertisers facilitate feedback by means of elaborate market research; public relations men obtain feedback by means of public-opinion polls and other devices for determining the effects of their messages. Such Cs as newspaper publishers sponsor readership surveys and, more recently, reader motivation studies to estimate and predict reader response. Radio's concern with "fan mail" and popularity ratings is well known.

Although feedback originates with B under most circumstances, it need not be assumed that B is necessarily trying to communicate back to C or A. When he does try to do so, we may think of this as purposive feedback. This is the case when an angry reader writes a letter "straightening out" the editor on some favorite issue. But there are also many ways B can feed back without intending to. These we will call non-purposive feedback. When a television fan decides to try a well-advertised detergent, his purchase becomes part of the data of a market survey, even though he may not have intended to let the sponsor know he had won a convert... (pp. 81-87).

Hardly any translation is required of Whestley's and MacLean's terms. As I have indicated by an interjection in the quotation from their work, they use 'purposive' as I would use 'intended', and 'non-purposive' as I would use 'purposive but not intended'.

Whestley and MacLean restrict the concept of an intermediary (C) to something that acts without purpose (in their sense) and without intention to affect the receiver (in my sense): "Cs serve as agents of Bs in selecting and transmitting non-purposively the information Bs require, especially when the information is beyond the immediate reach of B" (p. 87). It is not clear to me why they so restrict the function of the intermediary. It seems to me that the intermediary may alter the intended content of the sender's message (e.g., by censorship, editorializing, and so on) so as to change its effect on the receiver. In such cases the intermediary's behavior would be intentional in my sense and purposive in theirs. The intermediary obviously may act as a filter and as a condenser of messages as well as a distorter, collector, or transmitter of messages.

It seems possible to me to formulate a more general conception of social communication than has been developed by Whestley and MacLean. Let me begin with the obvious.

Messages from different sources (As) about the same X, even if intended for the same receiver (B), may be structurally or functionally dissimilar either (1) because of the differences in what two or more As observe even when they observe the same X, or (2) because of the difference in their relevant beliefs and attitudes involving X, Bs, Cs, and any other individuals in the system, or (3) because of differences in their abilities to formulate effective messages. Such differences create the need for evaluating alternative sources of information, instruction, and motivation.

When As intentionally send messages about Xs to Cs, they may intend the Cs to be receivers, not intermediaries. The neighborhood gossip may retransmit a message that the sender had not intended to go any further. C's may intercept messages not intended or intended not to reach them (e.g., a newspaper reporter overhears a conversation and reports its content). Indeed, Cs may conceive of their role as largely that of obtaining messages from As which the As do not intend to make available to Bs, or, for that matter to Cs. This appears to be the case where "private investigators," "secret agents," or "exposé journalists" are involved.

I should like to consider in detail the communication functions which intermediaries (and senders and receivers as well) can perform. First consider the production of a message.

10.15. Encoding: the act of producing a message.

Note that this is encoding in the pragmatic sense. It implies encoding in the syntactical sense, but such encoding does not imply pragmatic encoding. Syntactic encoding can produce a set of signs which are not capable of communicating. In pragmatic encoding a set of signs are produced which signify something the producer has experienced: perceived, thought, intuited, or felt.

10.16. Decoding: the production of a response by a message to that which it signifies.

Decoding in this pragmatic sense similarly implies syntactical decoding, but the converse is not necessarily true.

Although 'encoding' is often used synonymously with 'translation', I prefer to use them differently:

10.17. Translation: the act of changing the signs in a message from one language into another.

Thus translation presupposes encoding. The sender, receiver, or intermediary may translate a message.

A message is encoded by the sender and decoded by the receiver. It is <u>transmitted</u> from the former to the latter.

10.18. <u>Transmission</u>: the behavior by means of which a message produces a response in the receiver.

10.19. <u>Channel</u>: the instruments (objects, events, and their properties) which produce transmission.

For example, in sending a letter the postal service is the channel, in telephonic communication the telephone system is, and in speech the atmosphere is.

Now, intermediaries, as well as senders and receivers, can affect messages in a number of ways. It is convenient to consider these in connection with possible roles of intermediaries.

Intermediaries may be passive; that is, receive whatever is sent to them and transmit it without intentional modification. As I have already indicated, however, they may actively intervene in the communication between senders and receivers. Such intervention may serve the purposes of the senders and receivers either well or poorly. For example, rewriting news reports may be useful to the pressured sender and ultimate receiver; but censorship may serve neither's purpose well.

First consider <u>passive</u> intermediaries, ones whose only effect on messages is structural, not functional. In the most extreme case they serve only as a channel: they receive and deliver the message to the receiver. The post office, messengers, and the telephone system act in this way. The intermediary may also transform the signs of a message into structurally different but functionally equivalent signs, as when a secretary takes dictation and types a letter. Frequently such transformations are made to facilitate <u>storage</u> of the message until it can be received or is wanted by the receiver. The intermediary may store and subsequently retrieve the message from storage; for example, libraries and file systems.

Passive intermediaries, then, are ones which transmit, structurally transform, store, and retrieve messages. They can be classified into one of four types.

1. The <u>one to one</u> intermediary who receives messages from only one source (A) and transmits them to only one receiver (B). Such an intermediary can be one- or two-way. If two-way, it also receives messages from B which it transmits to A. There seem to be few intermediaries of this sort, except in contrived situations such as are constructed in laboratory experiments involving communication: I have known cases in which two persons who are not on "speaking terms" will communicate to each other through a third party. "Feelers" between two nations at war are frequently handled through a third party, often many third parties.

2. The <u>many-one</u> intermediary who receives messages from many sources and transmits them to only one receiver. In these cases the intermediary's function can be likened to that of a funnel. A secretary frequently performs this function. Some intelligence officers in the military do so for the senior officer to whom they report. Directors of marketing research often serve in this function for marketing executives.

3. The <u>one-many</u> intermediary, who receives messages from one source but transmits them to many receivers. Such an

intermediary disperses messages. A public relations or press officer may serve this function. A book publisher does so for its author.

4. The many-many intermediary who, of course, receives messages from many sources and transmits them to many sources. Newspapers, journals, and libraries are examples of such intermediaries.

Active intermediaries do more than affect the structure or transmission of messages. As already noted, they may translate messages, transform them from one language into another. In addition there are a number of other functions which they may perform among the most important of which are filtration, condensation, and editing.

10.20. Filtration: the selection of a subset from the set of messages intended for a receiver, for transmission to him.

An intermediary may filter messages with the intention of better serving the receiver's purposes; for example, transmitting only messages that he believes are of value to the receiver. Or the intermediary may filter for its own or another party's purposes. When it does so it engages in censorship.

10.21. Censorship: filtration that is intended to serve the purposes of a party other than the sender or receiver of a message.

The refereeing process used by most professional journals is intended to serve the receivers' purposes and hence is not censorship, but it is filtration. Filtration always involves evaluation of messages for their effectiveness. It attempts either to eliminate undesirable responses from the receiver's or someone else's point of view, or to eliminate messages that will produce no response (e.g., ones which are completely redundant).

10.22. Condensation: the reduction of the number of signs in a message or the transformation of them into a set of signs whose receipt requires less time than did the original message.

The intermediary may either reduce the message while trying to retain its essential content in order to reduce the receiving time required (i.e., digest the message), or provide a brief description of its content to that the potential receivers can decide whether or not they want to receive the full message (i.e., abstract the message). A digest is intended to replace the message. An abstract is intended to provide a basis for deciding whether or not to receive the message; thus it serves as an instrument for filtration.

10.23. Editing: the act of changing a message with the intention of increasing its effectiveness for the sender and/or the receiver.

Not only do editors perform this function but (at least good) secretaries do as well. The sender himself may perform the editorial function.

When there is an intermediary between A and B and A intends to communicate to a particular B or class of Bs, his beliefs about and attitudes toward C may also affect his formulation of his message. B's corresponding beliefs and attitudes involving C may also affect what message he receives and how he responds to it. This is particularly the case when two different Cs transmit inconsistent messages on the same subject (e.g., contrary accounts in different newspapers of the same event). Which of conflicting messages on the same subject the receiver believes is largely influenced by his beliefs about and attitudes toward the subject of communication, the senders and the intermediaries. As Whestley and MacLean have pointed out, a receiver may select that intermediary whom he believes will most efficiently serve his purposes.

The Arbitrator as Intermediary

Arbitrators in conflicts between two parties (e.g., labor and management or two nations in a dispute) serve as intermediaries operating in both directions. They may meet with each party separately until they have established a basis for direct (non-intermediated) communication between the two parties, or they may meet with both parties together and attempt to direct the communication between them. It is apparent that the attitudes of the conflicting parties toward the intermediator has a considerable effect on his effectiveness in producing agreement. In such a case A's attitude toward the arbitrator is likely to be strongly influenced by

(1) what A believes C's attitude toward A is: Ab(CaA)

and

(2) what A believes C's attitude toward B is: Ab(CaB).

The same is true for B. A's attitude toward C is likely to be favorable if A believes C's attitude toward A is more favorable (or no less favorable) than is C's attitude toward B.

One function of the arbitrator is to define the issue: the differences between A and B. Hence, he may go through a process much like that advocated by Rapoport for the parties of a debate. He may formulate A's and B's beliefs and attitudes in a way that is acceptable to them and he may try to find the conditions under which each believes the other's position is valid. Therefore, he can serve as a facilitator of the type of debating process that Rapoport advocates.

The arbitrator seeks a way of resolving or dissolving the conflict once the "problem" has been defined. He may not find any way of doing so. In such cases he may try to find an "equitable solution," one which removes exploitation and/or reduces the intensity of conflict. (See Chapter 11 for discussion of these concepts).

It should be apparent from this brief discussion that many hypotheses concerning the effectiveness of arbitration can be formulated within the conceptual system constructed here; for example, hypotheses that relate the arbitrator's effectiveness to the attitudes of the conflicting parties toward him, and his attitudes toward them.

CONCLUSION

The chain is obviously only one type of network by which multiple parties can be connected. A detailed analysis of different types of communication networks and ways of characterizing them can be found in Bavelas (1966). Experiments dealing with the effects of such networks on communication and task performance can be found in Leavitt (1966), Guetzkow and Simon (1966), Shaw et al (1966), Mulder (1966), and Macy, Jr., et al (1966). Note that a network is a property of a group, not of the individuals that compose it taken separately.

The most complex network is one in which every party can communicate directly with every other party. For a group of three, four, or five individuals such networks can be represented as is shown in Figure 10.1

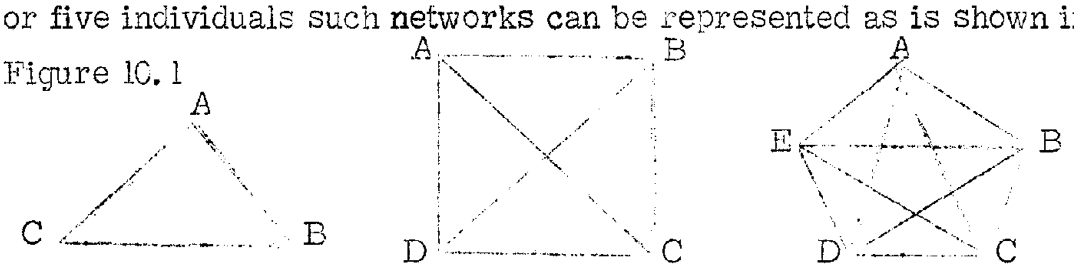


FIGURE 10.1. Three-, four-, and five'node networks.

Consider the simplest of these, the three-party network. The attitudes of each party toward the other two are now relevant to the behavior of each. Even if these attitudes are treated dichotomously (e.g., favorable and unfavorable), there are $2^6 = 64$ possible permutations of attitudes. In general, if there are n persons there are $2^{(n^2-n)}$ permutations of attitudes. Thus even for five persons there are more than a million permutations of just their attitudes. If we ignore the individuality of the participants and treat them as equivalent points in the network, then we can deal only with the combinations (not permutations) of attitudes. For n=3, there are seven such combinations:

Favorable	Unfavorable
6	0
5	1
4	2
3	3
2	4
1	5
0	6

In general, there are $n(n-1)+1$ such combinations. Hence, for five persons there are twenty-one combinations of dichotomously treated attitudes. But in order to get down to this number we must give up considering each party uniquely, and therefore much of the psychological content of the situation is sacrificed. The reason for making this sacrifice is apparent: a model of an interaction of individuals which treats each person uniquely would be too complex to handle. Remember that we have only considered their attitudes towards each other, and these only dichotomously. We have not considered their attitudes towards X nor any of their beliefs.

It becomes clear why groups are seldom conceptualized as the sum and interactions of their parts. For practical reasons it is necessary either to depersonalize the members of a group or to treat the group itself as an individual, hence the emergence of sociology.

Similar difficulties make physics arise from mechanics. Even if the behavior of bodies can be explained in principle when complete knowledge of each of their point-particles and their inter-relationships is available, it is not feasible to deal with bodies in this atomistic way. Bodies are themselves treated as individuals.

In the concluding chapter, I consider the conceptual transition from the purposeful individual to the purposeful group as an entity and indicate how the group can be treated as a teleological system, and how this conceptualization can be made completely compatible with that of the individual which has been developed here.

In the next chapter I consider the concepts or conflict which have been introduced here.

BIBLIOGRAPHY

Ackoff, R. L., et al, Model Study of Escalation: Report of 1964-1965 Studies. Management Science Center, University of Pennsylvania, Philadelphia, March 1, 1966.

Bateson, Gregory, "Information, Codification, and Metacommunication," in Smith (1966), pp. 412-426.

Bavelas, Alex, "A Mathematical Model for Group Structure," in Smith (1966), pp. 216-222.

Guetzkow, Harold, and Simon, H. A., "The Impact of Certain Communication Nets Upon Organization and Performance in Task-Orientation Groups," in Smith (1966), pp. 244-253.

Heise, G. A., and Miller, G. A., "Problem Solving by Small Groups Using Various Communication Nets," in Smith (1966), pp. 283-285.

Leavitt, H. J., "Some Effects of Certain Communication Patterns on Group Performance," in Smith (1966), pp. 222-243.

Macy, J., Jr., Christie, L. S., and Luce, R. D., "Coding Noise in a Task-Oriented Group," in Smith (1966), pp. 285-294.

Martin, M. W., and Ackoff, R. L., "The Dissemination and Use of Recorded Scientific Information," Management Science, 9, (January, 1963), pp. 322-336.

Mulder, Mank, "The Power Variable in Communication Experiments," in Smith (1966), pp. 259-274.

Newcomb, T. M., "An Approach to the Study of Communicative Acts," in Smith (1966), pp. 66-79.

Rapoport, Anatol, Fights, Games, and Debates. The University of Michigan Press, Ann Arbor, 1960.

Shaw, M. E., Rothschild, G. H., and Strickland, J. F., "Decision Processes in Communication Nets," in Smith (1966), pp. 253-259.

Smith, A. G., Communication and Culture. Holt, Rinehart and Winston, New York, 1966.

Weaver, Warren, "The Mathematics of Communication," in Smith (1966), pp. 15-24.

Whestley, B. H., and MacLean, Jr., M. S., "A Conceptual Model for Communications Research," in Smith (1966), pp. 80-87.

Chapter 11

CONFLICT, COOPERATION, AND COMMUNICATION

DISCUSSION, n. A method of confirming others in their errors (Ambrose Bierce, The Devil's Dictionary).

INTRODUCTION

In the discussion of communication in the last three chapters we have been concerned with one way in which one person can affect another. In this chapter I examine in more detail the nature of the effects that one person can have on another; that is, the nature of the interactions between purposeful individuals.

Two related concepts are central to this discussion: conflict and cooperation. From these concepts others will be derived, of which the most important are exploitation and competition. I will also consider the ways in which interactions between two individuals can be affected by what they and others do. One of the more important ways of affecting interactions involves communication.

