VALUES AND THE THEORY OF MOTIVATION

by George Edgin Pugh

Despite rapid progress in sociobiology and related areas of behavioral science there are still some very serious gaps in our theoretical understanding. The central premise of sociobiology is that each species is endowed with certain innate behavioral tendencies that are genetically inherited. But the theories do not offer any specific behavioral mechanisms to explain how such innate tendencies can influence conscious rational behavior in human beings or other advanced vertebrate species. When we attempt to apply the present concepts of sociobiology either to practical human problems or to the traditional, human-study disciplines such as psychology, sociology, jurisprudence, or social planning we find that the failure to deal explicitly with the mechanisms of human decision making is particularly serious. The major purpose of this paper is to show how a modernized theory of motivation can be exploited to correct this basic deficiency in the existing behavioral models. The new approach is based on some new insights concerning the theory of values that have developed in the fields of automation and decision science.

With a little reflection it should be apparent that the problems of adaptive control encountered in the design of complex, automated systems must involve many of the same basic problems that have been encountered by evolution in the design of the brain as a biological control system. Thus it should not be surprising if the evolution of the brain had been molded by many of the same principles of information processing and cybernetic efficiency that have dictated the design of the more sophisticated computerized systems.

About seven years ago I began to recognize some striking parallels between the functional design that evolution had employed in the brain and the design principles my colleagues and I were using for computerized decision systems. Indeed it appeared that our design techniques for the automated systems were converging toward a set of principles that seemed progressively more similar to the evolutionary

George Edgin Pugh is president of Decision-Science Applications, Inc., 1500 Wilson Boulevard, Arlington, Virginia 22209.

[*Zygon*, vol. 14, no. 1 (March 1979).]

design concept! This observation led to a systematic study of the analogy in an effort to clarify the basic, underlying cybernetic principles. As a result of this study it now seems clear that both the computerized decision systems and the brain are organized around the same fundamental design concept which I have described as a "value-driven decision system."

I published the basic concepts of the value-oriented theory of human motivation as *The Biological Origin of Human Values.*[1] Since this book provided the first public introduction of the basic concepts, it had to address a large number of theoretical and philosophical issues that might be raised as objections to the theory. As a consequence it was considerably longer than was needed for an introduction to the ideas.

This paper presents a summary of the main concepts in a form that is as close as possible to our commonsense understanding of human behavior and motivation. Such a summary and update of theory seem particularly appropriate at this time because of new developments both in the theory of values and in general behavioral theory which make it possible to present some of the key ideas in a clearer and more readable form. It is my hope that this will avoid some of the misconceptions that have developed from the original text.

In particular it is evident that many readers of the book have had difficulty relating the new theory of motivation to other recent developments in behavioral science. For this reason this paper begins by placing the theory within the broader context of other current developments. In this way I hope to show not only that the theory is compatible with the major, current trends in behavioral theory but also that it provides a unifying conceptual framework which resolves many apparent contradictions among the different schools of thought.

The Origins of Behavior: An Emerging Consensus

The behavioral sciences are in a period of rapid development which is leading, I believe, to a new consensus concerning the genetic, cultural, and cognitive origins of human behavior. This section outlines some of the main features of this emerging consensus to provide a background for discussing the theory of motivation within an overall theoretical framework.

Some of the most important concepts of the theoretical framework are incorporated in what E. O. Wilson describes as the new discipline of sociobiology.[2] These ideas include a formal statistical approach to the analysis of genetic inheritance, a recognition of the genetic and evolutionary basis of the behavioral as well as physical differences among species, and an emphasis on the cooperative and social aspects of behavior as well as the selfish and competitive aspects within the broader context of an evolutionary selection requirement called "inclusive fitness." These basic ideas of sociobiology are complemented by some key ideas in neurophysiology; for example, the brain now is viewed as an information-processing system which has evolved during the course of biological evolution as a behavioral control mechanism for the higher biological organisms. In efforts to understand mental behavior there is an increased emphasis on biologically determined intellectual functions, as in the work on the reward system of the brain or in the work of Charles Laughlin, Jr., and Eugene G. d'Aquili on "biogenetic structuralism."[3] When these trends are combined they lead to the conclusion that basic behavioral tendencies for each organism and each species (including the human species) somehow must be built into the genetic biological design of the brain.

The longstanding disagreements about the degree to which the determinants of behavior are genetic and biological, as opposed to cultural and environmental, seem to be converging toward a more balanced systems perspective which I have tried to show diagramatically in figure 1. In this figure the classical process of genetic selection as it affects behavior is illustrated in the processing loop at the left of the figure. This part of the figure shows that an individual's behavior is a product of his innate motivation system and his experience with the environment. In particular individual behavior is governed not only by an innate motivation system but also by an acquired system of personal values and personal knowledge which are developed by the individual as a result of experience within the physical and social environment. In each new generation the genetic mix (and the genetically determined balance of motivations) is determined by the degree of "reproductive success" experienced by individuals in the previous generation. But this traditional genetic selection process is elaborated in the new consensus in several ways. First, the environment (as illustrated in the diagram) is a social as well as a physical environment; and the social environment reflects the behavior of other individuals. Thus the process of genetic selection takes place not just in a physical environment but also in a dynamic social environment which is itself a product of the genetic evolution of the species. Second, the diagram shows that the social environment is conditioned by traditions and cultural values which influence the success of the culture and therefore are subject themselves to cultural evolutionary selection.

From this broad perspective it is apparent that at least three different time constants are involved in the adaptive evolution of human behavior. The longest time constant is determined by the genetic selection process, which permits the accumulation of experience at a very

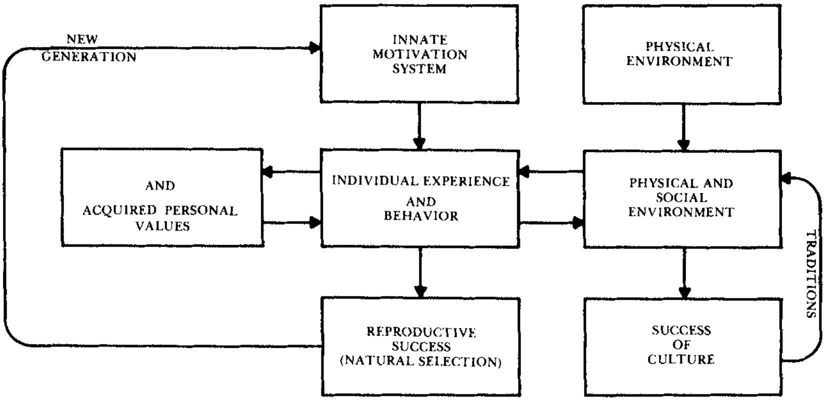

Fig. 1.—Illustrating the genetic, cultural, and cognitive origins of human behavior and human values

slow rate over millions of years. The behavioral wisdom accumulated genetically in the innate motivation system thus reflects the long evolutionary experience of the species with thousands of different cultural experiences. From the perspective of any individual (or any single culture) this genetically determined motivation system has to be viewed as an essentially invariant element.