Finally, I shall consider ways of conceptualizing or modeling interactions and emphasize the difference between an observer's view of such interactions and that of the participants.

The concepts of a choice situation and its components play a central role in this discussion, so let me review them briefly. I continue to use A and B to represent subjects, but will introduce T to represent "third parties." S continues to represent the environment of subjects. C_i $(1 \le i \le m)$ represents the courses of action available in the environment and O_j $(1 \le j \le n)$ the possible outcomes. Both courses of action and outcomes are considered to be so defined as to be exclusive and exhaustive unless otherwise noted. The parameters of the choice situation are P_i, the probability that C_i will be selected in $S(\sum_i P_i = 1.0)$; E_{ij} the probability that C_i will produce O_j in S $(\sum_j E_{ij} = 1.0)$; and V_j the relative values of O_j to the subject in S. I shall assume that relative values range from zero to one, but this assumption is not critical since appropriate adjustments can be made in what follows to take account of any scale of relative values, including ones with negative values.

The expected relative value (EV) of a choice situation to a particular individual (A) is given by

$$EV_A = \sum_i \sum_j P_i E_{ij} V_j .$$

EVs have a maximum value of one and a minimum value of zero.

COOPERATION AND CONFLICT

Consider two individuals, A and B. Let $(EV_A | B)$ represent the expected relative value to A of his choice situation when B is present in it; and $(EV_A | B')$ represent this value when B is not present in it. $(EV_B | A)$ and $(EV_B | A')$ are the corresponding expected relative values for B.

11.1. Cooperation, Conflict, and Independence. In a particular state (S) if

(a) $(EV_A | B) > (EV_A | B')$, then B cooperates with A,
(b) $(EV_A | B) < (EV_A | B')$, then B conflicts with A, and
(c) $(EV_A | B) = (EV_A | B')$, then A is independent of B

Therefore, if B's presence increases the value of A's state, B cooperates with A; if B's presence reduces this value, he conflicts with A; and if he has no effect on A's expected relative value, A is independent of B.

11.2. Degree of Cooperation and Conflict. The degree of cooperation of B with A is

$$DC_{BA} = (EV_A | B) - (EV_A | B').$$

The degree of conflict of B with A is

$$DC'_{BA} = 1 - DC_{BA} = 1 - [(EV_A | B) - (EV_A | B')].$$

This measure can take on values from -1 to +1. Negative values of the degree of cooperation represent conflict, and conversely. Note that cooperation and conflict exhaust the ways in which one individual can affect the expected relative values of another.

There is nothing in the definitions of cooperation and conflict that requires either of the parties to be conscious of, or to intend, his effect on the other. One person may inadvertently affect another of whose presence he may not even be aware; for example, when one person begins to use a telephone an extension of which is being used by another.

11.3. Degrees of Cooperativeness, Hostility, and Independence. If, in an environment occupied by A and B, (a) B's potential courses of action can be grouped into three exclusive and exhaustive classes,

C_1 : courses of action which have efficiency equal to 1.0 for increasing EV_A,
C_2 : courses of action which have efficiency equal to 1.0 for decreasing EV_A, and
C_3 : courses of action which have no affect on EV_A

(b) all the courses of action are equally efficient for all outcomes desired by B, and (c) B is aware of these efficiencies, then

P_1 = his degree of cooperativeness toward A,
P_2 = his degree of hostility toward A, and
P_3 = his degree of indifference toward A.

One individual may be cooperative or hostile toward another in a particular situation because of the affect that the other is having on him. Stimulated hostility is ascendancy, unstimulated hostility is aggressiveness.

11.4. Degree of Aggressiveness of one individual (B) toward another (A) is his degree of hostility toward A when A is having no effect on EV_B.

11.5. Degree of Ascendance (Submission) of B toward A is the degree B's hostility toward A when A is in hostility with B.

Ascendancy is stimulated hostility and reflects a desire to "get even" with an aggressor. One can be ascendant without being aggressive; that is, inclined toward hostility only if provoked. Although one could be aggressive without being ascendant, it does not seem likely to occur. This, however, raises a question of fact that remains to be answered. (See Appendix I for a detailed discussion of ascendance-submission.)

EXPLOITATION

The degree to which one individual (B) cooperates or conflicts with another (A) does not have to equal the degree to which A cooperates or conflicts with B. Thus two individuals may affect each other differently. This difference is a measure of exploitation.

11.6. Degree of Exploitation. The degree to which one individual (B) exploits another (A) is

$$DX_{BA} = DC_{AB} - DC_{BA}$$

and the degree to which A exploits B is

$$DX_{AB} = 1 - DX_{BA} = DC_{BA} - DC_{AB}.$$

This measure can range from -2 to +2. By use of this measure we can distinguish between three kinds of exploitation. If DC_{AB} and

DC_{BA} are both positive but unequal, then the two individuals cooperate with each other, but unequally. The one who benefits the most can be said to be a benevolent exploiter of the other. This is the type of exploitationthat most colonial powers have claimed for themselves when they have admitted to exploiting their colonies. Many employer-employee relations can also be characterized by this type of relationship.

If DC_{AB} and DC_{BA} are both negative but unequal, then A and B are in conflict with each other, but unequally. The one who suffers least can be said to be the malevolent exploiter of the other. Such an exploiter is one who is willing to suffer if he can make another suffer more than he is. This is usually the case where revenge is involved. Many wars are examples of malevolent exploitation.

Finally, if one of the parties cooperates with the other but the other is in conflict with him, we have a case of what might facetiously be called "normal" exploitation. This seems to characterize the historic relationship between slave and master.

The degree of exploitation is the difference between the degree of conflict of A with B and B with A, and hence is a measure of the asymmetry of the effects that two individuals have on each other. The sum of these degrees also has significance.

11.7. Intensity of Cooperation (Conflict) between two individuals is the sum of the degrees of cooperation (conflict) between them.

This sum has meaning only if A and B are in cooperation or conflict with each other (i.e., the signs of DC_{AB} and DC_{BA} are the same). Negative values represent intensity of conflict and positive values intensity of cooperation. Minimum and maximum values are -2 and +2, respectively. If $DC_{AB} = DC_{BA} \neq 0$, then even though there is no exploitation there is an intensity of conflict or cooperation.

Intensity can increase as exploitation decreases and exploitation can increase as intensity decreases. On the other hand they may increase or decrease together.

11.8. <u>Escalation (De-escalation) of Conflict (Cooperation)</u>: an increase in the intensity of conflict (cooperation) between two or more parties.

One seldom hears about the escalation (de-escalation) of cooperation but it is clearly as significant as escalation (de-escalation) of conflict.

COMPETITION

In the literature of psychology, social psychology, and sociology there is a good deal of qualitative discussion about the difference between conflict and cooperation. One of the more commonly cited differences is "the presence of physical force" in conflict, and its absence in competition. This difference does not seem valid to me because, for example, a prize fight is normally thought of as a competition while a street brawl is thought of as conflict. Although the use of force or physical contact does not seem to be essential to conflict, it can play an important role in it, a role that I consider below.

I make no attempt here to survey the copious literature on the distinction between conflict and cooperation, but I do want to cite the most suggestive difinition that I have found, that of Katz and Schanck (1937). In essence, they argued that competition is conflict <u>according to rules</u>, and hence is <u>contrained</u> conflict. This does distinguish between a prize fight and a street brawl but, although I can think of no case of competition that does not have rules, I can think of instances of conflict that also have rules. Wars, in contrast, to riots, have rules, but I do not believe war is competition, Waring nations are not supposed to use chemical and biological weapons, are supposed to treat prisoners and civilians in certain ways, and so on. But these rules are frequently broken and there is no authority to enforce them and to punish the offender. Therefore, although rules seem necessary for competition, they do not seem to be sufficient. However, I think the essential difference between conflict and competition can be found in the function of those rules which operate where competition occurs.

In a prize fight and other sporting events rules are imposed by an authority to protect the interests of both the participants and the audience. In economic competition governments impose rules to protect the public, if not the participants. The rules in economic competition do not prevent elimination of a participant but they usually reduce the likelihood of such an occurrence. In a private tennis match or chess game, rules are not imposed by an authority but they are accepted voluntarily by the players because doing so serves their interests. Therefore, although conflict appears in competition, it appears to be constrained by rules to serving the purpose of the participants or a third party. Let me try to make this more precise.

11.9. Competition. Two individuals, A and B, are in competition in an environment (S) if the following conditions are satisfied:

(a) A's degree of intention for outcome O_1 in S is greater than his degree of intention for another outcome O_2. The converse holds for B.

(b) Of the set of courses of action available to A and B in S $\{C\}$ there is a subset $\{C^*\}$ such that choices of either A or B of any member of this subset that increases (decreases) the probability of O_1 occurring in S, decreases (increases) the probability of O_2 occurring in S. (O_1 and O_2 cannot both occur simultaneously.)

(c) There is a third outcome (O_3) possible in S which may occur with O_1 or O_2 and which is intended either by both A and B or by a third party (T).

(d) The conflict between A and B relative to O_1 and O_2 in S is efficient relative to O_3 in S. *

(e) If A (B) selects a course of action which is not in the subset $\{C^*\}$ the other individual or T can punish him with respect to his pursuit of O_1 (O_2).

It should be apparent that this definition can easily be extended to cover more than two participants.

When the "third" (or cooperative) outcome (O_3) in competition is an objective of a third party (e.g., an audience) and <u>not</u> of the participants in the conflict, the competition can be said to be <u>extrinsic</u>. If O_3 is a common objective of the conflicting parties, the competition can be said to be <u>intrinsic</u>. Competition may therefore be both intrinsic and extrinsic as, for example, in a tennis match between friends before an audience.

In intrinsic competition the ratio of the degree of intention of each competitor for the conflicting objective (O_1 or O_2) to his degree of intention for the cooperative objective determines whether the competition is dominantly conflict- or cooperation-oriented for him.

The rules of competitive behavior which define the subset of permissable courses of action $\{C^*\}$ are accepted by, and/or imposed on, the participants in order to assure the effectiveness of the conflict

---

*In many cases a stronger condition is satisfied: the probability of O_3 occurring in S increases as the intensity of conflict between A and B relative to O_1 and O_2 increases. For example, the entertainment value of a sporting event generally increases as the intensity of the conflict between participants increases.

for the third (cooperative) objective. For example, in economic competition, conflict between rival companies is supposed to serve the interests of the consumer. Laws and regulations are enforced to make sure this is the case. In our economic system, for example, it is illegal for two competitors to "fix" prices so as to exploit the consumer. They may, however, exploit each other. Infraction of the rules can result in punishment of the guilty party by the government either by the imposition of fines or further restrictions on choice. In intrinsic competition an infraction of the rules by one party allows the other to impose some kind of penalty on him.

WAYS OF AFFECTING CONFLICT

The nature of conflict is such that either the environment in which it takes place, or the behavior of one or more of the participants must be changed if the conflict is to be removed or reduced in intensity. To attempt to remove a conflict by changing its environment is to attempt to dissolve it; to do so by changing the participants is to resolve it. These modes of affecting conflict are available to third parties as well as to the participants.

11.10. Dissolution of Conflict: a change in the environment of a conflict so that the participants no longer conflict with each other.

11.11. Resolution of Conflict: a change in the behavior of one or both of the participants so that they no longer conflict with each other.

To dissolve or resolve a conflict is to remove it. Curiously, however, when we speak of solving a conflict--as we do in the context of the Theory of Games--we do not necessarily imply removal of the conflict. To solve a conflict is to do as well as possible in the conflict situation.

11.12. Solution of Conflict: selection of that course of action by a participant from among those which are available to him which maximizes his expected relative value in the conflict situation.

Thus dissolving a conflict involves changing the environment, resolving it involves changing someone other than oneself, and solving it involves changing ones own behavior. Although these modes are exhaustive, they are by no means exclusive. Let us examine them in more detail.

Dissolving Conflict

There are several things that can be done to the environment which may change the nature of one party's effect on the other. First, the environment can be modified so that the behavior of one participant no longer has an effect on the other. This is equivalent to separating the opponent from the one affected. For example, if A is flashing a light in a room in which B is trying to read and thus comes into conflict with B, a screen can be placed in a position so that it blocks the light. Note that this may leave both A and B doing what they were previously, but the conflict is removed. The objective was not to change A's behavior, but to change its effect on B. The efficiency of such an effort clearly depends on whether A intends to conflict with B. If he does, such a separation of A and B is not likely to succeed because, for example, A may remove the screen or initiate distracting noises. Many conflicts that arise from unintended intrusions on the senses can be removed by modifying the environment.

Secondly, conflicts that arise out of scarcity can often be dissolved by making available more of whatever is scarce. For example, if two children want the same ball and are in conflict over it, the conflict may be removed by providing a second ball of the same type as the first. If it is impractical or impossible to duplicate what is wanted then, obviously, so is this mode of dissolving conflict. But where it is practical and possible it is an attractive way of removing conflict because it does not involve changing the behavior of the opponents. It effectively separates the opponents.

Separation can also be accomplished by removing one or both of the conflicting parties from the environment. If A can induce B to leave, A has resolved the conflict. If he cannot induce B to do so, he may use physical force to remove B or, what is equivalent, he may incapacitate B in the original environment. If force is used to remove or incapacitate a participant in a conflict we have what Rapoport (1961) called a fight. A fight may dissolve the conflict from the point of view of the victor, but it does not do so from the point of view of the vanquished. As a result the hostility of the vanquished toward the victor is usually increased so that if an opportunity later presents itself he is likely to initiate another conflict, one that is often more intense than the first. Hence a conflict is not usually dissolved or resolved by a fight. It is usually suppressed temporarily and subsequently escalated.

Resolution of Conflict

Note that in a fight one participant attempts to remove the opponent by changing some of his relevant structural properties (e.g., his location or physical ability to act). This is done in order to affect at least one of his functional properties, his probability of selecting conflicting behavior. One may change this or some other functional property of an opponent without affecting him structurally. For example, one can make the cost of an opponent's selecting a course of action that produces conflict greater than the gain that he can expect from it. The imposition of a threat on one or both parties of a conflict is to attempt to deter the conflict. The threat may be issued either by a participant or by a third party (e.g., the government). The law and police are intended to deter potential criminals.

The threat of punishment or retaliation can be effective only if two conditions are satisfied: (1) the recipient of the threat is aware of the deterrent and he perceives the expected costs imposed by it as greater than he can expect if he ignores it; and (2) he believes the deterrent will only be used if he selects the undesired course of action. If he does not believe the latter--that is, he believes the deterrent will be used against him no matter what he does--then it may very well increase his chances of selecting the course of action that it is intended to prevent. Hence the danger in a national policy based on massive deterrence lies in a nation's inability to convince others that this capability will not be used without the specified provocation. Similar remarks can be made about rewards for cooperation.

Note that the use of deterrents may not remove a conflict but only prevent it from escalating. It should also be observed that the use of deterrents is unlikely to reduce hostility even where it reduces the intensity of conflict.

To make an opponent aware of a deterrent or a potential reward may require communicating with him, but communication may affect conflict in other ways. Let us examine these.

<u>Resolving Conflict by Communication</u>. One party to a conflict may use communication to affect the other's behavior either by <u>informing</u> him (changing his probabilities of choice), by <u>instructing</u> him (changing the efficiency of his choice), by <u>motivating</u> him (changing the values that he places on outcomes), or by some combination of these. What Rapoport has called a <u>debate</u> is only one way of using communication to resolve a conflict: it is one directed toward changing those beliefs and/or attitudes that produce conflicting behavior.

Suppose one child (A) wants the ball that a second child (B) is playing with. There may be a second similar ball in the environment of which child A is not aware, but child B is. B may inform A of the availability of the second ball and thus resolve the conflict. Obviously, this information may also be conveyed by a third party; for example, a parent.

In general when one party (A) knows how another (B) can get what he wants without conflicting with A, and B does not know this, then A may inform B of the possible choice which will avoid or remove conflict. A third party may resolve a conflict by informing both parties of appropriate alternatives to what they are doing.

Suppose two persons want to use the one electronic computer that is available in the environment. If either tries to get exclusive use of the computer because he believes only one can make efficient use of it at a time, conflict results. If one party or an outsider instructs the one or ones who seek exclusive use of the computer on how to run both problems simultaneously, the conflict may be avoided or resolved. In general, instructions may be used to avoid or resolve conflict where by more efficient use of a course of action already selected by one or both parties, they can both obtain what they want without conflicting with the other.

Finally, if each of two persons in the same environment want something that both cannot have (e.g., two children who want the same ball), conflict may be avoided or resolved by changing the desire of one or both parties through motivational communication. For example, a parent may attempt to distract the child by interesting him in something other than the ball.

When both conflicting parties communicate with each other in an attempt to resolve or prevent escalation of conflict they can be said to be <u>negotiating</u>.

11.13. Negotiation: communication between the parties of a conflict, which they intend either to dissolve or resolve their conflict, or to prevent its escalation.

The way in which a negotiation is organized and the environment in which it is carried out can have considerable effect on its chances for success. Even such things as the arrangement of the room in which negotiations take place can influence the outcome. Rapoport's (1960) discussion of ways to make a debate more productive of conflict resolution is relevant to negotiation as well.

Negotiation is often facilitated by a third party, a mediator.

11.14. Mediator: an individual who is present at negotiation of of a conflictof which he is not a participant, whose function it is to increase the probability that communication between the conflicting parties produces a resolution of the conflict.

Many conflicts cannot be resolved without outside intervention. It has been pointed out, for example, that one of the reasons that many conflicts between nations are so difficult to resolve is that there is no "third" nation that the conflicting nations respect equally as a neutral. Even in such cases it would still be possible to resolve conflicts if there were a third party that was strong enough to impose its will on those involved. If there were such a third party--for example, an effective world government--national conflicts (like many labor-management disputes) could be arbitrated.

11.15. Arbitrator: an individual who resolves a conflict, to which he is not a party, or prevents its esclation by selecting the courses of action to be followed by the participants in the conflict.

The courts often serve as arbitrators; for example, in civil cases. Even in criminal cases the courts can be looked at as arbitrating a conflict between the accused and the state. Governments or governing bodies of organizations may appoint arbitrators for disputes between their members and impose such arbitration on them, using their power of punishment to make the imposition. In some cases (e.g., in many labor-management disputes) the parties to the conflict themselves agree to the selection of an arbitrator and to be bound by his decision.

In many negotiations and arbitrations the objective is not so much to resolve conflict as it is to prevent its escalation. Escalation is likely to occur when one or both parties to a conflict believe they are exploited by the others. Hence most negotiations and arbitrations are directed to removing exploitation, not conflict. The parties involved are primarily motivated by a desire "not to be taken advantage of." A conflict in which neither escalation nor de-escalation takes place can be said to have reached equilibrium or stability.

Inducing and Imposing Cooperation. In a fight at least one participant attempts to impose his will on the other. To impose behavior on someone is to give him no choice. Deterrents, rewards, and communication are used to induce (not impose) behavior which is preferred by the user. To induce behavior is not to remove choice. Rapoport (1961) argued that it is not possible to induce (produce choice of) desired behavior by use of physical force:

> To induce an action...is most physically impossible. The most you can do is offer a choice between alternatives, for example, "Sign this or die." We call such an offer intimidation by use of force, but in the last analysis, it is the Other who makes the choice. If he chooses not to sign, he cannot be forced to do so, because his nervous system and his muscles cannot be controlled by another in coordinated fashion (p. 215).

It is because of this apparent inability to impose cooperation on conflicting parties that pacifists have to be so passive. They can be passively against conflict and war, but not aggressively for peace. Put another way, one cannot impose cooperation on another without a fight, or at least so it seems.

The ancient Greeks endowed Cupid with the ability to impose love on another without the use of physical force. He was equipped with a unique bow and arrows for this purpose. If such instruments were generally available it would change the entire logic of war and peace. For example, if one person, A, behaved aggressively toward another, B; B might "shoot him" with cooperativeness and thus impose a change of attitude on A without denying him choice. Then if A wanted to retaliate, he would shoot B with cooperativeness toward him.

Even Cupid's bow and arrows could be used as instruments of conflict. For example, one person could inflict cooperativeness on another in order to make it easier to destroy him. The ideal instrument of peace, therefore, would be one that is so designed that its user could not impose cooperativeness on another without doing so to himself. Instruments such as the Greeks gave Cupid, or the peace pills or gases that many have dreamed of, could not provide a permanent removal of conflict by themselves. The way in which they would be used is critical.

Instruments to impose cooperation are becoming a reality. In the May 1966 issue of Esquire, in an article entitled "Mind Control is Good, Bad (Check One)" (pp. 106-109), A. J. Budrys reviewed recent technological developments which make it possible to impose cooperativeness and other functional properties on men and animals, at least under laboratory conditions. We already have the makings of "conflict decontamination chambers." The question of how well we will use the power of Cupid, once we have it, remains open.

Solution of Conflict

When an individual finds his effectiveness reduced by the behavior of another, or his own behavior reducing the effectiveness of another, he may either remove himself from the conflict environment or change his behavior in that environment. If he restricts himself to looking for a course of action which is intended to minimize the undesirable effect which another person has on him, then he treats the conflict as what is now commonly called a "game" and he seeks what is called a "solution" to it. In such cases, a "game" is used as a representation or model of the conflict situation. Many, if not most, theories of conflict are based on such representation, a consequence of which I now examine.

REPRESENTATIONS OF CONFLICT

Theories dealing with conflict behavior are frequently classified as normative or non-normative. Normative theories attempt to determine what choice a participant in a conflict ought to make. Non-normative theories attempt to predict, and sometimes explain, what choices individuals actually make in such situations. This distinction is not as clear in practice as it is in principle. For example, when a participant in a conflict does not do what a normative theory says he ought to do, then some explanation is required, and only a non-normative theory can provide it. Furthermore, normative theories of conflict have been used repeatedly as though they were predictive theories, however inappropriate it has been to do so.

Whichever type of theory a researcher attempts to construct, it is apparent that he must employ some way of representing conflict situations. The most common way of doing so was developed by von Neumann and Morgenstern in their work on the Theory of Games. Each participant is assumed to have a well specified set of alternative courses of action from which he can select only one at a time. It is also assumed that for each combination of choices by the participants there is a well-defined outcome, the relative value or utility of which to each participant is known by the researcher. Consequently, conflicts are represented by what is called a payoff matrix in which the possible choices and the relative utilities or values of each possible outcome are shown. Figure 11.1 is a payoff matrix for a very simple conflict (game) which involves two participants (A and B), to each of whom two choices (X and Y) are available. (The choices available to the participants need not be the same, or the same in number.) The first number in each cell is A's payoff and the second is B's. For example, for the pair of choices, X by A and Y by B--which I represent by (X, Y)--A receives 3 units of relative utility or value and B loses 3 units, which I represent by (3, -3).

		B Chooses: X	B Chooses: Y
A Chooses	X	1, -1	3, -3
	Y	2, -2	4, -4

Fig. 11.1 A Payoff Matrix

This representation of conflict appears to be relevant only to simple one-play games. However, von Neumann and Morgenstern (1953) have shown that if one considers rules of choice (i.e., strategies) in any interaction involving a finite sequence of choices, then these can also be represented in principle by the so-called "normal form" described above. In such a representation the choices are from among strategies, not plays. In practice it is still not possible to so represent many complex interactions (e.g., the game of chess) because of the very large number of possible strategies which are involved.

Using such a representation of a two-person conflict situation, various normative game theorists have deduced what each participant ought to do from certain assumptions about his state of knowledge and what consitutes "rationality." The participants are usually assumed to have the knowledge represented in the payoff matrix. They are assumed to be rational if when confronted by a choice between outcomes over which they have a preference ordering, each selects that one which he most prefers if he believes that he will get it if he selects it and will not if he does not. From this it follows that a rational person is one who tries to maximize his expected relative value or utility. It is assumed, of course, that parties to a conflict ought to act rationally in this sense.

One of the more important concepts that emerges in normative theories of games is that of an equilibrium point. Such a point consists of a combination of choices (one by each participant) such that if either participant were to alter his choice without the other doing so, he would do no better for himself and might do worse. For example, in the conflict situation represented in Figure 11.1, the combination of choices (Y, X) yields an equilibrium point. If A were to change his choice to X and B were to retain X, A's payoff drops from 2 to 1. On the other hand, if B were to change his choice from X to Y and A were to retain Y, B's payoff would drop from -2 to -4. (X, Y) is the only equilibrium point in this matrix. (Y, Y) is not an equilibrium point, for example, because if B changes from Y to X, A retaining Y, B's payoff increases from -4 to -2. A payoff matrix may have more than one equilibrium point, or none.

The normative theorists argue that both participants, if rational, will not be satisfied with an outcome that is not at an equilibrium point because if the outcome is not at such a point, one or both of the participants can improve his payoff by changing his choice. If he is rational, by definition he will do so. Hence, it is further argued, rational players ought to select courses of action which yield an equilibrium point.

This much is argued for one-play non-cooperative interactions. It so happens that in some games, one of which (the Prisoner's Dilemma) is discussed in detail below, in which repeated choices are made, all equilibria consist of repetitions of one-play equilibria. For such games, the theory asserts that no other type of outcome can be stable in the long run (i. e., if repeated choices are allowed). When the players converge on an equilibrium point the game is said to be solved; each has made the best choice possible under the circumstances which prevail.

Not every conflict has a solution in this sense, not even simple two-person two-choice conflicts. For example, see the conflict represented in Figure 11. 2.* It has no equilibrium point.** In some cases even a simple conflict can have two equilibrium points

		B Chooses X	B Chooses Y
A Chooses	X	1, -1	4, -4
	Y	3, -3	2, -2

Fig. 11. 2 A game with no equilibrium point.

and, hence, two solutions.

---

*The games represented in Figures 11. 1 and 11. 2 are called "zero-sum" because the sum of the payoffs in each cell is equal to zero. The game represented in Figure 11. 3 is not zero-sum.

**This is true if the participants are restricted to "pure strategies," but not so if they can select "mixed strategies;" that is, ones which involve a random choice from the alternatives with pre-determined probabilities of selecting each.

For example, see Figure 11.3.

		B Chooses X	B Chooses Y
A Chooses	X	3, -1	1, -2
	Y	2, -4	4, -3

Fig. 11.3. A game with two equilibrium points.

Game theory leads to both logical and empirical difficulties. Logical difficulties arise in a type of conflict situation first noted by Merrill M. Flood in 1951 and later explicitly formulated by Albert W. Tucker who gave it the name, "Prisoner's Dilemma." The payoff matrix for this situation is shown in Figure 11.4 where the numbers in the cells represent A's and B's preference orderings of outcomes, 4 being the most preferred and 1 the least preferred.

		B Chooses X	B Chooses Y
A Chooses	X	3, 3	1, 4
	Y	4, 1	2, 2

Fig. 11.4. Payoff matrix for Prisoner's Dillemma.

The one equilibrium point in the Prisoner's Dilemma occurs at (Y, Y). Furthermore, Y is the best choice for either player, no matter what the other chooses. But clearly (X, X) yields an outcome whose payoff (3, 3) is preferred by both. Therefore, it seems rational for the participants to select (X, X). However, (X, X) does not yield an equilibrium point because each player can increase his payoff if he alone changes his choice. Hence the paradox: one application of the principle of rationality dictates that the participants should select (X, X), and the other that they should select (Y, Y).

The empirical difficulty arises in long sequences of "plays" of this game, such as have been carried out by Rapoport (1965), in which subjects often stabilize at (X, X) rather than at (Y, Y).

Commenting on the logical problem, Rapoport (1967) wrote:

> Like other paradoxes, this one was denied by some logicians and worshipped by others. The deniers declared the unfavorable outcome of the game was a realistic fact of life that in no way changed the fact that the players' choices were based on the "rational pursuit of self-interest." The worshippers saw the impasse as a new manifestation of the unsatisfactoriness of the human condition. A number of decision theorists, however, undertook to wrestle with the paradox, and as far as I know [Nigel] Howard was the first to succeed (p. 54).

Curiously, Howard (1966) was not so much interested in solving the paradox as he was in predicting correctly those combinations of choices that would produce (empirical) stability, long runs of the same choice combinations.

As far as I know Howard was the first to question the way in which conflict is represented in the Theory of Games. Others have concerned themselves with the assumptions incorporated into the theory or the deductions made from them or aspects of the conflict situation which they believed to be omitted from the theory.* But the way of representing a conflict seemed so appropriate and undebatable that it was not brought into question.

Howard's basic insight was that the payoff matrix is the <u>researcher's way of conceptualizing a conflict</u>, but <u>not</u> the participants' way of doing so. Even when a participant is presented with a payoff matrix to represent the situation he is in, he transforms it, consciously or unconsciously, into another type of matrix.

Each participant in a conflict predicts what his opponent is going

---

*This was true of Rapoport and the author. For example, see Rapoport (1959) and Ackoff (1959).

to do even if he has never met his opponent. As a minimum he will predict that the opponent will do as he would if he were in the opponent's position. Even in contrived experimental situations each opponent usually has some information about what the opponent is like (e.g., a fellow student or another housewife). In real situations a great deal of information about the opponent is usually available. Whatever the situation, real or contrived, each participant operates with some predictions of what the other will do, and is aware of the fact that his opponent does likewise.

Each participant formulates a set of possible <u>policies</u>; that is, conditional rules of choice. For example, if each player has two possible choices (X and Y), then each player has four possible policies:

(1) choose X no matter what opponent does (X/X),
(2) choose Y no matter what opponent does (Y/Y),
(3) choose X if he does and Y if he does (X/Y), and
(4) choose X if he chooses Y, and Y if he chooses X (Y/X).

Therefore, each participant (say A) can represent his conception of his opponent's (B's) conception of the conflict (e.g., the Prisoner's Dilemma) as is shown in Figure 11.5.

A's Choices	B's Policy-Choices (X/X)	(Y/Y)	(X/Y)	(Y/X)
X	3, 3	1, 4	3, 3	1, 4
Y	4, 1	2, 2	2, 2	4, 1

Fig. 11.5. A's conception of B's conception of the Prisoner's Dilemma.

The entries in the cells are the payoffs associated with A's choices and B's policy-choice. For example, if A chooses Y and B follows policy (X/X), B will select X, and the payoff associated with (Y, X) is

(4, 1). (See Figure 11. 4.) Note that in this expanded matrix (Figure 11. 5) there is still only one equilibrium point: where A plays Y and B follows policy (Y/Y) and hence also selects Y. This is the same combination of choices that yields equilibrium in the original payoff matrix.

Now it is natural for A to formulate for himself policies with which to meet B's policies; that is, meta-policies, rules of choice conditional on B's policy choice. (Investigations show that each participant actually does predict what policy his opponent will follow as well as what choice he will make.) There are sixteen meta-policies that A can formulate. These are shown in Figure 11. 6 together with the payoffs associated with each combinatinn of policy and meta-policy choices. Howard calls such a matrix a metagame payoff matrix.

Note that three equilibrium points appear in this matrix, including two with payoffs of (3, 3). These clearly are preferable to the equilibrium with payoff (2, 2). Howard's theory asserts that if stability is reached it will be reached at one of the metagame equilibrium points. Furthermore, he predicts longer-run stability will be reached at the equilibrium point(s) prefered by both participants; that is, at (3, 3) rather than (2, 2).