The shortest time constant in the behavioral system is provided by individual learning and the acquisition of personal values from individual experience, as shown in the box at the left of the figure. This capacity for individual learning makes it possible for human behavior to adapt very quickly to radically changed situations and permits some anticipation of potential problems.

The evolution of cultural values and traditions (as shown in the right-hand loop), provides an intermediate time constant which permits the accumulation of experience over many generations. Whereas the innate motivation system contains wisdom accumulated over millions of years, the cultural traditions contain wisdom about the more recent past in the range of hundreds or thousands of years.

In most social species other than man the normal social structure in the wild can be viewed as an equilibrium social state which reflects the genetic inheritance of the species. Even socially inexperienced animals (when transplanted to a new environment) will tend to regenerate a new social environment characteristic of the species. This of course is less true for modern man because the accumulated cultural knowledge of hundreds of generations cannot be replaced quickly. However, it seems likely that it was essentially true for man's prehuman ancestors—before the emergence of a linguistic capability adequate to transmit extensive cultural information.

One of the important insights of sociobiology is that even within the equilibrium social structure of any species certain conflicts of interest are inevitable. The behavioral strategies that are genetically most productive for males may be in conflict with the strategies that are most productive for female members of the species; thus there is an almost inevitable sexual conflict of interest. There is also an inevitable reproductive competition among males. And finally the behavioral strategies that are genetically most productive for the individual usually will not coincide exactly with the strategy that would be most effective for the survival of the society or the species as a whole. Thus the social environment itself generates new selection forces which tend to favor those individuals who can optimize their own genetic success despite the conflicts of interest.

The recent quantitative analysis of evolutionary stable strategies (ESS) based on genetic inheritance suggests that many of the elaborate behavioral patterns such as courtship and mating rituals, social hierarchies, and social leadership roles for mature adults have evolved as mechanisms for coping with these genetic conflicts of interest.[4]

Whereas in the less intelligent species these behavioral strategies are defined largely by genetic inheritance, the adjustment in the human species has been more dependent on individual learning and cultural traditions. Thus human cultural traditions often have served to inhibit individual behavior that is in conflict with the best interests of the society. As Donald T. Campbell observes, the subject matter of folk moralizing in a society can be predicted approximately on the basis of the conflict between the genetically defined motivations of the individual and the best interests of the society.[5] In modern human society this conflict is probably more severe than for most other species not only because of the tendency of human evolution to rely on cultural solutions but also because the modern urban environment is very different from the primitive society for which the innate human motivation system has evolved. Because of this dependence of human society on cultural inhibitions and controls it appears that the recent tendency of social scientists to discredit the traditional cultural inhibitions may be in conflict with the best interests of society.

It appears that a realistic approach to human planning must recognize explicitly the existence of inherent conflicts of interest. The practical objective therefore must be not to eliminate the cultural inhibitions but to devise a modernized social structure in which the necessary social inhibitions can operate and the innate cooperative human motivations can be stimulated and rewarded more effectively. The real goal of such social planning should be to provide a human social environment in which life can be as satisfying, meaningful, and secure as possible within the practical limits defined by the human genetic inheritance.

Limitations of Prevailing Behavioral Models

It is evident that rapid progress is being made toward a more realistic and objective theoretical understanding of human behavior and human society. Nevertheless, when we attempt to apply the present concepts of sociobiology to the traditional, human-study disciplines such as psychology, sociology, jurisprudence, or social planning we find that there are serious deficiencies in the available scientific models. The most significant problems result from the failure of the theories to deal explicitly with the human-decision process and other mechanisms of behavior.

As David P. Barash has observed aptly, sociobiology is concerned with distal rather tha proximate causes.[6] It relates behavior to genetic inheritance without addressing the mechanisms of behavior. In contrast the traditional human studies have sought to understand behavior in terms of the decision process and the interactions of the individual with his environment. Since sociobiology has had little to say about these mechanisms of behavior, its present contribution to traditional human studies is necessarily somewhat limited. When its concepts are applied to practical human decisions its failure to provide a credible behavioral model becomes particularly serious. From our introspective mental experience we have the strong impression that we behave rationally and that our behavior is frequently, if not always, the result of a conscious, "rational" choice. But how can we reconcile these subjective impressions of consciousness, rationality, and free will with sociobiology's concept of genetically inherited behavioral tendencies?

The "behavioral models" that are most widely accepted do not provide a satisfactory resolution of the problem. The prevailing scientific theories of behavior seem to fall into two broad categories. First, there are the formal, behavioral models based on simple cybernetic concepts such as stimulus-response, feedback, homeostatic response, and behavioral conditioning. These models have the merit of being relatively objective and uncontaminated by subjective, "mentalistic" concepts. Unfortunately they yield a vision of man that is flat—without personality or purpose. They seem more suitable as a behavioral model for an amoeba than for a human being. Second, there is a wide variety of relatively informal scientific theories such as Abraham H. Maslow's theory of motivation which explicitly incorporate mentalistic constructs and correspond more closely to the commonsense view of human nature.[7] Although these theories are of considerable value in practical clinical and psychological work, they lack the formal structure that is necessary for a quantitative scientific theory.

Human behavior and experience cannot be described adequately in terms of behavioral tendencies or operant conditioning. Human beings have desires, ambitions, objectives, and purposes; they experience pain and loneliness, joy and sorrow; they have disappointments, pleasures, and creative inspirations; they make decisions and they make moral judgments. An adequate theory of human behavior should somehow incorporate and account for these important aspects of human experience. Ideally we would like these familiar aspects of human experience to emerge naturally from a formal and quantitative behavioral model.

The absence of a valuative dimension also must be viewed as a key weakness in the prevailing scientific models. Somehow all human beings make value judgments, and they make decisions in terms of personal value criteria. A satisfactory theory of human behavior should account for this tendency to make decisions in terms of values. It should explain somehow the intuitive mental processes that allow human beings to make the transition from "is" to "should" or "ought."

Overview of a New Motivational Model

This section provides a very brief introduction to the main concepts of a value-oriented theory in a form that is as close as possible to common-sense understanding of human behavior and motivation. Such a non-technical introduction seems desirable to avoid a variety of misconceptions that can arise when the concepts are developed in a more formal and technical way. Unfortunately when the concepts are presented in this informal way the match to common sense and subjective mental experience is so close that it is easy to assume that there is no new content to the theory. However, as I hope to show in later sections, the theory provides a formal and potentially quantifiable model of motivation and behavior which parallels common sense much more closely than existing scientific models and also leads naturally to many of the "mentalist" concepts of classical psychology.

To place this theory of motivation within the context of the genetic, biological, and cultural origins of behavior the reader is referred again to figure 1. Within this broad framework the present theory is concerned primarily with the behavioral mechanisms that provide the link between genetic inheritance and species-specific behavior. In figure 1 therefore the focus is primarily on the box labeled innate motivation system.