A's Meta-Policy Choices	B's Policy Choices (X/X)	(Y/Y)	(X/Y)	(Y/X)
(X/X/X/X)	3, 3	1, 4	3, 3	1, 4
(X/X/X/Y)	3, 3	1, 4	3, 3	4, 1
(X/X/Y/X)	3, 3,	1, 4	2, 2	1, 4
(X/Y/X/X)	3, 3	2, 2	3, 3	1, 4
(Y/X/X/X)	4, 1	1, 4	3, 3	1, 4
(X/X/Y/Y)	3, 3	1, 4	2, 2	4, 1
(X/Y/X/Y)	3, 3	2, 2	3, 3	4, 1
(Y/X/X/Y)	4, 1	1, 4	3, 3	4, 1
(X/Y/Y/X)	3, 3	2, 2	2, 2	1, 4
(Y, X, Y, X)	4, 1	1, 4	2, 2	1, 4
(Y, Y, X, X)	4, 1	2, 2	3, 3	1, 4
(X/Y/Y/Y)	3, 3	2, 2	2, 2	4, 1
(Y/X/Y/Y)	4, 1	1, 4	2, 2	4, 1
(Y/Y/X/Y)	4, 1	2, 2	3, 3	4, 1
(Y/Y/Y/X)	4, 1	2, 2	2, 2	1, 4
(Y/Y/Y/Y)	4, 1	2, 2	2, 2	4, 1

Fig. 11. 6. Metagame payoff matrix for Prisoner's Dilemma.

In the description given above only A's conception of the conflict has been considered, but clearly B's conception can be similarly developed. In the case of the Prisoner's Dilemma, both A's and B's metagame payoff matrices are identical. This is not so for all conflict situations. Howard has considered such cases as well.

It is apparent that the reasoning process which produces the metagame payoff matrix can be extended. B can formulate policies for response to A's meta-policies, then A can formulate policy responses to these, and so on. But Howard has shown that if there are n participants in a conflict then any expansion beyond the n^{th} policy level will reveal no equilibrium points not revealed in the n^{th}. Hence there is no need for a player's conception in a two-person game to go beyond his meta-policies.

It has also been shown that any equilibrium point in the original payoff matrix will appear as an equilibrium point in the metagame payoff matrix. But the metagame payoff matrix may reveal equilibrium points which were not revealed in the original payoff matrix (as is the case in the Prisoner's Dilemma).

Howard has conducted experiments which support his theory. The empirical as well as the logical paradoxes appear to be removed by it. The moral in his effort is that the researcher of purposeful behavior should not assume that his conception of a subject's choice situation corresponds to his subject's conception of it. Understanding the subject's conception may provide the key to predicting and explaining his behavior.

Metagame Theory identifies the points at which stability of conflict will be reached, <u>if</u> it is reached. It does not predict that stability will be reached, or if reached, by what path. In Appendix III I describe a theory developed by James Emshoff that provides explanations and predictions of individual choices in simple conflict situations. As will be seen, Emshoff's work makes liberal use of metagame concepts and parts of the conceptual system which has been developed here.

CONCLUSION

In experimental conflict situations in which communication between participants is prevented, it has been observed that the participants attempt to make their intentions known to, and influence, the other by their actions alone. People tend to cooperate more when they can communicate with each other than when they cannot do so. An even stronger observation has been made: even when communication is possible but is not used, people tend to cooperate more than where it is prohibited. (These and related findings are reported in Management Science Center, 1967.) This indicates that the mere possibility of communication tends to reduce the hostility between conflicting parties. In the experimental work in this area with which I have been involved, communication has had a greater positive effect on cooperativeness than any of a number of other variables tested.

It is a curious characteristic of our culture that we have expended much more research effort on increasing our effectiveness in conflict than on doing so in cooperation. An ability to do the first does not imply an ability to do the second. This imbalance in allocation of research effort and resources is reflected in our greater ability to wage war successfully than to so wage peace.

Cooperation is not merely the absence of conflict. Furthermore, even if two parties cooperate with each other, one may be the benevolent exploiter of the other. Even this kind of asymmetry breeds conflict (e.g., the outcome of benevolent colonialism). Hence reduction of exploitation among cooperators is as difficult a problem as any involving the control of conflict.

Social groups are normally held together by cooperative interactions among their members. It is not surprising, therefore, that most of the research on cooperation that has been done, has been done as part of research on group behavior. It is to such behavior that we now turn.

BIBLIOGRAPHY

Ackoff, R. L., "Games, Decisions, and Organizations," General Systems, 4(1959), 145-150.

Howard, Nigel, "The Theory of Meta-Games," General Systems, 11(1966), 167-186.

________, "The Mathematics of Meta-Games," General Systems, 11(1966), 187-200.

Katz, Daniel and Schanck, R. L., Social Psychology. John Wiley and Sons, New York, 1938.

Management Science Center, A Model Study of the Escalation and De-escalation of Conflict (mimeographed report). University of Pennsylvania, Philadelphia, March 1, 1967.

Rapoport, Anatol, "Critiques of Game Theory," Behavioral Science, 4(1959), 49-66.

________, Fights, Games and Debates. The University of Michigan Press, Ann Arbor, 1960.

________, "Three Modes of Conflict," Management Science, 7(1961), 210-218.

________, The Prisoner's Dilemma. The University of Michigan Press, Ann Arbor, 1965.

________, "Escape from Paradox," Scientific American, July 1967, 50-56.

Shubik, Martin, "Bibliography of Simulation, Gaming, Artificial Intelligence and Allied Topics," Journal of the American Statistical Association, 55(1960), 736-751.

von Neumann, J., and Morgenstern, O., Theory of Games and Economic Behavior. Princeton University Press, Princeton, (3rd edit.) 1953.

Chapter 12

SOCIAL GROUPS AS TELEOLOGICAL SYSTEMS

> MAN, n. An animal so lost in rapturous contemplation of what he thinks he is as to overlook what he indubitably ought to be. His chief occupation is extermination of other animals and his own species, which, however, multiplies with such insistent rapidity as to infest the whole habitable earth and Canada. (Ambrose Bierce, The Devil's Dictionary).

INTRODUCTION

It is possible, in principle, to reduce any property of a physical body to some function of the properties of the point-particles of which it is composed. For example, the temperature of a body, as we observed earlier, is a function of the velocity of the particles of which it is composed. It is obviously easier to determine temperature (and other properties of the collection of particles that form a body) holistically than to do so atomically.

The same is true for social entities. Their properties can be expressed, in principle, as a function of the properties of the (psychological) individuals that make them up, but it is often easier to do so holistically. The relative ease of a holistic approach increases as the size and complexity of the group's structure increases. We can, for example, usually determine the attitude of an audience toward a performance by observing its collective behavior rather than by determining the attitudes of each of its members and aggregating them in some way.

From knowledge of the velocity of each particle making up a body we can determine its temperature. But from knowledge of its temperature we cannot determine the velocity of each of its particles. Hence, temperature is truly a collective property of a body. The attitude of an audience is similarly a collective property; from a knowledge of an audience's attitude we cannot infer the attitude of any particular member of it.

Collections of entities can themselves be conceptualized as entities. Whether or not it is fruitful to do so depends on our interests. For example, a teacher may be more interested in the uniqueness of each member of her class than in the class' collective properties. The school's principal may not have this interest; from his point of view only knowledge of the collective properties of each class is necessary. To the superintendent of schools, the school itself may be a more suitable unit with which to deal.

Although it is possible to infer a property of a collection from properties of its elements, this does not mean that properties of the elements are in some (ontological or epistemological) sense more basic or fundamental than the collective property. For example, although knowledge of attitudes of each member of a group may yield knowledge of the group's attitudes, we may not be able to explain a member's attitude unless we know the collective attitude. The individual both influences and is influenced by groups. For example, we have already considered the fact that language, a group product, influences the thought processes of the group's members. Hence to understand how individuals think we must understand their language.

We have already seen how a purposeful individual can be conceptualized as a teleological system. My objectives in this chapter are (1) to show that the groups of purposeful individuals can be similarly conceptualized and (2) to indicate how the concepts required to study groups can be related to those developed here for study of the individual. In addition, I also hope to show how feedforward from psychology to sociology and feedback from sociology to psychology can be facilitated. Such feedforward and feedback between mechanics and physics has been a major factor in the development of these sciences.

SOCIAL INDIVIDUALS AND GROUPS

12.1. Social Individual. Any collection of psychological individuals that can be individuated. *

Individuation of a social individual requires specification of the rule(s) for inclusion and exclusion of psychological individuals from the collection over a specified period of time. For example, "the people whose names appear on a specified page of a telephone directory" constitute a social individual. "Residents of New York City who own automobiles" do also. In the former the specification of membership is denotative since the members are identified by name; in the second the specification is connotative since membership properties are given. It is possible in principle, of course, to translate any connotative specification into a denotative listing of members.

The "populations" used in social surveys of any kind (e.g., censuses or market surveys) are social individuals. Identification of those who are members of such a collection may be difficult, as those who have conducted social surveys are well aware. Defining a population is the same thing as identifying a social individual.

What I have called a social individual is sometimes broken into two different types of entities: a category and an aggregation.

*See definition 4.25.

According to Cuber (1959):

> A category is any number of persons who are thought of together, whether they are in communication or not.
>
> An aggregation is a collectivity of persons who are held together in a physical sense by some factor other than intercommunication (p. 298).

Thus, "the people whose names appear on a specified page of a telephone directory" would be a category in this sense whereas "those visiting a sea-side resort" on a specified day would be an aggregation.

A social individual is the most all-inclusive type of social entity. The principle concern of the social sciences, however, is with a particular type of social individual, the social group. This is a less general concept because although all social groups are social individuals, not all social individuals are social groups. The identifying characteristics of social groups have been treated without precision in the literature of sociology; but there is an apparent agreement among many sociologists as to what these characteristics are.

"By a group itself we mean any collection of social beings who enter into distinctive social relationship* with one another" (MacIver, 1937, p. 13). For Gillin and Gillin (1953) "A group is any collection of two or more individuals who are in social interaction*; that is, who have social relations* with each other" (p. 19). As Mizruchi (1967) points out, "the typical definition of a group includes the assumption that two or more persons are in interaction" (p. 113). To define "social group" in terms of "social interaction" or "social relationship" is to define circularly; hence to leave "social interaction" and "social relationship" undefined, as is usually done, is to leave "social group" undefined. The blatancy of this circularity is reflected in the following statement from Gouldner and Gouldner (1963):

---

*Italics mine.

> A group consists of two or more people in interaction; the term <u>group</u> refers to repeated and patterned social interaction. We shall use the terms <u>social interactions</u> and <u>group</u> more or less interchangeably throughout the text... (p. 98).

Discussion of social interactions seem to involve two concepts: (1) cooperation with respect to common objectives, and (2) reciprocal communication. For example,

> A group is an identifiable, structured, continuing collectivity of social persons who enact reciprocal roles according to social norms, interests, and values in the pursuit of common goals (Fichter, 1957, p. 110).
>
> ...the sociological group involves consensus, concert, communication (Faris and Ellsworth in Coser and Rosenberg, 1957, p. 300).
>
> A group is taken to be any aggregation of two or more people who have similar interest or interests and who thus in this more or less narrowly defined aspect of their lives participate in what amounts to a common area of social interaction on common terms (Lee, 1964, p. 112).
>
> ...a group is any number of human beings in reciprocal communication (Cuber, 1959, p. 297).

First, let's consider the meaning of "common objective."

> 12.2. <u>Common Objective</u>: an outcome intended by each member of a social individual.

We must be careful to distinguish between <u>common</u> and <u>analogous</u> objectives. For example, if each member of a collection of people wants a car for himself, they have analogous objectives. If each member wants every member to have a car for himself, this is a common objective. Analogous objectives are ones which differ only in the individual(s) involved in the outcome. Objectives are not the same unless the individuals involved in them are the same. Analogous objectives are similar but not the same.

All the members of one of two competing teams have a common objective: to beat the other team. The members of the second team also have a common objective: to beat the first team. The common objectives of the two teams are analogous.

Members of a group may have more than one common objective. Groups whose members have many common objectives are sometimes referred to as "multi-purpose."

The common objective is what the interactions in a group are about. The interactions themselves are cooperative. This does not mean that each member has continuous face-to-face interaction with every other member. It only means that over time each has contact with some of the others some of the time. Nevertheless, insofar as such contact furthers the common objective, each member cooperates with all others.

At least one sociologist, Cuber (1959), does not think it necessary for members of a social group to have a common objective:

> A...popular.. fallacy pertaining to groups is the "common interest" cliché. Men are said to be found everywhere functioning in groups because they have common interests... Undeniably, <u>some</u> of man's interests are common, but others are individualized or specialized, while some are openly antagonistic...Courts, strike mediation boards, and legislative bodies are only a few of the many groups which come into existence because of conflicts among men.(p. 299).

Cuber's examples do not seem to support his point. Members of a mediation board do have a common objective: to settle the conflict to which he refers. Similarly, members of legislative bodies have a common interest: to provide adequate government. To say the members of a social group have a common objective is not to say they do not conflict with respect to other objectives or even with respect to the means by which the common objective should be pursued. If these conflicts prevent cooperation with respect to pursuit of a common objective, then the aggregation has no cohesiveness and hence is not a group.

The members of a social group can also communicate with each other, directly or indirectly (i.e., through intermediaries). Again this does not imply that each actually communicates with all others; but it does imply that the means for doing so are available to each member.

12.3. Social Group: a social individual all of whose members can communicate with each other and have a common objective with respect to which each cooperates with the others.

Social scientists find it useful to distinguish between different types of social individuals and groups. The principal variables used to differentiate between different types of social entities are (1) space-time properties (e.g., whether they are transient or permanent, and dense or sparse), (2) whether they are stimulus-oriented or response-oriented, and (3) the types of organizational structure that they have. I will not deal with all the types of social entities which even a simple dichotomous classification of these variables would yield, but only with those which have received most attention in the literature of the social sciences: organization, crowd, mob, gang, team, audience, public, family, community, and state.

ORGANIZATION

One of the most important characteristics of a social group is the extent to which its activity is organized. We also speak of the organization of a group as well as of its activity. Certain types of groups are called organizations. As we shall see the activities of organizations are organized, but not all groups whose activities are organized are called organizations. Furthermore, "The social aggregate is <u>not organized</u>; it does not have a structure with a hierarchy of positions and functions (Fichter, 1957, p. 87)."

The concept "organization" is often discussed but is seldom defined in the sociological literature. It seems to be treated as a social group whose members are either social groups or psychological individuals.

> A social organization may be defined as an integrated system of interrelated... groups formed to accomplish a stated objective (Krech *et al*, 1962, p. 384).
>
> Like other kinds of groups, the modern organization is a social system, composed of mutually influential, interdependent parts-elements and structures such as departments and individuals. None of its parts can be understood in isolation from the others (Gouldner and Gouldner, 1963, p. 396).

Such statements, however vague, provide useful insights which I shall try to exploit.

An organization is a social group and hence, contains at least two purposeful entities who have a common objective and who (actually or potentially) interact. A social group is an organization if it satisfies two additional conditions:

(1) it has a functional division of labor, and

(2) it is capable of some self-control.

These characteristics require clarification and definition.

<u>Functional Division of Labor</u>

A functional division of labor occurs in a group when a task to be done is divided into functionally dissimilar subtasks and these are assigned to different parts of the group (subgroups). Let me make this concept more precise.

(1) Pursuit of the common objective of the group can be decomposed into a finite set of functionally different subtasks $(t_1, t_2, \ldots t_n)$

each with different subobjectives, such that if these subtasks are performed (or subobjectives are obtained), the common objective will be obtained at least some of the time. No subset is ever sufficient for obtaining the common objective. Hence, each subtask in the set is necessary. There are always alternative decompositions (and, hence, alternative sets of subtasks) which are sufficient for obtaining the common objective in some environments. This accounts for the fact that different groups with analogous objectives may (and do) organize themselves differently, or that one group may reorganize itself.

To take a very simple case, suppose a car is stalled on a highway and the common objective of its occupants is to move it to the side of the road. This task can be decomposed into two subtasks, steering the car to the side of the road and pushing it, and hence two people can orgainze themselves to do the job. A baseball team divides its task of beating the opposing team into nine different subtasks: pitching, catching, and so on. A company divides its task into research and development, purchasing, production, marketing, personnel, finance, legal, and so on.