Obviously a balanced understanding of human behavior must take into account all of the interactive processes shown in figure 1. These processes interact with one another not only within the life of a single individual but also over the course of human evolution, so that the design of the innate motivational system itself reflects all of the interactive processes. Since the present theory of values is concerned primarily with the motivational system, however, the discussion tends to focus primarily on this aspect of the problem. The other factors are taken into account only to the extent necessary to allow the reader to relate the ideas to the other fields of interest.

In figure 1 the innate motivational component has been labeled "innate motivation system" rather than "innate motivations" because according to the theory only the motivation system is innate or genetically inherited. Motivations are not innate but generated in the cognitive processes of the organism as a result of experience with the environment and with the innate motivation system. The resulting cognitive motivations are so numerous and complex that, as Maslow comments, they defy any simple or systematic classification. However, the output signals from the innate motivation system which lead to the cognitive motivations are not nearly so diverse. In our conscious minds we experience the output signals from this innate motivation system in the form of specific valuative sensations such as pain, hunger, joy, sorrow, or shame.

The motivational effect of these signals derives from their positive or negative contribution to a unified, internal (utility) scale of pleasure or displeasure. The organism is so designed that (in its cognitive decision processes) it will try to avoid situations where the valuative signals are negative (or aversive), and it will seek situations where the valuative signals are positive (or rewarding). The evolutionary design of the innate motivation system therefore is concerned primarily with the way these subjective valuative signals are related to the objective experience of the organism. The way the valuative signals are related to experience can be very simple and direct (as in the case of pain or hunger), or it can be very complex (as in the case of joy, sorrow, or shame).

During the course of biological evolution the functional relationship between these valuative signals and the objective experience of the organism is adapted to produce a balance of motivations in each species that contributes as effectively as possible to inclusive genetic fitness. From a system-design perspective we can say that the valuative sensations are related to the experience of the organism through functional relationships (analogous to mathematical functions) which are an essential part of the evolutionary design. For example, there is a functional relationship between the pressure of a pin on a finger and the subjective sensation of pain. Similarly there is a relationship which determines the intensity of the hunger sensation as a function of the fullness of the stomach and the state of nourishment of the organism. At a less obvious level the valuative sensations such as joy, sorrow, shame, and pride which motivate human social behavior seem to be functionally related to certain types of social experience. For example, evidence of social approval seems to contribute to happiness, whereas evidence of scorn or disapproval contributes to shame or sorrow. Since the innate motivational system exerts its influence on behavior by means of these valuative sensations, the motivational system can be described as a value-generating system (or a system of innate values) which provides the organism with its primary criteria of decision. Thus the species-specific "behavioral tendencies" in the higher vertebrates appear to be encoded in a genetically defined set of valuative sensations. These valuative sensations serve as the primary motivators for the cognitive decision processes which control conscious behavior. The fact that the primary motivational information is supplied in the form of a system of values helps to explain the tendency of the brain to develop personal and cultural value associations and to use these value associations to facilitate the decision process.

In traditional psychological literature the basic value sensations have been classified not as "values" but as basic drives, emotions, sensory experiences, etc. Nevertheless, from a theoretical decision-theory perspective they must be classified as values. They are valuative (i.e., scalar) quantities that are associated with "outcomes" for the purpose of guiding a decision process.

The consciously generated behavior of the organism is not genetically specified but governed by a cognitive learning and decision process. Of course the capacity of the brain to support such rational decision processes and many of the commonsense procedures for learning and deciding appear to be defined as a part of the human genetic inheritance. As a result of experience with the environment (and with the valuative signals received from his innate motivation system) the individual is able to use his cognitive abilities to develop a mental model of his social and physical environment, including the cause and effect relationships within that environment. Because the valuative sensations are delivered to his cognitive consciousness as an integral part of his experience with the environment, the valuative sensations are incorporated automatically as a part of his mental model. When the mental model is used to project probable outcomes for alternative courses of action, the appropriate valuative sensations (fear, joy, pain, sorrow, or shame) are incorporated automatically as a part of the anticipated outcome. Thus these innate valuative sensations serve as a primary guide to the organism's decision processes.

The Interaction between Innate and Cognitive Values in the Evolution of Social Behavior

Although the innate valuative sensations serve as a primary guide to the organism's decision processes, figure 1 shows that behavior is not governed solely or directly by them. Individual learning and acquired personal values are vital. The majority of decisions are made on the basis of habit, rules of thumb, ethical principles, and personal values which are learned from experience and serve to simplify the decision process. Most practical decisions take too much mental effort (and too much time) for their outcomes to be projected far enough into the future so that they can be assessed in terms of the primary value sensations. Consequently the individual gradually develops cognitive associations which lead to simple rules of thumb, ethical principles, and personal values which simplify the decision process. In routine situations these simple decision rules evolve into subconscious habits. But the individuals' ultimate measure of the "quality" of the rules of thumb, the ethical principles, the personal values, and the habits is the extent to which they lead to outcomes that are "desirable" when evaluated in terms of the primary value sensations.

From a decision-theory perspective the value sensations can be viewed simply as a system of "primary values" that have been built into the human decision system as an essential part of the evolutionary design. The rules of thumb, the ethical principles, and the personal values can be viewed as a network of secondary values that are derived from the primary value system on the basis of experience with the physical and social environment.

This theoretical classification of values into primary values (those that are an essential part of the system design) and secondary or cognitive values (that are learned from experience) has been criticized by some because it does not explicitly include cultural values.[8] Actually of course when cultural values are adopted by the individual they become a part of his system of secondary values. This valuative interaction between the individual and the society with its cultural values and traditions plays a very important role in the development of human behavior. Although these issues were discussed in some depth in chapter 14 of my book, it appears they were not given sufficient emphasis.[9]

In human society the secondary value criteria (the rules of thumb, the personal values, and the ethical principles) tend to be socially communicated. They become a part of the social environment, and over a period of generations they develop into cultural values and traditions, which influence the success of the culture and are subject to a cultural evolutionary process as shown in the right-hand loop of figure 1.

But the development of the cultural values and the cultural tradition involves far more than just an accumulation of individual learning and personal values. Indeed most of the traditional cultural values are concerned not with what is rewarding to the individual but with how the individual would like others in the society to behave toward him! In all probability this intense concern with the behavior of others is a genetically inherited behavioral tendency that is built into the human motivation system. This concern with how others should behave is a continuous source of gossip, conversation, and storytelling within the primitive human society. It defines the character traits (bravery, honesty, sincerity, etc.) and the behavior patterns that are admired within the community. To understand the operation of primitive human society it is important to recognize that this concern with the behavior of others is not just "idle gossip." The gossip serves an important functional role in the operation of the society since it actually influences the way other individuals behave.