(2) The members of the social group are divided into subgroups with one or more members in each. An individual member may be part of one or more subgroups, but no two subgroups have the same composition. Every member of the social group must be a member of at least one subgroup.

(3) Each task or subobjective is assigned to one subgroup and each subgroup has at least one task or subobjective assigned to it. Assignment of a task (or subobjective) to a subgroup involves giving the subgroup "responsibility" for performing (or obtaining) it. To accept such responsibility is to accept the <u>right</u> of the group or its agent to punish members of the subgroup if they do not perform

satisfactorily.

12.4. Responsibility: one psychological or social individual (A) is responsible to another (B), if when A's behavior fails to satisfy B, B can punish A.

For example, if a member of a baseball team fails to perform satisfactorily he may be "fired" or otherwise penalized. Responsibility usually involves acceptance of the right of someone to punish. The person or group to whom responsibility is assigned may not recognize this right; for example, a criminal may not recognize the right of society to punish him, and hence he feels no responsibility to it. Society may nevertheless hold the criminal responsible and impose punishment on him whether he accepts society's right to do so or not.

A psychological or social individual can be responsible to himself or itself. This involves punishing oneself for failure to meet an expectation. Such behavior is not uncommon, although it may not be as common as rewarding oneself.

Each subgroup in an organization may organize itself to perform its subtask more effectively. There may be many layers of organization in a complex social group. In an army or a large industrial corporation it is not unusual to have as many as ten organizational layers.

We can now summarize this discussion in the following definition:

12.5. Functional Division of Labor. A social group has a functional division of labor if (1) its common objective is divided into a set of different subobjectives each of which is necessary and all of which are sufficient for the attainment of the common objective in some environment, (2) each member of the group is a member of some subgroup, (3) no two subgroups have identical membership, (4) each subgroup is responsible to the group for attaining one or more subobjectives, and (5) each subobjective is assigned to only one subgroup.

Self-Control

An organization must be capable of improving its performance when it is not satisfactory. This does not imply that it always does so. Therefore, it must be capable of, but not necessarily exercise, self-control.

12.6. Self-Control. A social group has self-control if (1) one or more of its members are conscious of its common objective(s), (2) they can observe the outcome(s) of the group's behavior and compare it with what the group intends, and (3) when the outcome is unsatisfactory they can produce changes in the behavior of the group, changes which have greater probability of producing satisfactory outcomes, than the behavior which is replaced.

Satisfaction of these conditions provides a social group with a feedback control system.

If a social group does not have a functional division of labor, it is unorganized. To the extent that it cannot exercise self-control, it is disorganized. Thus, "disorganized" implies poor organization and "unorganized" implies lack of organization. If the members of a group do not know what they are supposed to do in pursuing a common objective, the group is unorganized. If everyone knows what to do, but some required tasks are not carried out or others are not well coordinated, the group is disorganized.

Organized Activity and Organizations

Now consider the difference between an organized activity and an organization. A group of people may organize to carry out a task and as soon as it is completed, disband; or it may reorganize itself for a completely different task. For example, a group of boys may organize itself to play a game of baseball and disband after it is over, or reorganize itself for another activity. In neither case would we call such a group an organization because of the temporary character of the group and its structure. An organization is a non-temporary social group whose division of labor, although subject to change, does not change from task to task. Organizations are occupied with the repeated or continuous pursuit of common objectives. Thus a baseball team that is organized "on the spot" and shortly thereafter disbanded, is not an organization; but one which plays repeatedly against other teams in a league is an organization.

Bringing together all the preceding considerations the following definition of "organization" can be formulated:

> 12.7. Organization: a social group which (1) has a functional division of labor, (2) can exercise self-control, and (3) repeatedly or continuously pursues its common objective.

Organizational "Structure"

The way a task is decomposed and assigned to subgroups of an organization is usually called its organizational structure, but "structure" is used here in a different way than I have used it previously: as the contradictory of "function." "Organizational structure" is a functional concept.

Any task may be decomposed in a number of different ways, some more efficient than others. Measurement of the efficiency of an organization's division of labor is not a simple matter. I should like to develop such a measure for a very simple organization and by so doing indicate how it can be done for more complex organizations. (Discussion of a general measure can be found in Sengupta and Ackoff, 1965).

The objective of any organization can be described in very general terms as one of maximizing its gains (G) minus its losses (L): max (G-L). Even in the simplest organization there must be at least two controllable variables (X and Y), otherwise there would be no need or advantage to dividing its activity into parts. In pursuit of its objective the group attempts to select values of the controlled variables (X and Y) that maximize (G-L). The group's objective function, then, can be represented by

$$\max_{X,Y} (G-L). \tag{1}$$

Suppose the gain (G) is a function (f_1) of only one controlled variable (X):

$$G=f_1 (X); \tag{2}$$

and the loss L is dependent only on Y:

$$L=f_2 (Y). \tag{3}$$

Substituting the values of G and L in equations (2) and (3) in equation (1) yields the following reformulation of the group's objective function:

$$\max_{X,Y} [f_1 (X) - f_2 (Y)] \tag{4}$$

Now suppose we want to divide pursuit of this objective into two tasks. One group can be assigned control of X and the other control of Y, and their respective subobjective functions could be $\max_X [f_1 (X)]$ and $\min_Y [f_2 (Y)]$.

Then, because the gain and the loss are independent,

$$\max_{X,\,Y}\,[f_1\,(X) - f_2\,(Y)\,] = \max_{X}\,[f_1(X)] - \min_{Y}\,[f_2\,(Y)]. \qquad (5)$$

The division of labor (organizational structure) in this case has no inherent inefficiency: if each subgroup obtains its subobjective the parent group will obtain its objective.

Note that in this case the two subgroups are independent because the variable controlled by each has no effect on the performance of the other. Now consider another simple but more realistic organization. This organization has the same objective function but its gain and loss depend on both controllable variables. Hence (3) and (4) become

$$G=f_1\,(X, Y) \qquad (6)$$

and

$$L=f_2\,(X, Y), \qquad (7)$$

It may seem reasonable to assign to one group the subobjective function:

$$\max_{X}\,[f_1\,(X, Y)] \qquad (8)$$

and to the other

$$\min_{Y}\,[f_2\,(X,\ Y)] \qquad (9)$$

But for most functions (f_1 and f_2) the following inequality holds:

$$\max_{X,\,Y}\,[f_1\,(X,\ Y) - f_2\,(X, Y)] \neq \max_{X}\,[f_1\,(X,Y)] - \min_{Y}\,[f_2\,(X, Y)]. \qquad (10)$$

Therefore,

$$\max\,[f_1\,(X,\ Y) - f_2\,(X, Y)] - \{\max_{X}\,[f_2\,(X,\ Y)] - \min_{Y}\,[f_3\,(X,\ Y)]\} = K > 0. \qquad (11)$$

The difference (K) between the best that the organization can accomplish given its division of labor and the intended outcome, is a measure of the inefficiency of the organization's division of labor (structure). This measure can be generalized to cover any number of controllable variables to apply to any subobjective functions, and to take uncontrolled variables into account.

Using this same approach we can also derive measures of an organization's inefficiency due to faulty communication and poor decision making. Consider communication first. Subgroup A which controls X requires information on what value of Y subgroup B selects, and subgroup B requires information on what value of X subgroup A selects. Suppose they obtain incorrect information: subgroup A believes the value y is used by B where $y \neq Y$, and subgroup B believes the value x is used by A where $x = X$. Both use incorrect values. Then their actual performance, assuming they optimize correctly would be

$$\max_X [f_1 (X, y)] \text{ and } \min_Y [f_2 (x, Y)]$$

The difference,

$$\{ \max_X [f_1 (X, Y)] - \min_Y [f_2 (X, Y)] \} -$$

$$\{ \max_X [f_1 (X, y)] - \min_Y [f_2 (x, Y)] \}, \tag{12}$$

is a measure of inefficiency due to communication. (As we shall see below, if the organizational "structure" is inefficient, communication inefficiency can be negative.) The magnitude of this inefficiency due to communication depends on the functions f_1 and f_2, and hence on the organization's "structure." This is consistent with the widely held belief that some organizations are more sensitive to poor communication than others because of differences in their "structures." In this simple case the joint contribution of "structure" and communication to the organization's inefficiency can be measured by

$$\max_{X,Y} [f_1 (X, Y) - f_2 (X, Y)] - \{ \max_X [f_1 (X, y)] - \min_Y [f_2 (x, Y)] \} \tag{13}$$

Finally, suppose that the subgroups do not maximize and minimize their subobjective functions correctly. Let max* and min* represent such "faulty" optimization. Then

$$\{ \max_X [f_1 (X, Y)] - \min_Y [f_2 (X, Y)] \} - \{ \max_X{}^* [f_1 (X, Y)] - \min_Y{}^* [f_2 (X, Y)] \} \tag{14}$$

represents decision-making inefficiency. Note that here, too, the effect of decision-making on the system's efficiency depends on the mathematical properties of the functions f_1 and f_2 and, hence, on the system's "structure." In fact, where "structural" inefficiency exists, "faulty" decision-making and communication may be desirable. This is illustrated by the following simplified version of a real business situation which I once encountered.

Consider a retailing organization that has two subgroups: a purchasing and a sales department. The purchasing department buys a product at the beginning of each month in a quantity X which it determines. The purchased items are placed in stock until sold. The sales department sets the price (Y) at which the item is to be sold; the lower the price, the more can be sold, on the average. The amount that will be sold in any period can only be predicted subject to a known distribution of errors. This yields a "Price Demand" curve such as is shown in Figure 12.1. In this case both departments know this curve. Only items in stock can be sold; back orders are not permitted; that is, customers will not wait for the item.

Suppose the purchasing department is assigned the subobjective of minimizing the cost of inventory while at the same time providing

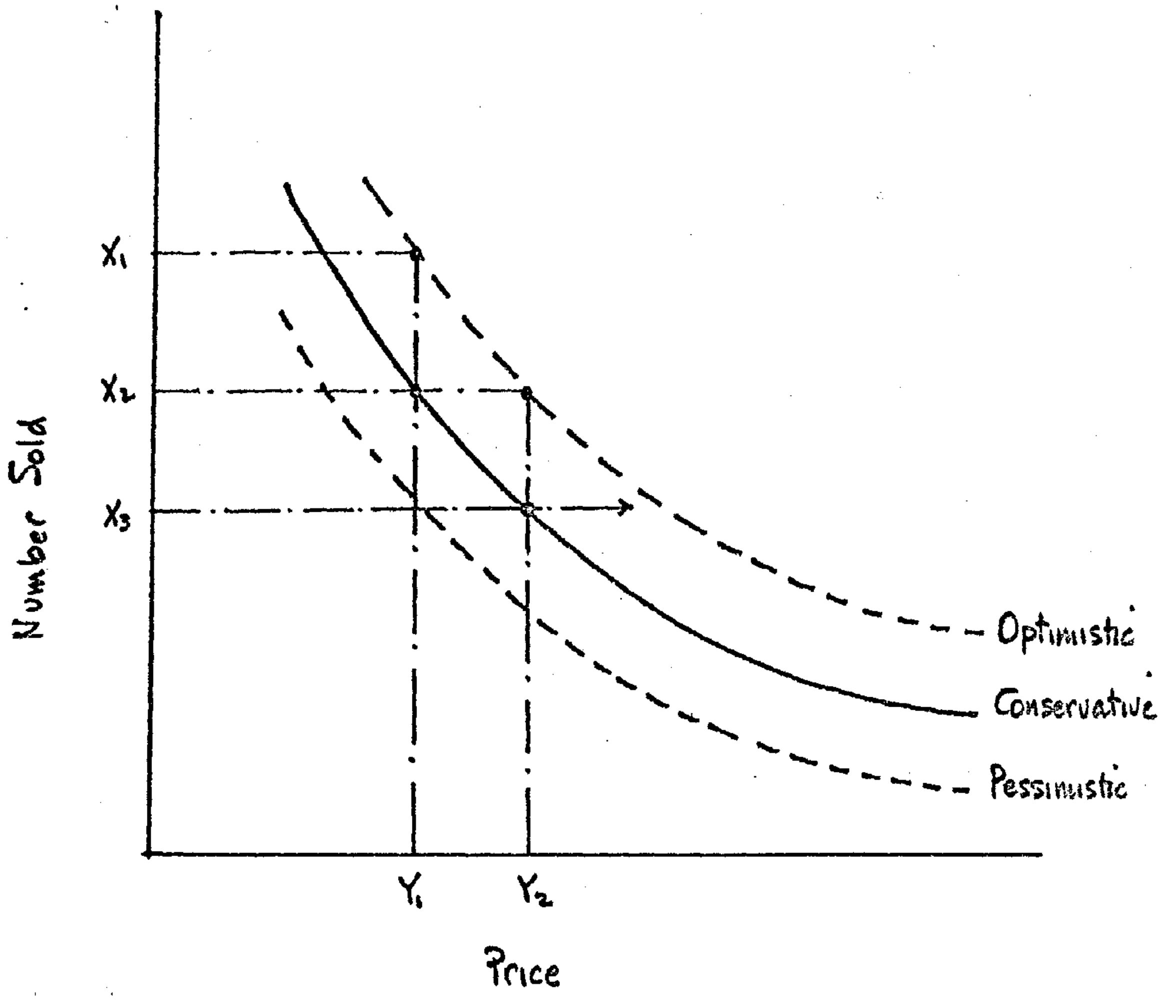


FIG. 12.1. Price-demand curve.

sufficient stock to meet its own estimate of demand. The sales department is assigned the subobjective of maximizing gross profits where this profit is equal to [number of items sold (sales price--cost of purchase)]. Now if the sales department sets a price Y_1 for the next period it forecasts that an optimistic quantity, X_1, will be sold. (See figure 12. 1.) It tends to overestimate sales and order too much because its performance suffers if it cannot meet demand, but not if items are left in stock. The purchasing department, on the other hand, makes a conservative estimate of sales (X_2) and buys only an amount sufficient to meet this forecast because, if it over-buys, the inventory carrying costs increase and the purchasing department suffers. The sales department, of course, wants the purchasing department to use an optimistic forecast of sales because its own performance suffers if orders are not fulfilled, but not if anything is left over in inventory. When the purchasing department selects an order quantity X_2 to meet a conservative forecast of sales based on the price Y_1, the sales department is informed of the fact and responds by raising its price to Y_2 for which X_2 is an optimistic forecast of demand. When it does so, the purchasing department revises X_2 to a lower value, say X_3, which corresponds to a conservative estimate of sales for price Y_2, and so on. The limit of this process is reached when the purchasing department buys nothing and, hence, nothing can be sold.

In the real situation, the limit was not reached because both departments wanted to keep the company in business,and they did so by restricting communication between them. The sales department did not tell the purchasing department what price it was going to set, and the purchasing department did not tell the sales department how many items it was going to buy. Each had to predict what the other would do. In this way, stability was obtained.

It is also apparent in this simplified example that less-than-optimization of the subobjective functions was better than optimization because the organization's "structure" is deficient.

In this case a slight change in the subobjective functions could remove the problem and the need for deliberately withholding information and not suboptimizing. If the sales department were made responsible for inventories produced by optimistic sales forecasts and the purchasing department were made responsible for lost sales, the organization's "structural" deficiency would be removed.

"Structural" inefficiency can be reduced or removed by re-organizing the group. But this is not the only way of doing so. It is possible to control subgroups, without changing their subobjectives, so that they make decisions which minimize such inefficiency. For a discussion of ways of doing so see Sengupta and Ackoff (1965).

TYPES OF SOCIAL INDIVIDUALS AND GROUPS

I turn now to a consideration of some of the principle types of social individuals and groups discussed in the literature of sociology.