Some of the most important, human value sensations (pride, joy, sorrow, and shame) respond most directly to the way we are viewed by others. When we are admired we feel pride and joy. When we are scorned we feel shame and sorrow. These basic, social value sensations generate the motivation to be admired. They lead almost automatically to the cooperative patterns of behavior that are so characteristic of primitive human society. Although it is obvious that these cooperative patterns of behavior contribute to the effectiveness and survival probability of the society, it may be less obvious how the innate value sensations that motivate these patterns of behavior could have been sustained genetically as an evolutionary stable strategy.[10]

Like the social behavior in most species, the evolutionary development of these behavioral mechanisms seems to involve the interaction of a number of behavioral traits. The evolutionary stability of the motivations cannot be explained without explicitly considering the interactions between different behavioral traits. It seems obvious that as long as the desire to be admired is present in a society an individual's tendency to admire cooperative and altruistic behavior on the part of others will motivate them to behave in ways that are more favorable to his own interests and in fact will contribute to survival probability and enhance genetic fitness. The real problem therefore is how the motivation to be admired has been sustained during human evolution as a stable strategy. The most obvious answer seems to be that individuals who failed to measure up to acceptable social standards were targets for anger and were likely to be attacked, ostracized, or otherwise denied the protection and benefits that the society could offer. Such a tendency to become angry in the face of antisocial behavior and to penalize it by exclusion from the community clearly would contribute to an enhanced genetic fitness for the remaining members of the community. Moreover, given such behavior on the part of the community, it seems clear that an innate desire for social acceptance (or a desire to be admired) in fact would contribute positively to genetic fitness. Thus the evolutionary, stable social strategy for the human species seems to involve the interaction of three separate factors: a tendency to anger in the face of antisocial behavior, a tendency to admire altruistic cooperative behavior, and the desire to be admired.

In the overall pattern of human motivation the desire to be admired (as motivated by sensations such as pride, joy, shame, and sorrow) is balanced by a large number of essentially selfish motivations such as avoidance of pain and hunger, etc. Decisions in any specific situation will tend to reflect the relative strengths of the different value sensations. Evidently the evolutionary stable strategy requires a balanced behavior in which the concern for the opinions of others is neither too weak nor too strong. If the concern is too strong it leads to an overly altruistic behavior which negatively would affect inclusive genetic fitness. If the concern is too weak the dominance of personal selfish interests in the pattern of behavior can result in ostracism from the mainstream of the society with a corresponding loss in inclusive genetic fitness.

When the structure of the innate human motivations is viewed broadly from such a system-design perspective, it seems clear that the evolutionary process has arrived at a balance of motivations which strongly encourages cooperation and reciprocal altruism in primitive human society but which nevertheless tends to discourage any excesses of pure altruistic behavior.[11] In traditional human society the effectiveness of the cooperative motivations tends to be enhanced by the accumulation of cultural values and traditions, which over a period of several generations tend to be sanctified as an almost absolute source of moral authority.

Obviously it is within the social context of his own society that each individual learns his own personal values. Because his innate valuative sensations make him want to be admired and accepted within the society, the individual also is motivated strongly to accept the society's values as his own. It is this type of learning process that accounts for the extreme receptivity of small children to "cultural conditioning."

From this perspective the moral conflicts that play such a prominent role in folk literature appear to be an inevitable consequence of the human motivation system. In an abstract logical sense it might appear that there should be no such conflict since the individual should be motivated to select the course of action which will be most rewarding to him over the long run. In practice, however, the choice is not so easy. On the one hand the individual is motivated by certain essentially selfish values: avoidance of pain and hunger, avoidance of the discomfort of unnecessary labor, etc. On the other hand he is motivated by the desire to be admired and to achieve a position of respect and eventually leadership in the society. In many cases these two sets of motivations are in conflict. In terms of his selfish motivations the individual "wants" to do what is easy and pleasant. But because of his desire to be socially accepted he knows that he "ought" to conform to the ideal behavior standards as defined by social tradition. The conflict between what he "wants" to do and what he "ought" to do is the basic moral dilemma of human tradition.

Because folk moralizing is fundamentally concerned with what others "ought" to do it tends to set moral standards that are somewhat higher than the average individual actually will achieve.[12] This serves a functional purpose in the culture since it tends to motivate behavior which is more beneficial to the society as a whole, but it also intensifies the individual's moral dilemmas. Because the rewards of selfish behavior tend to be rather immediate, whereas the rewards of altruistic behavior (in terms of personal reputation) tend to be cumulative and long term, much of folk moralizing is concerned with the importance of thinking ahead rather than yielding to short-range temptation.

In primitive human societies where social interactions are on a personal, face-to-face basis and where each person is known as an individual, the innate, human social motivations operate as an integrated system to motive effective cooperative behavior. They also go a long way toward resolving the underlying genetic conflicts of interest that otherwise would disrupt the operation of the society. The innate social motivations do not operate so effectively in the modern urban society. In a large urban center most contacts tend to be on an impersonal basis in which individual reputation and personality seem irrelevant. Moreover, because of the transient nature of most urban living, personal reputations do not have the long-term importance that they have in small, stable communities. In such a society the human social-motivation system cannot function as effectively; the individual tends to feel that life is meaningless and without purpose, and the society tends to be much more dependent on formal law-enforcement mechanisms to maintain the social order.

THE VALUE-DRIVEN DECISION SYSTEM

The basic concepts of the theory were developed in a nontechnical form in the two previous sections. However, the theory would be of very little scientific interest if it were limited to such an informal discussion of essentially commonsense concepts. Although it is not possible in a short paper to give a full discussion of the technical concepts, it seems important to provide some insight concerning the theoretical foundations of the theory. This section provides a brief introduction to some of the concepts of computer automation that stimulated the initial development of the theory.

Fundamentally the theory is built around the design concept for a specific class of automated systems, which I describe as value-driven decision systems. In the design of computerized automation systems we find that as the decision environment becomes more complex the feasible design concepts seem to converge toward this single, basic design principle. Although simpler design principles such as feedback loops or stimulus-response systems are efficient and appropriate for relatively simple control problems, they become progressively less useful as the decision environment becomes more variable and complex. In a complex and variable environment only the value-driven decision system seems to be capable of generating reasonably efficient and satisfactory behavior. It seems reasonable to suppose that evolution's design of the brain may have been molded by very similar considerations so that it also has tended to converge on the value-driven approach as a basic design concept.

In a recent effort to codify the alternative design concepts for automated control systems we found all of the systems studied could be classified within the framework of two basic design concepts: the familiar, stimulus-response design concept and the value-driven decision system, which specifically considers alternatives and chooses among them on the basis of value criteria. Although other design principles such as feedback control, associative learning, and the equivalent of a conditioned response also were used in various parts of the systems, we found that the highest level of control for a fully automated system dealing with a complex problem environment could be described always in terms of one of these basic concepts or as some kind of hybrid between the two design concepts. Most of the early automation systems tend to fall in the stimulus-response category, whereas the more recent systems that deal with more complex problem environments tend to follow the paradigm of the value-driven decision system.

Since both biological evolution and the design of automated computer systems appear to have gone through this progression from stimulus-response to value-driven designs, it may be worthwhile to compare the basic structure of these two concepts as they appear in the computerized system designs. Figure 2 shows a side-by-side comparison of the two design concepts in the form of a very simple, stylized flow diagram. The flow diagram on the left corresponds to the stimulus-response design concept, while the one on the right illustrates the value-driven design. Notice that there is almost an exact functional correspondence between the first three steps in both design concepts. Thus the final steps in the value-driven design can be viewed as a processing filter that is added to improve the quality and adaptability of the decision process.

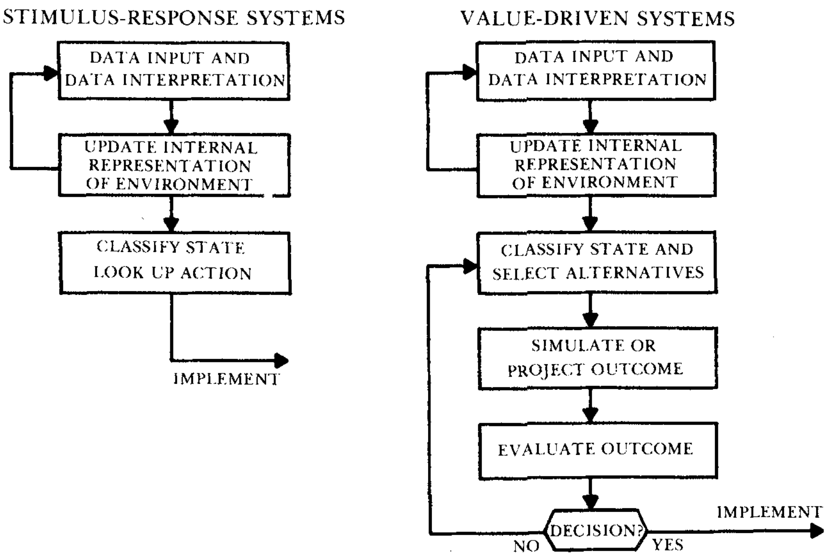

Fig. 2.—Comparison of stylized diagrams for computerized decision systems

The first two steps in both system designs serve to update the internal symbolic representation of external reality. In the simplest stimulus-response system this representation may be nothing more than a very simple classification system which recognizes certain key stimuli for which some specific response is appropriate. As the desired pattern of behavior becomes more complex in a stimulus-response system, it becomes necessary to recognize and classify a larger and larger number of stimuli (and combinations of stimuli) that are to serve as action releasers. In the limit of very complex stimulus-response behavior systems this leads to the requirement of an almost complete classification of states of the environment, which begins to correspond to an internal symbolic model of external reality.

The third and final step in the stimulus-response decision mechanism consists essentially of a table lookup process which allows each classified state of the environment to be linked to a specific, predefined muscle response. To complete such a behavioral system one of course will require another cybernetic system which contains the repertoire of predefined muscle responses.

The essential difference between the computerized, stimulus-response decision system on the left in figure 2 and the value-driven system on the right concerns the final steps in the decision process. Whereas in the diagram on the left the classification of states of the environment leads to a single, well-defined course of action, in the diagram on the right it leads to a number of alternative action possibilities that are suggested as promising courses of action. To determine which of these possibilities should be implemented the system goes through two additional steps which serve to evaluate the desirability of the alternatives. First, a mental model is used to project probable outcomes for each of the alternatives. Second, the projected outcomes are evaluated in terms of some predefined value criteria. The alternative that scores highest in terms of this value criteria is chosen for implementation. Of course in the design of specific systems there can be a wide variety of variations on this basic flow process. Often the decision process is broken down into a hierarchy of decision steps in which broad strategic decisions are made at the highest level, followed by tactical decisions at an intermediate level and detailed implementation decisions at the lowest level. Nevertheless, the simplified, basic flow diagram of figure 2 can serve to illustrate the essential features of the value-driven decision mechanism.

In the design of computerized decision systems the use of the value-driven concept requires two very important cybernetic components which add complexity to the design process but which pay great dividends in terms of the adaptability and quality of the decision process. These two components are (1) a "mental model" (in symbolic form) which is detailed enough to be used to interpret incoming sensory data and to project the probable "outcomes" for alternative courses of action and (2) a "value function" which can be used to assign a value (or measure of desirability) to the projected outcome for each course of action. The quality of the resulting decisions is of course critically dependent on the quality of both of these processing elements.

The importance of the mental model of the environment as a key functional element within the value-driven decision system cannot be overemphasized. The functional requirement for such a component helps to explain why biological control systems have tended to evolve into systems capable of maintaining and using such a mental model. Moreover, it is worth noting that systems using such a model will tend to behave purposefully as if they were both conscious and rational. Experience with the artificial decision systems has shown clearly that the quality of performance attainable with such a system depends critically on the quality of the mental model. The model must be able to project probable outcomes with reasonable reliability; otherwise satisfactory decisions cannot be made. It seems clear that biological evolution also has encountered this same relationship between the quality of the model and the quality of the resulting decision (or behavior). The tremendous cybernetic resources within the brain that are devoted to provide a capability to maintain a good mental model of the world environment stands as effective testimony to the importance evolution has assigned to a good mental model.

The second striking characteristic of the design concept lies in its dependence on values as an integral part of the control mechanism. From a formal mathematical perspective a value is simply a number (or scalar quantity) that is associated with outcomes for the purpose of making a decision. When we design an artificial decision system we must devise a value function (i.e., a mathematical function) that can be used to calculate a numerical "measure of desirability" for the outcome of each alternative course of action. This design of the value function is usually one of the most difficult, and it is always one of the most important parts of the system design.

Experience with artificial systems has shown that the behavior of such a system is critically dependent on the design of the value function. To provide satisfactory behavior for an artificial decision system the designer often must devise a very complex value function which may be time dependent. Thus the value criteria themselves (like human emotions) may be complex functions of both the recent experience and the current state of the decision system. The structure of the value criteria chosen by the designer turns out to be the dominant factor governing the behavior of such a system. It is the value criteria that define the motivation of the system and determine which outcomes are judged as "desirable" or "undesirable." Small changes in the value criteria can make major changes in the behavior of such a system. Whereas the mental model determines how well the system will be able to predict outcomes, the value function determines which outcomes the system will "prefer."

Experience in the design of artificial decision systems demonstrates clearly that the value criteria that provide the system with its ultimate decision criteria must be supplied externally by the designer or a user of the system. The decision system cannot create or change these ultimate criteria of decision. If it were to do so the resulting criteria almost certainly would fail to serve the objectives of the system designer. Consequently the ultimate value criteria for a value-driven decision system must be viewed as an essential part of the system design. Moreover, in order to predict the behavior of such a system one must be able to relate the behavior to these basic value criteria. Thus the built-in value criteria are also an essential part of any scientific description of the system.