The Crowd

Crowds are a type of social individual that received attention from sociologists at an early date (e.g., in the last century from Le Bon). There has been sufficient time for many definitions to accumulate, a small sample of which follows:

> The crowd we distinguish as a physically compact aggregation of human beings brought into direct, temporary, and unorganized contact with one another. It is quickly created and quickly dissolved. The units in it are not organized in relation to one another. In the crowd mere conjuncture takes the place of any definite order controlling the relation of each to each (MacIver, 1937, p. 6).

> ...a temporary aggregation of human beings at a particular spot, whether called together or responding simultaneously to like stimulus... (Eliot, 1944, p. 79).

> The crowd...may be defined as a contiguous and spatially distributed group which has a circularity of response in common language and gesture toward each other, with shoulder to shoulder massing and polarization toward some object of attention (Barnes, Becker and Becker, 1940, p. 369).

> The crowd is an ordered, relatively non-interacting aggregate of persons. From the point of view of the totality it is aimless and is not performing any common function; it is simply occupying physical space...The crowd is peacable and non-excitable; it is amorphous and exhibits only a kind of "external unity" (Fichter, 1957, p. 90).

These definitions, though not identical, overlap with respect to several properties: (1) a crowd is a temporary aggregation, (2) it is physically compact, (3) it is unorganized (not disorganized), and (4) its members respond to a common stimulus. Let us explore these properties further.

A crowd's members are potentially or actually intercommunicative. Frequently, however, the members of a crowd have no actual effect on each other. Intercommunication, when it takes place in a crowd, is frequently slight. The activities of the members are relatively independent; the members are seldom in either conflict or cooperation with each other. If, for example, one individual blocks the vision of another, interaction (though not necessarily inter-communication) can take place. The members of a crowd do not have a common objective although they usually have analogous ones. For example, each member of a crowdat a beach may have his own recreation as an objective. Because the members of a crowd do not have a common objective they do not form a social group and there is no need for them to organize their activities.

The members of a crowd are all in the same physical environment and respond to the same properties of that environment. Their responses to this stimulus need not be the same; for example, some may bathe, some play ball, others walk, and so on. The members may respond to some of each other's behavior. For example, if it begins to rain at the beach, the departure of some usually stimulates the departure of others.

The members of a crowd are densely concentrated in the same environment. The criterion of density is relative. What is required is that the number of people in the environment who respond to the common stimulus, is greater than usual.

Finally, crowds are temporary aggregations. The specification of "temporary" is also relative. Even during its short life its membership may change.

> 12.8. Crowd: a social individual whose members densely occupy the same environment for a short time and respond to the same properties of that environment.

We say of some places that they are "crowded." This does not mean that the occupants form a crowd, only that the occupation of the area is dense. For example, we say New York City is crowded, but its occupants do not form a crowd. Crowds can, of course, form in New York (e.g., at Times Square on New Year's eve). A place occupied by a crowd can always be said to be crowded.

A Mob

A crowd may and occasionally does convert into a mob. "When a crowd changes from the passive state or from one of mere interaction among its members, into a state of aggressive collective action toward some unreasoned object, it becomes a mob" (Eubank, 1942, pp. 154-155). Or again, "Mob is a crowd in motion" (MacIver, 1937, p. 191). Similarly, "A crowd in active motion in relation to a *common objective**, usually violent..." (Eliot, 1944), "A mob is essentially a crowd in positive action usually motivated by anger or joy" (Gillin and Gillin, 1943, p. 264).

---

*Italics mine.

A mob is "a congregate group of individuals who feel strongly that certain of their values are threatened and whose attitudes direct their overt behavior toward a common goal*" (Cantril, 1941, p. 80). Finally, and in more detail, Fichter (1957) wrote

> The mob is a social aggregate said to "get out of hand" because it lacks both internal and external control. It is disorderly rather than unorderly. It tends to act as a social unit on a short-lived, large-scale basis. The persons making up this aggregate are usually charged with intense emotions. The term is almost always used in a pejorative sense, indicating that the mob is destructive, antisocial, and belligerent. It is usually a protest phenomenon (pp. 90-91).

Like a crowd, a mob is a temporary aggregation. It is more mobile than a crowd and can change its location. At any moment of its duration, however, its members densely occupy the same environment.

All the members of a mob react to a common stimulus as do the members of a crowd, but the stimulus is not necessarily in the same environment as is the aggregation. For example, a mob can form in one country to protest the behavior of someone in another part of the world. A mob's members share a common dissatisfaction with whatever stimulated its formation. Therefore, a common feeling is the basis of a mob's formation. A mob responds to what its members believe (correctly or incorrectly) to be an aggressive act on them or someone with whom they sympathize. For example, mobs in foreign countries have protested against believed American aggression in Viet Nam. An anti-segregation mob takes racial equality as its objective and responds to what it believes to be white aggression on civil rights. A revolutionary mob believes its government is aggressive on them.

A mob is more active than a crowd. Unlike a crowd it attempts to change or influence others to change the producer of the stimulus

---

*Italics mine.

to which it responds. For example, it tries to stop believed aggression. A mob may organize its activity as when it divides to attack different targets. Hence, unlike a crowd it may have leadership through which it exercizes some self-control.

A mob's behavior is of a type that is disapproved by the larger social group of which its members are a part. This explains for example, why a sheriff's posse may have all the other characteristics of a mob, and yet not be considered to be one; its activity is socially approved. It is necessary to clarify "social disapproval." Mobs do not come into existence until organized societies exist. Where there is no law, regulation, and custom there are no mobs. A mob is always formed within a larger social group. This larger social group collectively has a disapproving (unfavorable) attitude toward the courses of action which a mob uses to accomplish its objective.

12.9. Mob: a temporary social group whose members may move from one environment to another but who densely occupy the same environment at any one time; it responds to a believed aggression on its members or one(s) with whom they sympathize; and it employs socially disapproved means which are intended to reduce the believed aggression.

A Gang

A gang is similar to a mob in many respects. It is mobile, aggressive, and employs socially disapproved means. It differs from a mob in that it need not be dense, that is, it may contain only a few members who may be highly organized. It does not require a believed aggressive act to produce it. Its common objective, however, involves aggression on others. The gang may merely aggress for the sake of gain for its members and not for any "revenge." A gang has a longer duration than a mob, but need not*.

12.10. Gang: an organized social group whose members cooperatively aggress on other individuals by using socially disapproved means.

A Team

A team is like a gang in that its members cooperate in conflicting with others, but it does so in a competitive situation in which it employs socially approved means. Its members cooperate in the pursuit of a common objective. Some teams (e.g., a baseball team) have highly organized activities; others (e.g., a tug-of-war team) are unorganized since there is no functional division of labor in them. The members of a competitive business organization, as well as of athletic groups, can be, and frequently are, justifiably referred to as a team.

12.11. Team: a social group which competes with another social group, using socially approved means.

Audience

An audience is similar to a crowd in several respects. The usual definitions of an audience do not generally make the differences between an audience and crowd clear, or where they do, they do not seem to be justified. For example, F. H. Allport (1924) wrote: "The audience consists of a number of individuals attending to some common object arranged, usually in rows" (p. 301). The row arrangement is clearly not essential, however, since audiences (such as those attending a concert in the park) may be arranged otherwise. Another definition emphasizes that "The audience is primarily a listening group" (Bogardus, 1941, p. 407). Etimologically the word "audience" does involve listening. But the meaning of the word has generalized to include those attending a silent film, for example, or a pantomine, or a circus, and so on. This

*"Gang" is sometimes used in an entirely different sense, as when we speak of a work gang, but this is not as the usage with which I deal here.

is reflected in the following discussion by Fichter (1957):

> The audience is a social aggregate of persons who deliberately assemble to watch and listen to a performance of some kind. We use the term here only in its strict reference to a physical collectivity within a limited spatial area. The people in an audience differ from the mob in that they are listeners and spectators rather than active performers in any joint action. They differ from the crowd in that they endure longer and their attention is more closely focused. Audiences are expected to react to a common stimulus...(p. 91).

One difference between a crowd and audience is rather apparent; an audience need not be dense, but it may be. A crowd must be dense. An audience may be sparse and small. A crowd may be small, but not sparse. But this too does seem to be the essential difference between them.

Another difference which appears to be more essential is contained in the following definition: "A crowd coming together for a specific purpose, to be instructed or entertained" (Britt, 1941). But this will not do because an audience frequently does not have this property. For example, at a party it is discovered that someone present is an accomplished pianist, and this person is induced to play. The gathering turns into an audience as he performs though its members had not come to the party for the purpose of hearing the pianist. Furthermore consider the case of firemen fighting a fire. A crowd forms at the sound of the sirens for the specific purpose of watching what is going on; but we do not call this assemblage an audience, but a crowd.

Considerations such as these indicate the fuzziness surrounding the term, "audience." Since the development of the radio and television, the term has been even further confused, for we speak of a radio or television audience, where the individuals are not even in the same environment. Such an audience is not a social group, for the listeners may not be intercommunicating or even potentially intercommunicating.

If they are, it would be more precise to call them a public, which concept we shall presently consider.

A suitable basis for differenting between a crowd and an audience seems to lie in the difference between the following two situations. (1) An accident occurs on a busy corner, a crowd gathers; we would not call it an audience. (2) A street vender on a similar corner begins to talk about his product and a group forms (we are tempted to say a crowd forms). But we refer to this group as the vender's audience; we do not refer to the group forming around the accident as the drivers' audience. In the case of a crowd the stimulus of its function was not intended to attract attention, but in the case of an audience it was. We can speak of an audience at a prize fight, but not of an audience at a street brawl. Nevertheless we also refer to the attendants at the prize fight as a crowd, as we do to any audience if it is dense. This is evidence of the looseness with which we use "crowd" to refer to any dense collection of individuals.

It is not sufficient, however, merely to say of an audience that it gathers to respond to a stimulus intended by its producer to be responded to. If it were, a group of people riding in a bus would have to be said to constitute an audience, since they are all responding to a stimulus, the driving of the bus, to which the driver intends to have them respond. Britt's definition (given above) suggests, however, that audiences gather to be entertained or instructed. Generalizing on this suggestion we realize that members of an audience gather to have their functional, not structural, properties changed. The bus driver merely intends to change the physical properties of his passengers. If the producer of the stimulus intends to change the functional properties of the responding invididuals and they intend to be so stimulated then they would constitute an audience. For example, those attending a fashion show are an audience having their familiarity patterns changed. Those attending a college lecture are presumably being informed, instructed, or motivated. Those listening to a sermon are having their intentions modified. Those who watch a motion picture are being relaxed or excited. These would all constitute audiences.

> 12.12. Audience: a social individual all of whose members intentionally respond to the same stimulus which is produced with the intention of stimulating them.

A more restrictive definition would require that all members of an audience occupy the same environment.

An audience leaving a theatre is no longer an audience, but it may be a crowd. If it engages in an aggressive protest against the performance, as some audiences have, it becomes a mob.

Public

> The public...is...a mental construct in which persons are thought of as a social unit because they posess certain common characteristics...
>
> In a technical, scientific sense, a public does not refer to the total general population, nor does it refer to an organized social group, although both of these meanings are sometimes erroneously applied to the term. A public differs from an aggregate because the latter is marked by physical proximity and the former is not (Fichter, 1957, p. 74).

The members of a public do not seem to me to be quite as disconnected as Fichter takes them to be. They, like the members of an audience seem to respond to the same stimuli. The stimuli to which a public responds are not necessarily produced for the purpose of evoking their response as in the case of an audience, nor do the respondents expose themselves to the stimuli for the primary purpose of being stimulated.

When the President of a nation speaks over radio or television, all those who listen to the broadcast or rebroadcasts are part of his audience, but all are not necessarily part of his public. If his broadcast is carried abroad, his foreign audience is not part of his public. On the other hand, much of his public is not likely to be in his audience. The public consists of individuals to whom the President is responsible; he is supposed to be serving their interests. They can respond by supporting (cooperating with) him, or opposing (conflicting with) his actions, or being indifferent. This is why public response matters.

An audience may respond to a performance which involves no communication; for example, acrobats and other circus performers. Not so for a public; it responds to communications about what a person or group does.

> 12.13. Public: a social individual whose members are dispersed over many environments and respond to communications about the behavior of a (psychological or social) individual which affects their expected values, an individual who is responsible to them.

The members of a public are always members of the same community and the stimulus to which they respond affects their community, to the nature of which I now turn.

The Community

> ...a community...[is] a group of people who occupy a common land area within which they perform their major life activities (Gouldner and Gouldner, 1963, p. 421).
>
> The term "community" is another of those sociological words which has come to have a wide variety of meanings. It is sometimes used interchangeably with words like "society," "city," "neighborhood," and even in expressions like the "Catholic community" or the "Negro community," to designate loose social categories in the larger cities... A community is a territorial group of people in reciprocal relations, using common means in the pursuit of common goals (Fichter, 1957, p. 141).

It is apparent that Fichter's definition of a "community" is much more specific and restricted than is that of Gouldner and Gouldner.

What the latter has called a community, the former would call a neighborhood.

Most definitions of "community" refer to only two properties, geographical proximity of members, and common interests. But such definitions do not serve to distinguish a bridge club from a city. Some social scientists (e.g., Osburn and Newmeyer, 1933) attempted to correct for this by relating the community to "the chief concerns of life" or "basic needs." But these concepts are themselves left vague. Nevertheless, they are suggestive.

Let us consider what might be called "primitive communities." Such communities consist of a collection of interacting individuals whose common objective it is to provide and maintain instruments (including facilities and services) for the satisfaction of its members' analogous objectives (e.g., self-preservation). Each of the members have access to these instruments in satisfying these ends. Further, each member of the social group is responsible to the group for the cooperative provision and maintenance of these instruments.

An example of such a community would be a group of pioneers who combine efforts in tilling the soil, caring for crops, and protecting themselves against hostile men and animals. If a member of such a community does not do his "share," he may be deprived of his allotment of food or protection.

This is a very simple type of community. For one thing we have not considered the very young, the very old, or the infirm, the incarcerated, and others who cannot or are not permitted to contribute to producing instruments but who may, nevertheless, be members of a community. For another thing, the members of a community may only produce these instruments in an indirect sort of way. For example, they may merely contribute to providing means (e.g., money) for purchasing facilities and instruments from other communities. Therefore, it is necessary to generalize the meaning of a community to take care of such considerations.

The young, the old, the infirm, and so on may neither produce or maintain nor contribute directly or indirectly to the production or maintenance of the shared instruments and facilities and yet have access to them. A community accepts responsibility for providing such access to some who cannot contribute to making these instruments available but who live in the region occupied by the community. All who have access to the communal facilities and services are responsible for using them in such a way as not to deprive others of access to them.

The members of the community need not actively cooperate in producing these instruments as where each works the land. They may merely pay taxes which are used to pay for labor which produces streets, markets, water supply, and so on. But even where they contribute indirectly, tax payers are co-producers of these communal instruments and their maintenance.

The responsibility of the community to its members and its members to it is the basis of whatever measure of autonomy it has. It may not be (and frequently) is not completely autonomous. Its members may be responsible to another social group which contains the community as a part (e.g., as a state contains a city). But the group as a whole is responsible only to its members or those within its boundaries. A visitor may be required to contribute to the production and maintenance of public instruments through taxes (e.g., sales tax), but his responsibility for doing so is usually limited.

We sometimes speak of one social group as having more community spirit than another. This indicates that we employ, however unconsciously, measures of communality. These measures can throw light on the nature of a community.

(1) One aspect of the degree of communality is the range of analogous objectives for the satisfaction of which instruments are cooperatively produced. A community which is organized only to provide public roads, is much less communal than one which also provides shelter, purified water, sewage disposal, public transportation, schools, parks, police protection, and so on.

(2) Another aspect of the members' objectives which is important in measuring the degree of communality, is the average intensity of interest in the objectives for which the instruments are produced. For example, if we have two communities, both producing facilities usable in the pursuit of only one objective, but one provides food and the other flowers for decorative purposes, we would say the first is more communal since the desire for food is generally more intense than is the desire for decorative surroundings.

(3) The amount of cooperation that takes place among the members in the production of the instruments is another important measure of communality. The more cooperation among its members the more communal is the group. It is this measure that is the usual basis of attribution of community spirit.

(4) Finally, the amount of cooperation in the use of the public facilities and services is important as a measure of communality. If all the members cannot equally share the public instruments, facilities, and services--that is, if their availability is preferential--the communality is not as high as it would be if access were equal. Because access to these facilities and services is an essential part of membership in a community, its members must occupy an area from any part of which there is access to these instruments, facilities, and services.

12.14. Community: a social group all of whose members occupy a bounded area within which the group provides them with, or provides them with access to, instruments for the satisfaction of some of their analogous objectives, instruments which some of its members are responsible either for producing and maintaining or for providing the group with the means for acquiring and maintaining them, and which all of its members are responsible for using in a way which does not reduce the access of any others in the group to them.

The State

The state is a social group and a special kind of community. By state, I do not refer to such political units as New York, Pennsylvania, and so on, but rather to autonomous social organizations (e.g., a nation). The state may exist in the form of a primitive tribal community, an urban or rural community (e.g., the city-states of the ancient Greeks or in such "free cities" as was Danzig), or a national community or empire. The essential characteristic of the state is its autonomy. It is necessary, therefore, to define this notion of autonomy; but this is not difficult since the concepts necessary to do so have already been developed.

In a community, as I have defined it, each member has certain responsibilities to it. But the community may itself be part of a larger social group so that its members also have responsibilities to the larger group. This is obvious enough in the United States, where a citizen has obligations to city, state (in the other sense), and nation. Each of these represents a community, with the United States a "national community." But the United States has a property which its member communities do not have: its members have no responsibility with respect to public facilities, and services to any community which is not contained within it. In this sense the United States is currently an autonomous community, though it was not prior to its revolution in 1776.

12.15. Autonomy. A social individual is autonomous if it has no responsibility to any psychological or social individual that is not part of it.

12.16. State: an autonomous community.

Historically the state has been developing into a larger and larger community. The trend toward larger and larger autonomous communities is unmistakable. History has seen autonomy move from clan, to tribe, to village, to city, to nation, and now is watching the struggles toward internationalism. In a sense, empires have already achieved one type of internationalism. But such efforts as that of the League of Nations and the United Nations are moving toward establishing one all inclusive state, one autonomous community. At present the members of the United Nations are still autonomous units, but if and when the member nations passes more control over to it, a new autonomy will begin to arise in much the same way as it did in this country when the original thirteen states united and sacrificed their autonomy to form this nation.

Society

"Society" is as fundamental a concept in sociology as is "social group" but its meaning is even less clear. For example, "Society is a group of human beings cooperating in the pursuit of several of their major interests invariably including self-maintenance and self-perpetuation" (Fairchild, 1944, p. 300). This definition is equally applicable to the family. The same is true for the definition of society as "a group of people who have lived together long enough to become organized and to consider themselves and be considered as a unit. more or less distinct from other human units" (Cuber, 1959, p. 68).

A different emphasis can be found in the two following definitions:

> A society...represents the largest grouping in which common customs, traditions, attitudes, and feelings of unity are operative (Gillin and Gillin, 1943, p. 19).
>
> A society refers to the broadest grouping of people who have a certain common set of habits, ideas, and attitudes, that is, a social and cultural content, living in a definite territory, and often set off from other societies by attitudes and actions of indifference or antagonism (Young, 1942, p. 19).
>
> ...a society...is an organized collectivity of <u>interacting</u> people whose activities become centered around a set of <u>common goals</u>, and who tend to share <u>common beliefs</u>, <u>attitudes</u>, and <u>modes of action</u> (Krech *et al*, 1962, p. 308).
>
> A society is an organized collectivity of people, living together in a common territory, co-operating in groups to satisfy their basic social needs, subscribing to a common culture, and functioning as a distinct social unit (Fichter, 1957, p. 135).

Note the emphasis on <u>largeness</u>, <u>organization</u>, and <u>common culture.</u> What is not made clear in these definitions is how a state and a society differ. The difference is noted but not clarified by Fichter (1957) as follows:

> The society exists within a common geographical area. In the highly organized modern world, this usually means that certain physical limits fix the boundaries of a nation in which a complete society exists. It is possible, however, that separate societies exist within a nation so that the word "nation" is not synonymous with "society" (p. 134).

There has been an increasing tendency in history for "society" and "state" to be treated as synonyms. The development of autonomy has more and more paralleled the development of common traditions, attitudes, and so on. But, as Fichter noted, the state and society are not completely synonymous. For example, the British Empire contained many different societies: the Indian, South African, New Guinean, and so on. Even the United States contains different societies within its own geographical boundaries; for example, some Indian societies are not yet "integrated" into the dominant Anglo-Christian society which characterizes most of it. In ancient Greece individual cities and their satelite communities constituted states, and yet there was a Greek society of which these states were parts. Greece and Rome differed in this respect: Greece was a society containing many states, and the Roman Empire was a state containing many societies.

Societies, it seems to me, are not so much based on common objectives, as they are on similarity of means and instruments used by most of its members in pursuit of similar objectives. Not only do we find different societies with similar objectives, but in any one society we find many conflicting objectives pursued by its members. The means and instruments commonly used by most members of a society are usually taken to be part of its "culture;" and "common culture," as we have seen is often used to define "society." There are similarities and differences between all cultures, and nowhere has the critical amount or kind of similarity and difference been set down. For example, some would maintain that there is one Anglo-American society, others that there are two societies, Anglo and American, although this was not so before the end of the eighteenth century. The American and British are alike and different in many respects. And until we make precise what are the critical similarities and differences in terms of which societies are to be individuated and identified, argumentation on such an issue is academic in the worst sense.

The definition of "society" may not lie in developing a set of critical standards in terms of which social groups are joined into one society or separated into two. For some purposes it seems fruitful to consider all the peoples of Europe as constituting one society, for others to break Europe into national or ethnic societies, and on other occasions it is useful to make even finer distinctions. In effect, the concept of "largest social group" as it appears in the definition of society, does not seem to involve size so much as the investigators purposes: it is the largest social group he is willing to consider as an individual in the particular investigation he is conducting. Relative to one particular investigation certain cultural differences may be important, which in another investigation may be incidental. Then in this sense, even a social group which is relatively small in size may be considered as a society. For example, a small community which is built around some unique economic function which it shares with no other community in the area (say whaling) develops cultural patterns which distinguish it from other communities to some of which it may be bound by many other similarities. Nevertheless, for some investigations, it may be the differences rather than the similarities which are of importance, and in this case the community would be taken to be a societal entity.

As I see it, society is not a type of social group; the properties which define it are not properties of the group but of the researcher. It is the largest social group on which a researcher concentrates his attention. Therefore, any social group might, under certain conditions, be considered to be a society. Using "society" in this sense makes it clear why sociology is so often called the "science of society."

GROUP BEHAVIOR AND SOCIAL SCIENCE

At the beginning of this chapter the similarity of the relationship between (1) mechanics and physics, and (2) psychology and sociology was pointed out. But these pairs differ in an important way. The mechanics-physics relationship is affected by the fact that it is easier to observe physical bodies and their properties than point-particles and their properties. In the psychology-sociology relationship, however, it is the psychologist who appears to have the simpler observational task. As a consequence most social scientists, it seems to me, do not observe group behavior, but the behavior of individuals in groups. In fact, a great deal of sociology has nothing to do with group behavior, only with the behavior of individuals in groups. Whereas psychologists are primarily occupied with the uniqueness of each individual's behavior, many sociologists, equally occupied with individual behavior, are preoccupied with similarities of behavior of different members of a group. Therefore, the objects observed are the same in psychology and much of sociology, but the properties of concern differ; one seeks to describe and explain differences and the other similarities.

The fact that many sociologists are concerned with individual behavior and similarities of different individuals' behavior is apparent in the way they deal with central concepts of sociology.. Almost any standard text in sociology could be used to support this observation. I use that by Gillin and Gillin (1943) as illustrative.

Gillin and Gillin define "culture" as "the learned reactions in common practice by members of a social group" (p. 127). Note the emphasis on individual behavior. This emphasis is quite self-conscious, for the authors write later, "The only form in which the culture of a group is available for scientific study is in the learned, common behavior of the individuals who compose the group" (p. 27). This same orientation is reflected in their treatment of various important aspects of culture. For example,

> Folkways are behavior patterns of everyday life, which generally arise unconsciously in a group, such as tipping the hat, calling on strangers, and shaking hands, and without planned or rational thought (p. 134).
>
> Mores, on the other hand, are those customs and group routines which are thought by the members of the society to be necessary to the group's continued existence. These customs are "right". Under this head come such customs as religious rituals, respect for authority, marriage, sex tabus, and so on (pp. 134-135).

Although the latter definition hints at group behavior the observations that are taken as relevant to the study of mores are still observations of recurrent patterns of individual behavior.

This individualistic treatment of "culture" is not unique to Gillin and Gillin; for example, Cuber (1959) wrote as follows:

> Culture is the continually changing patterns of behavior and the products of learned behavior (including attitudes, values, knowledge, and material objects) which are shared by and transmitted among the members of society (p. 60).

The study of common patterns of behavior among members of a group is, of course, a legitimate scientific enterprise, but it is not the study of group behavior. It seems to me that the historian has been more sociologically oriented than the sociologist. He deals with the behavior of nations, alliances, political parties, revolutionary movements and so on. The economist and management scientist deal with companies, industries, and even nations as entities. To study these entities they do not observe the behavior of their individual members, but the behavior of the groups taken as a whole, as an entity. For example, the pricing policy of a company, its growth in sales volume, profits, acquisitions, diversifications, and so on are group behavior. To be sure, knowledge of them could be obtained by observing individual members of the firm, but this is not the way it is done. A contract between labor and management is the product of negotiation between individuals, but is also a group product. The question involved is which of the two ways of looking at groups, atomistically or holistically, we should use.

I am in no position to say which way of looking at group behavior is the better. I don't think this is an issue because experience in science has indicated that phenomena should be studied from as many different point of views as possible. Different points of view interact and enrich each other.

My objective, therefore, is not to preclude the current type of study of social groups, but to facilitate a science of social groups which studies groups as entities and not as collections of psychological individuals; and I press this point because I believe it will yield new insights and understanding of collective behavior.

I have already observed at the beginning of this chapter that a property of a social individual (and hence of a social group) can be expressed as a function of the properties of the individuals who make it up. I also noted that a property of a group need not be determined by observing the individual members. I used the analogy of observing temperature in physics. The profitability of a firm is a similar property; it is much easier to observe at the collective level than to try to build it up out of properties of individual behavior. We can characterize certain communities as being aggressive or submissive without observing any of its individual members; in fact we do this daily when reading the newspaper.

Given that we can observe the properties of a social group without observing the properties of its members, it follows that we can also observe its behavior since behavior is simply a change of properties over time. We can and do speak of a nation selecting war as a means to an end, or negotiation. Since we can talk meaningfully of a group's behavior, we can discuss the outcomes that it does and does not produce; that is, its function. Hence we can also consider the choices of a group and characterize them by use of the same conceptual scheme that I have developed for study of an individual's choice. A re-examination of the conceptual system developed in this book will reveal that it has no properties which restrict its application to persons (i. e., to psychological individuals). It is equally applicable to groups (i. e., to social individuals).

Therefore, it is meaningful to talk of a gouup's personality, but I would prefer to call this its culture; for I believe culture is to a group what personality is to an individual; its general choice function. We can also speak of the familiarity, knowledge, understanding, intention, vacillation, traits, attitudes, observations, beliefs, and memory of groups, as well as the many other properties discussed in these pages. In many cases we may prefer to call these properties by other names when they refer to groups, but in trying to find other names we run into the fact that the social sciences have hardly dealt with such properties of groups, at least not in a scientific way.

One might argue against the point of view that I have taken here as follows: Even if we can observe group properties and behavior holistically and use the concepts you suggest, we cannot begin to approximate the kind of controlled environmental conditions called for by your definitions when the object being observed is a social group, particularly a large one. The task of inferring from the environments in which we can observe these groups to the idealized ones involved in your definitions is very complex and, indeed, probably impossible now and for a long time to come, if not forever.

Such an argument has a great deal to it. In the next section I describe a methodology for social research which I believe can overcome these difficulties. However, it is not the only methodology which can do so.

A METHODOLOGY FOR SOCIAL RESEARCH

In attempting to develop understanding through research of social phenomena, particularly ones involving large social groups, the investigator is confronted with what initially appears to be an almost hopeless task. For example, each instance of large-scale social conflict--a war, a strike, or a riot--appears to be infinitely complex, unique, and characterizable only by intangibles. Dealing with such problems, however, is not new to science whose progress can be measured largely by the extent to which it has converted what initially appeared to be hopelessly complex into what eventually appeared to be relatively simple. Simplicity comes at the end, not at the beginning of research. At one time heat and electricity seemed to be as difficult to understand as large-scale social conflicts seem to be today.

At the beginning of scientific inquiry into a new area every theory that is proposed, no matter how complex, seems too simple. Once science has achieved some measure of success in an area, however, every theory, no matter how simple, seems too complex. As understanding of a class of phenomena increases, the number of variables required to explain it decreases, and the explanation of their interactions and effects becomes increasingly "obvious."

The principle method by which science has explored the unknown is experimentation. It is not possible, however, to conduct experiments on large-scale social systems. For example, we cannot bring social conflicts into the laboratory, nor experiment on them in their natural environment, nor do we have the right or capability of intervening in them; we cannot run the risk of intensifying them by experimental manipulation. Furthermore, we cannot perform quantitative analyses on past conflicts, because histories and descriptions of these conflicts have not been recorded reliably or in a quantitative form. Records of past conflicts do not provide us with sufficient "facts" to allow us to find in them dynamic regularities or consistent causal principles.

In a sense the researcher into the operations of many social systems is in a situation similar to that of the early astronomers; the system they studied also seemed to be infinitely complex and yet incapable of being subjected to experimentation. Astronomers, however, eventually developed mathematical representations (models) of the systems and analyzed, or conducted experiments on these models. Today such experiments are called "simulations."

In order to proceed as the astronomer has it is necessary to have ordered, accurate, quantitative descriptions of the behavior of the system under study. Newton's work depended upon Kepler's and Kepler's upon Brahe's. Without Brahe's detailed and fastidious accumulation of relevant facts, Keplerian laws and the Newtonian theory could not have been developed. The corresponding type of quantitative descriptions of large-scale social phenomena which are required before theoretical work can be begun is not available. For example, there are few impersonal and objective descriptions of past or current conflicts because different observers seldom record contrary "facts," and analysts seldom draw the same conclusion from even the same set of "facts." Therefore, one might first attempt to understand the dynamics of large-scale social conflict by seeking accurate descriptions of real conflicts. But even today this is very difficult, if not impossible. However, there is an alternative method recently developed for just such situations in which the problem of preparing quantitative descriptions of real large-scale social phenomena has a secondary role. If this method succeeds, it will provide the criteria of relevance and techniques of data-evaluation that are required before accurate and reliable descriptions of complex social phenomena are possible. I continue to use the example of large-scale social conflict in developing the characteristics of this method.

Conflict, like many other social phenomena, has been studied extensively. In previous research three approaches to the problem have been taken. The first, [exemplified by Anatol Rapoport's simple Prisoner's Dilemma conflict games (1965),] involved two-person groups in laboratory situations. Rapoport has developed a mathematical model which explains this particular conflict game. But he recognizes that inferences cannot legitimately be drawn from highly controlled but simple conflict situations to the very complex uncontrolled ones found in reality. His work simply provides what he calls "insights" into real large-scale social conflicts (Rapoport, 1960 and 1965). This is not a meagre accomplishment, but until we can learn how to infer from conflict situations that can be studied in the laboratory to real situations, it is not likely that we will develop a scientific theory that applies directly to the dynamics of real large-scale social conflicts.