Since biological evolution played the role of designer for the brain, the innate value structure must be included as a basic component of the genetic design. This of course is entirely consistent with the basic premises of sociobiology. The built-in value criteria provide a specific behavioral mechanism which allows the genetically inherited "behavioral tendencies" of sociobiology to be incorporated within a "rational" decision process.

In human experience these built-in value criteria are perceived subjectively as quite distinct valuative sensations, namely, pain, hunger, good- or bad-taste sensations, joy, sorrow, pride, shame, and anger. Some of these valuative sensations such as pain or hunger are related to the experience of the organism by rather simple, direct functional laws. Others such as pride, sorrow, and anger are related to experience through a much more complex, functional relationship. These genetically determined, innate human values of course are not usually referred to as "values." They are described more commonly as drives, or emotions, or sensations. Nevertheless, from a system-design perspective they are values. And these innate values are the primary determinates of human motivation.

Human beings behave characteristically in the way we expect human beings to behave because they are motivated by this genetically inherited, innate value system. Human behavior and motivation can be viewed thus as a natural consequence of a complex structure of innate values. From a system-design perspective the purpose of these valuative sensations is to motivate human beings to behave in ways that during the course of human evolution have been found to be genetically productive. If we are to develop a scientific understanding of human motivation, we will need to understand this complex structure of built-in value sensations; and from the perspective of sociobiology we will need to identify the functional role of the various components of this value structure during the biological evolution of the human species.

Although the brain operates as a value-driven decision system and uses some of the same cybernetic principles as our computerized systems, it also differs from the artificial systems in a number of important ways. In a computerized decision system the mental model that is used to project outcomes must be supplied by the system designer. In effect it is a part of the original system design. The biological decision systems in contrast are able to develop and refine their own mental models. (Actually a closer examination reveals that some important components of the human mental model are built in genetically. But the overall model is an achievement of the individual decision system, and it reflects the individual's personal experience with the environment.) At present we are a long way from being able to duplicate this process of creative model development in any computerized system.

There is one other very important difference in the way the basic value-driven concept is implemented in the biological system. Note that, in the computerized system (on the right of fig. 2), values are calculated for the projected outcomes as a final step in the decision process. In this computerized design the built-in value system is exercised explicitly to assign values for every alternative that is considered.

The evolutionary system design in this respect is quite different. The innate or built-in value structure operates in the first step of the process as if it were part of the data input process. Figure 3 illustrates a rearrangement of the idealized flow chart to accommodate this change. With this change in the design the value information appears at the data-input phase in such a way that the valuative elements (pain, hunger, joy, sorrow, etc.) are delivered to the rational analysis system as if they were an integral part of the external experience of the decision system. Consequently in step 2, when mental models of the decision environment are developed by the decision system, the models automatically include the valuative component of the experience. When the models are used to project outcomes for alternative courses of action, the projected outcomes inevitably will include the valuative component. Thus in this revised system design there is no need for an arbitrary assignment of values to the outcomes. Since the innate value sensations are included in the mental model as an integral part of the predicted outcome, all that is required to make a decision is to assess the various components of these values in the predicted outcomes to determine an overall valuative score for each alternative.

Although such a preprocessing of input data in the brain to add the valuative information may seem surprising at first, it is entirely consistent with standard design techniques used in the brain. Indeed a great deal of preprocessing of the input data is accomplished before the information is delivered to the rational or conscious mind. For example, the visual input data are processed to assign colors to objects and to attach depth-perception information. This additional information then is delivered to the rational mind as if it were an integral part of the original sensory data. We are not conscious of this preprocessing of sensory information and have no introspective mechanism that will allow us to recover the logic that is used. In a very similar way the value information is added to the experience (before the information reaches the rational mind) in such a way that we are completely unaware of the cybernetic processes that accomplish the value assignment.

In the diagram shown in figure 3 the last four steps of the decision process take place within the rational conscious mind. Consequently we have a subjective awareness of these processes, and we gradually develop an intuitive understanding of how they operate. The mechanisms that assign the innate valuative sensations, however, take place outside the conscious mind, so that this value assignment is a fait accompli by the time the information reaches the conscious mind. Thus we perceive this valuative information as if it were an essential part of objective reality.

In the foregoing discussion I oversimplified rather grossly the theory of values. In fact even a computerized, value-driven decision system usually will make use of two types of values which I refer to as primary and secondary values. Primary values are those supplied by the user or the designer; these are the values I have been discussing so far. In contrast, secondary values usually are developed and refined by the decision system itself. They serve as a practical aid to decision. They help the system make decisions that would be evaluated as "good" when measured against the primary or built-in value criteria. They simplify the process by allowing the system to make good decisions with less computational effort.

The values we assign to chessmen are a good example of this principle of secondary values. Through the use of these secondary values we can judge the desirability of an exchange of pieces without having to project the detailed moves of the game through to the final move. Although the only real objective of the game is to achieve a checkmate, the secondary values that we assign to the chessmen make it easier to make good decisions during play. To provide a good secondary-value criteria the secondary values must reflect the primary goals. They must yield decisions that are "good" when measured against the primary-value criteria.[13]

Our analysis of human behavior indicates that our cognitive values such as ethical principles, social norms, social taboos, and the routine dos and don'ts of daily living are all forms of human secondary values. If these secondary-value criteria are to be valid and useful, they must lead to decisions that will be generally "desirable" when evaluated

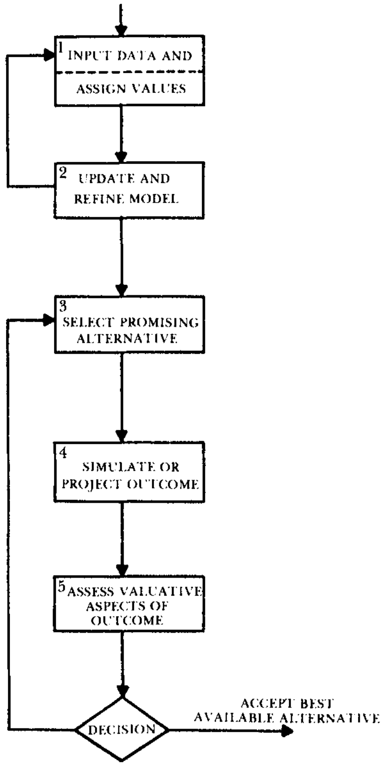

Fig. 3.—The human decision process (a simplified canonical taxonomy)

against the innate or built-in human-value structure. In ordinary conversation "values" almost always are concerned with these "secondary" or cognitive values. Since these secondary-value criteria are generated by the rational mind (as a result of experience with the physical and social environment), they are a natural subject of discussion and are likely to be somewhat controversial. The innate or primary-value sensations appear to be such a part of our physical experience that they usually are accepted as natural and inevitable. They are not controversial, and they rarely need to be discussed.

However, to provide a valid scientific explanation of the secondary values we must start with the innate or primary human values, for these are the basic source of human motivation and they determine what secondary-value criteria the individual will be willing to accept.

Before completing the discussion of figure 3 I must emphasize that it is a grossly oversimplified representation of the human decision process. It is used here only to illustrate the basic decision principles that are involved. In fact the actual, human decision process uses many complex combinations of the steps shown in the figure and many different forms of mental models ranging from very simple and aggregated to very complex. Part of my present work is an effort to analyze and classify in more detail the principles of cybernetic efficiency that are employed to guide our "commonsense," mental decision processes. But it is not necessary to pursue them at this time.

Let me summarize now what appears to me to be the significance of this new perspective. It shows clearly how our "rational" decisions can be motivated in fact by genetically inherited "decision criteria." It provides a concrete link between the concepts of sociobiology and the mechanisms of human behavior and thus ultimately a more scientific foundation for the traditional human studies such as psychology, sociology, ethics, and social policy.

It also offers a surprisingly simple resolution of the perennial controversy concerning the role of nature versus nurture as determinants of behavior. With regard to conscious or voluntary human behavior the theory suggests that the genetically defined guidance is contained primarily if not entirely in the subjective valuative sensations. Actual behavior is governed by a rational learning and decision process which responds adaptively to experience with these valuative sensations within the context of the physical and cultural environment. But even the innate motivational system is not free from environmental influence. For example, the development of "sentiments" requires an associative learning process within the motivational system itself so that specific emotional signals can be generated in response to specific persons or places.

Because of space limitations I cannot attempt a comprehensive discussion here of the structure of the innate values or the principles of cybernetic efficiency that seem to have molded the design of the brain. For a more detailed discussion the reader should consult my book.[14]

Neurological Evolution of the Decision Mechanism

At the beginning of the previous section I discussed the essential functional differences between the simple stimulus-response decision mechanism and the more sophisticated value-driven decision system. At least in computerized systems these differences are great enough to make it difficult to understand how such a transition from the stimulus-response to the value-driven design concept could have been accomplished within the constraints of a gradual evolutionary process. This section outlines in nontechnical terms a specific hypothesis about how the transition was accomplished in the evolution of the vertebrate brain. Readers not interested in this material can skip the section without missing any essential ideas.

We recall that even a stimulus-response decision system must be able to classify external stimuli or states of the environment according to the appropriate response. In order to generate a complex stimulus-response behavioral repertoire this internal system for classifying states of the environment must begin to approximate a rudimentary mental model. A rather sophisticated stimulus-response system of this type probably corresponds roughly to what existed in primitive fish before the emergence of the forebrain (when the brain consisted of only two rather than three neurological enlargements at the head of the spinal column). In all probability the system for classifying states of the environment was contained in the frontal enlargement of this primitive brain (where the sensory channels were attached), whereas the predefined muscle responses were encoded mostly in the posterior enlargement which is known as the hindbrain.

It seems likely that the third enlargement at the head of the spinal column, which is now known as the forebrain, may have developed originally as a result of some mutations which affected genetic control and caused the pattern classification system (i.e., the rudimentary mental model) with its associated sensory input channels to be duplicated—producing the third enlargement at the head of the spinal column. During subsequent evolution the primitive, stimulus-response behavioral mechanisms of the midbrain came to be moderated by higher-level control processes in the forebrain. Thereafter the stimulus-response mechanism in the midbrain began to relinquish direct control of action and evolved into a system of motivating values. At the same time the mental model and the associated neurological control mechanisms of the forebrain became more elaborate and ultimately evolved into what we now can recognize as a rational decision mechanism driven by motivating value sensations.

In the primitive vertebrate brain the valuative sensations appear to have been generated in the midbrain and were transmitted via the hypothalamus to the forebrain for cognitive processing. In later evolution portions of the forebrain cortex including the frontal lobes and at least part of the limbic cortex seem to have been annexed into the motivational system to provide a more adaptive and flexible primary value system.

In the early versions of the system the input data required by the motivational system were supplied directly to the midbrain via the original sensory channels. However, this arrangement required an inefficient duplication of sensory processing systems both in the midbrain and the forebrain. Later the sensory processing system in the cortex of the forebrain developed so that it became far superior to that in the midbrain. The motivational system then was modified to obtain most of its input data indirectly by tapping into those areas of the forebrain cerebral cortex where the necessary processed sensory data are available. As a result in the higher vertebrate brain the independent sensory processing systems of midbrain have begun to atrophy.

Although the interpretation of the brain as a value-driven decision system began as a theoretical cybernetic concept, there is now a growing body of neurological data which seem to support the theory. For example, my conjectures about the role of the midbrain and the frontal lobes in the motivational system now appear to be supported by solid neurological data, and neurologists are beginning to analyze the motivational system from a functional perspective which is quite compatible with value-theory concepts.[15]

Relation to Existing Scientific Concepts

When the value-theory concepts are presented informally without much theoretical background, the approach may seem to offer little that is different from prevailing commonsense ideas. In fact it involves a number of important departures from the prevailing behavioral theories. For this reason it may be desirable to review some of the key differences and show how they can lead to a better understanding of behavior and more accurate theoretical predictions.

Superficially the value-theory interpretation of behavior appears to be a rather straightforward extension of the familiar theory of biological drives, but in practice there are a number of very significant differences:

1. In traditional psychology the emotions and valuative sensations such as taste and pain are treated as totally different psychological entities, but in the present approach they all are incorporated as an integral part of a unified theory of motivation.

2. In traditional psychology the emotions are treated usually as intuitively obvious consequences of experience, whereas in value theory they are treated as a functional part of the human motivation system. This change in perspective suggests new questions for psychological research. For example, we can ask for a functional understanding of each emotion. What specific types of behavior is it supposed to motivate? Quantitatively how does the emotion respond to specific social stimuli? It also raises questions concerning the genetic and evolutionary origins of each emotion. How did the emotion originate? How is it maintained as a part of an evolutionary, stable motivational inheritance?

3. Traditional behavioral theories generally have tried to explain behavior in terms of a minimum number of motivating factors. In contrast, value theory predicts on a priori grounds that the number of motivating factors shall be large, that the relationship of many of the factors to experience will be complex, and that the specific structure of the motivating values cannot be predicted on a priori theoretical grounds. The structure of the innate motivating values used for illustration purposes here is grossly oversimplified. Although the actual structure is not known with any precision, the basic outlines of a plausible motivation structure are developed in much greater detail in part 2 of my book.[16]

The value theory also suggests some significant changes in the research strategy for behavioral science. Assuming that the other higher vertebrates are motivated also by an innate value structure and that they utilize a mental model to make their choices, it suggests that behavioral research might be directed usefully toward understanding the similarities and differences in both the innate value structure and the quality of the mental model used by each species. The traditional approach in which the brain is treated theoretically only as a black box (stimulus in—behavior out) seems much less likely to yield useful new insights.