The second approach to the study of large-scale social conflicts involves the use of relatively complex experimental situations; for example, international political games. Examples of this approach are in the work of Harold Guetzkow (1963) and Bloomfield (1965).

Although the gap between these games and reality appears to be, and may be, smaller than in simple two-person games, the inferential problem remains for two reasons. First, these games resemble reality because they reproduce many of its properties, but there is no assurance that these properties are related to each other in the games as they are in reality. Therefore, inferences cannot legitimately be drawn from games whose structure is not known, to a reality whose structure is not known. Second, because of the complexity of these games, precise quantitative description of what happens in most of them has not been possible. Again, such comments do not minimize the value of the insights these games have provided.

The third type of approach involves analysis of real conflict situations by either (1) traditional historical analysis; (2) new techniques of analyzing communications between conflicting parties; or (3) statistical analysis of political, social, and economic variables. Examples of rigorous efforts using this approach include the work done at the Foreign Policy Research Institute of the University of Pennsylvania and the work of Yale's Dimensionality of Nations Project.

The low degree of relevance and reliability of available data can make analysis of real situations difficult. At best the types of statistical analyses involved in this third approach yield descriptions, not explanations, of what has taken place. Hence, even if completely successful, they can only yield accurate predictions, not control, of what will take place. Most of the conclusions reached by any analysis of reality have not been reproducible in any objective way. In brief, such analyses have not yet produced a body of knowledge that can be called scientific. The results obtained are often vague and frequently inconsistent, and seldom justify a status higher than that of a conjecture.

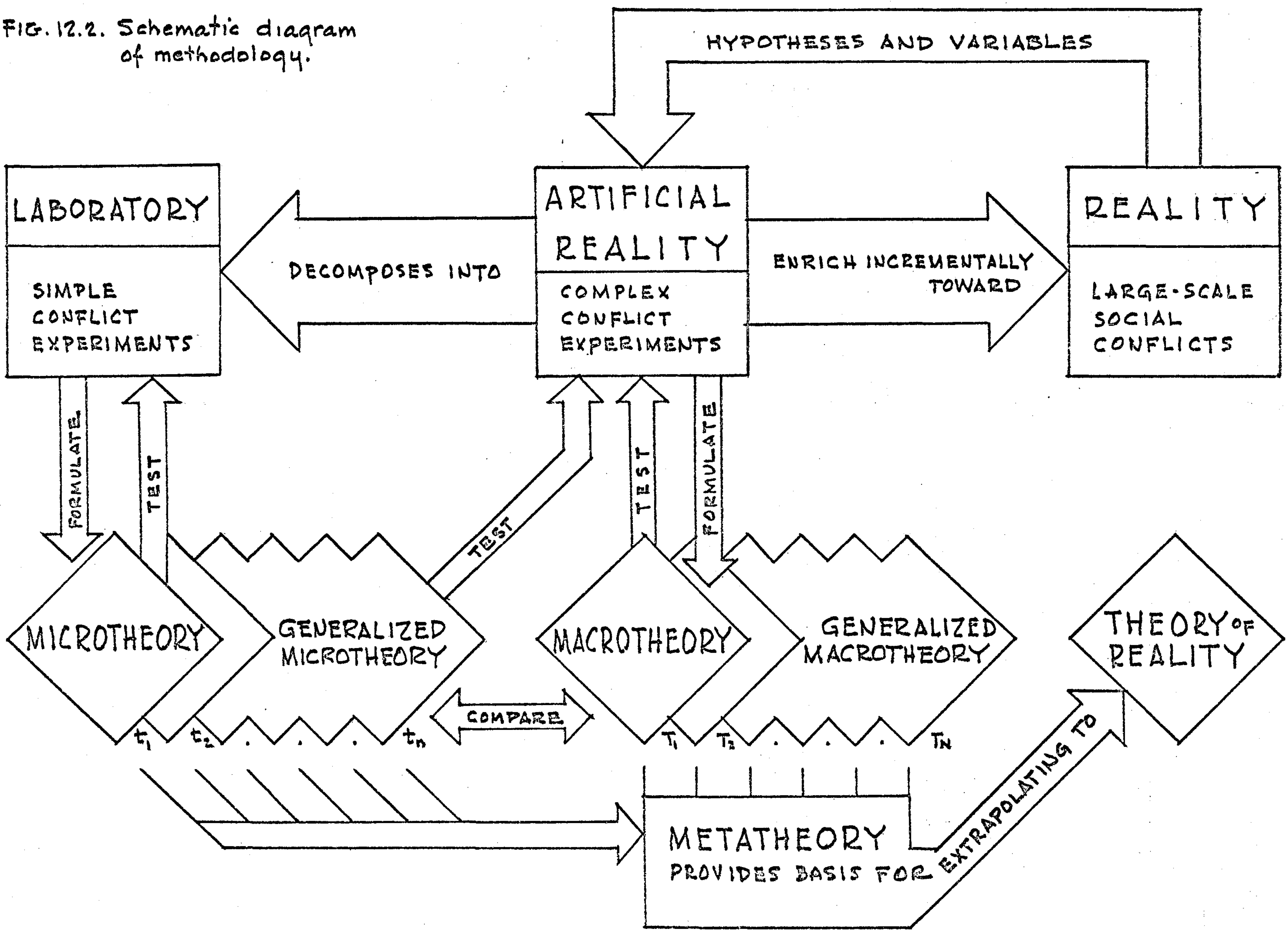
In the method that is proposed here an effort is made to incorporate the strengths of each of these three approaches and to avoid or minimize their difficulties. It tries to retain the value of both controlled experimentation and rigorous quantitative analysis contained in the first type of approach and also the realism of the second and third approach. If anything other than real social phenomena is to be studied, however, the principal methodological problem that must be overcome is that of infering from a situation that substitutes for reality to reality itself.

The method that I propose is shown schematically in Figure 12.2, again using research into large scale social conflict illustratively. *

First, the literature relevant to the real situation under study is reviewed and all hypotheses and conjectures concerning the phenomenon in question are extracted from it. Since some of these statements will overlap, the resulting list is edited and condensed. (In the case of large-scale social conflict well over a hundred such statements were yielded by this process.)

*A complete account of the conflict research that is making use of this method can be found in Management Science Center (1957).

FIG. 12.2. Schematic diagram of methodology.



Next the variables involved in each of these hypotheses and conjectures are extracted and listed. This list is also edited. The variables in the final list are then ordered in terms of the frequency with which they appear in the hypotheses and conjectures. (For example, in the conflict case, "communication" was the most frequently cited variable.) This ordering provides an initial priority that can be used in the experimental work described below.

As will become apparent in a moment, these variables must eventually be given operational definitions so that they can be used experimentally. The conceptual system constructed in this book may provide some of the required definitions, and others can be derived from those that are provided here.

Now, a relatively complex experimental situation is constructed, one that I call an "artificial reality" (or "rich game"). It should be as simple a situation as possible and yet satisfy the following conditions:

(1) It is "rich" enough to test a large number of hypotheses that have been formulated about whatever type of phenomenon is relevant (e.g., the dynamics of large-scale social conflict). (Clearly, such tests cannot confirm any hypotheses about reality, but they can limit the generality of hypotheses or show how they can or should be generalized.) The purpose behind this condition is to assure use of an experimental situation that is realistic enough so that most assertions made about the real situation are applicable to it.

(2) There must be explicit operational definitions of the variables manipulated in the situation including the scales used in measuring them, and of the variables by which simplification of reality has taken place (e.g., by holding a variable constant). Identification of these factors makes it possible to design successively enriched experimental situations by the addition of complexities, one at a time or in controlled combination.

(3) The relevant behavior in the experimental situation must be describable in quantitative terms.

(4) The situation must be decomposable into a set of simpler experimental situations and where possible, these simpler situations should be ones which have already been experimented on, or closely resemble situations which have been researched. This enables one to relate the results obtained here to previous work.

The experimental situation which satisfies these conditions is not used as a model of reality, but rather as a "reality" to be modelled; hence, its name, "artificial reality." It is used to generate a "history" which is to be explained by the first "macrotheory" to be constructed. The history is generated by experimentation (e.g., by playing the rich game under laboratory conditions) which is designed to test hypotheses about real conflict that have been translated into operational and quantitative terms and adapted to this artificial world.

Experiments are also conducted using decomposed parts of the artificial reality; that is, using simpler conflict situations. These "laboratory" experiments are used to construct a "microtheory" to explain their results. A generalized microtheory which explains a variety of simple conflict situations is then sought; the essential differentiating characteristics of a variety of simple situations enter the theory as variables. Finally, generalization of the microtheory to the artificial reality is attempted. Such a generalization is called "macrotheory" in this context. *

A simultaneous effort is made to formulate a "macrotheory" of the "artificial reality" by direct analysis of the history which it generates. These two theoretical efforts interact until a satisfactory macrotheory (T_1) of the artificial reality is developed.

*Howard's Meta-Game Theory that was discussed in Chapter 11, and Emshoff's Microtheory that will be discussed in Appendix III were developed out of this phase of the research on large-scale social conflict.

Once a satisfactory macrotheory (T_1) is developed, the initial "artificial reality" (R_1) can be modified to provide a more realistic conflict situation (R_2) e.g., by converting something that was held constant in R_1 into a variable. Efforts can then by made to generalize the earlier theory, T_1, so that it applies to R_2. The output is a more general macrotheory T_2 of which T_1 is a special case. T_2 is tested against "history" generated by experimentation with R_2. This procedure is continued, hopefully producing a sequence of successively more general macrotheories, $T_1, T_2, \ldots, T_n$.

As this set of theories expands, it can be analyzed to find principles which explain how the theories must be generalized in order to apply to more realistic "artificial realities." That is, a metatheory is sought. The metatheory yields a procedure for generating T_{n+1} given $T_1, T_2, \ldots, T_n$, which can be tested in R_{n+1} which is a modification of R_n. The development of such a metatheory should eventually make it possible to take larger jumps toward theories of real conflicts; hopefully to a theory that applies to reality in all its complexity.

The plan encompassed in this methodology cannot be carried out in a short period of time. Its complete realization even for one type of phenomenon will take many years. The time required depends on the amount of research effort devoted to its realization. The methodology provides a framework for organizing and integrating the efforts of a large number of research units.

CONCLUSION

This chapter completes the conceptual system which I have set out to construct. This is not to say that all the relevant concepts in the study of human behavior have been covered. Obviously this is not the case. However, I hope that enough has been included in this effort to provide others with a basis for defining into this system any other behavioral concepts of importance to them in their research. In a sense, then, what I have tried to do is provide a framework on which the edifice of the behavioral sciences can be hung. It will undoubtedly be necessary to change many of the definitions that I have formulated. I am not prepared to defend any particular definition in the same way that I am prepared to defend the method of approaching the definitional problem which I have used. One of the principal purposes of this method is to provide a basis for progressive improvement of definitions of behavioral concepts. Therefore, I expect to revise a number, if not most, of the definitions offered here over time, and I hope others will join me in doing so, as well as in adding concepts on to this framework.

Since I started this effort with the concepts of geometry, kinematics, and mechanics and worked my way into behavioral concepts, I may seem to have supported a hierarchical notion of science in which it is maintained that some concepts are more fundamental than others. I do <u>not</u> support such a concept of science and have written to this effect in other places. (See, for example, Churchman and Ackoff, 1950.) I believe strongly in the complete interdependence of concepts in science and maintain that the apparent logical hierarchy of concepts is, in fact, a historical (not a logical) ordering.

To prove this point C. West Churchman and I once began with behavioral concepts much like those developed in this effort and derived the concepts of logic, geometry, arithmetic, kinematics, and mechanics from them. These structural concepts were defined in terms of the functions they performed for the purposeful entities who created them. For example, the "basic" concepts of "time" and "place" were shown to derive from the need to individuate objects and events which are alike in all other respects that are relevant to the observer (e.g., identical twins, or two copies of the same book). Such definitions of structural concepts were functional in nature; the concepts were defined in terms of what they do for people not in terms of what they represent. Such definitions state what these concepts mean in the pragmatic sense of "meaning" developed in Chapter 9.

Although it is a great temptation to "round the circle" here and show how structural concepts can be derived from ones that are functional, I have resisted because it is not essential for my purposes here.

Up to this point I have only suggested how the conceptual system which I have developed can be used in the research process. Whatever value this system has can only be proven in such a process. Productivity in research, not polemic, will establish its value or lack of it. In a second volume to follow this one, three types of application will be described in some detail. The first will deal with the development of psychological tests for such personality properties as have been defined here. The second will deal with measurement and experimentation on the communication process. The third will deal with experimentation and theory construction in the realm of behavior in conflict situations.

I am very grateful to those readers who have reached this point without turning to it first, and hope that the hard trail over which I have tried to guide them has provided them with some reward for their efforts.

BIBLIOGRAPHY

Allport, F. H., Social Psychology, Houghton Mifflin Company, Boston, 1924.

Barnes, H. E., Becker, H., and Becker, F. B., Contemporary Social Theory. D. Appleton-Century Company, New York, 1940.

Baron, M. L., (ed.) Contemporary Sociology. Dodd, Mead & Co., New York, 1964.

Bloomfield, L. P., "The Political-Military Exercise: A Progress Report," ORBIS, 8 (1965), pp. 854-870.

Bogardus, E. S., Fundamentals of Social Psychology. D. Appleton-Century Company, New York, 1941.

Britt, S. H., Social Psychology of Modern Life. Farrar & Rinehart, New York, 1941.

Cantril, Hadley, The Psychology of Social Movements. John Wiley and Sons, New York, 1941.

Churchman, C. W., and Ackoff, R. L., Methods of Inquiry. Educational Publishers, St. Louis, 1950.

Coser, L. A., and Bernard Rosenberg (eds.), Sociological Theory: A Book of Readings. The Macmillan Co., New York, 1957.

Cuber, J. F., Sociology: A Synopsis of Principles. Appleton-Century-Crofts, New York, 4th edition, 1959.

Eliot, T. D., in Fairchild (1944).

Eubank, E. E., The Concepts of Sociology. D. C. Heath and Company, Boston, 1942.

Fairchild, H. P. (ed.), Dictionary of Sociology. Philosophical Library, New York, 1944.

Fichter, J. H., Sociology. The University of Chicago Press, Chicago, 1957.

Gillin J. L., and Gillin, P. G., An Introduction to Sociology. The Macmillan Company, 1943.

Gouldner, A. W., and H. P. Gouldner, Modern Sociology. Harcourt, Brace, & World, New York, 1963.

Guetzkow, Harold, et. al., Simulation in Inter-Nation Relations. Prentice-Hall, Englewood Cliffs, New Jersey, 1963.

Krech, D., R. S. Crutchfield, and E. L. Ballachey, Individual in Society. McGraw-Hill Book Co., New York, 1962.

Lee, A. M., "Level of Culture as Levels of Social Generalization," in Baron (1964), pp. 109-119.

MacIver, R., Society: Its Structure and Changes. Farrar & Rinehart, New York, 1937.

Management Science Center, A Model Study of the Escalation and De-escalation of Conflict, (mimeographed report). University of Pennsylvania, Philadelphia, March 1, 1967.

Mizruchi, E. H., The Substance of Sociology. Appleton-Century-Crofts, New York, 1967.

Osburn, L. D., and Neumeyer, M. H., The Community and Society. American Book Company, New York, 1933.

Rapoport, Anatol, Fights, Games and Debates. The University of Michigan Press, Ann Arbor, 1960.

_______, The Prisoner's Dilemma. The University of Michigan Press, 1965.

Sengupta, S. S., and Ackoff, R. L., "Systems Theory from an Operations Research Point of View," IEEE Transactions on Systems Science and Cybernetics. 1 (Nov., 1965), pp. 9-13.