With regard to research on human behavior it suggests that the effort to exclude mentalistic concepts from behavioral theories is a mistake. The effort instead should be focused on the development of objective theories which naturally incorporate the mentalistic concepts. In the absence of a detailed neurological understanding of the brain that would allow us to measure objectively the intensity of the subjective valuative sensations, studies with human subjects using verbal reports of subjective sensations may provide some of the best clues for the development of a comprehensive motivational theory.

The theory also has a bearing on the analysis of genetic inheritance within the context of sobiobiology. A single value sensation or motivational mechanism typically will contribute to the motivation of many different types of behavior. Some of these behaviors may correspond to the main evolutionary "purpose," while others may be simply irrelevant "by-products" of the motivational mechanism. The assessment of whether a specific motivational mechanism contributes to inclusive fitness must take into account the total effect of the mechanism on all aspects of behavior. For example, the enjoyment of music could be simply a by-product of a behavioral mechanism whose main function is to motivate conversation and the learning of language. In such a case an assessment of the survival benefits of musical interest could be completely misleading. Thus a valid analysis of the genetic benefits of behavioral mechanisms should deal with the actual motivational mechanisms that are inherited and should take into account all the effects of each mechanism.

There recently has been a considerable amount of theoretical work in genetic statistics concerning the theoretical limits of behavioral altruism. It is important to recognize that these results are valid only so long as they are applied to the actual mechanisms that motivate behavioral altruism. The results can be very misleading if they are interpreted as applying to specific, isolated elements of behavior. In general the actual mechanisms of behavior are very imperfect in recognizing kinship relationships. The genetic value of an altruistic behavioral mechanism can be evaluated only when the probability distribution of kinship is known for the circumstances in which the mechanism actually might be stimulated. For example, a little girl's altruistic maternal instincts may be stimulated by the "cute" characteristics of a puppy. Rationally she has no difficulty in distinguishing the puppy from a human baby; nevertheless her behavior toward the puppy tends to be protective and altruistic. The genetic benefits of this behavioral mechanism, however, should be measured in terms of its overall effect on her behavior toward human babies as well as toward the puppy.

Potential Implications for Human Decision Making

The discussion of the value theory of motivation so far has focused only on the application of the theory within the traditional framework of a descriptive science. Obviously any scientific theory must survive first as a descriptive theory, and it will be many years before a scientific verdict can be rendered on the present theory.

Nevertheless, because many of the most interesting implications of the theory are concerned with potential applications to human decisions, it may be appropriate to consider how the theory may influence our approach to issues such as personal behavior and social planning. It appears that, at least to a first approximation, the theory should have no effect. First, it corresponds so closely to the commonsense intuition that most people use as a basis for practical decisions that one would expect the incorporation of the theory into the mental model of most people to have only minor effects on their actual behavior. Second, it is really nothing more than a descriptive theory of human motivation. Assuming that people are motivated fundamentally by their own internal valuative sensations, we would expect that they would continue to be so motivated. To the extent that the theory differs from these natural motivations for specific individuals it will be ignored, and to the extent that it corresponds it should have no effect.

With or without the theory we can expect that people (at least in small communities) will continue to admire and commend cooperative and altruistic behavior traits. We can expect that those who are most interested in social approval will continue to be motivated to conform to the social ideal for their community while those who are less motivated by social approval will continue to behave selfishly and will be deterred from antisocial behavior only by the expectation of formal law enforcement or community reprisals.

But this basic observation that a theory cannot change innate motivations can be made also about any religion or any cultural tradition. Thus the observation that the theory should have no effect on behavior is only a first-order approximation.

In practice cultural traditions and theoretical concepts do influence behavior not because they change basic motivations but because they affect the mental model of the environment and thus change the way people project and assess the probable outcomes for alternative courses of action. A scientifically validated theory of human motivation in principle can affect behavior in two basic ways. First, it allows people to make generally better predictions about the behavior of others, and in particular it enables them to make better guesses about how other people respond to different personal behavior choices. Second, it helps people make better guesses about how they themselves respond emotionally to alternative future outcomes. Specifically which courses of action and which outcomes are most likely to contribute positively to personal happiness, contentment, and a sense of achievement and personal pride?

I think there is reason to believe that such basic psychological knowledge can provide a scientific foundation from which some of the fundamentals of personal ethics can be developed and taught. Because the basic ethical concepts are not now incorporated in any generally accepted scientific theory, many intelligent young people are inclined to reject the concepts as superstitions which (like astrology, ghosts, and black magic) are to be ignored in any rational, human decision process. Obviously the effect of such scientific ethical knowledge would be not to change or distort basic motivations but to provide individuals with the knowledge necessary for making better decisions relative to their own personal system of values.

In terms of the development of public policy such a theory can be expected to have rather similar effects. By providing a better understanding of human social motivations it can help communities in structuring an environment within which the innate motivations can be used more effectively to support the necessary social restraints and inhibitions. The theory may also provide a useful service by helping society distinguish between those social restraints and inhibitions that continue to be relevant in the modern environment and those that are simply irrelevant inheritances from an earlier cultural environment.

Another question that can be raised is how the theory may affect the procedures for developing and adjusting social norms and social policy. Campbell discusses an approach which corresponds very closely to my views.[17] Basically he suggests that the appropriate process is to define a set of objectives that can be agreed on by the group and then to use an objective or scientific analysis of man's nature and the status of the world to derive suitable rules of behavior for the achievement of the chosen objectives. He notes that the resulting rules of behavior are normative in that they tell people how to behave, contingent on the assumed set of objectives and the validity of the scientific analysis of human nature and the state of the world. He cites as examples of such reasonable ultimate objectives the continued survival of the human race, avoidance of nuclear war, limitation of pollution of the environment, and avoidance of overpopulation. I do not believe that a theory of values can change in any substantial way this basic prescription for developing rules of behavior. However, if the theory is valid, it should permit a more accurate scientific assessment of the structure of human nature and the state of the world. In addition, a scientific theory of human values is likely to facilitate a better selection of realistic and appropriate objectives.

It is on this final point that the theory of values becomes most controversial and that the greatest conflict with prevailing philosophical and ethical systems is to be expected. Without addressing any of the troublesome philosophical issues let me consider briefly some of the reasons why such scientific assistance in the choice of objectives may be useful. The broad social objectives used by Campbell in his examples are all so obvious and intuitive that it seems unlikely that a scientific theory can be of much help in the selection of such objectives. In the case of such objectives the real problem is finding a practical way of achieving the objectives, not the selection of the objectives themselves.

However, there is a wide variety of social-policy problems in which the heart of the problem is the proper definition of the objectives. This particularly tends to be true when the essential problem involves an allocation of resources among competing objectives. But there are also a large number of situations in which we are really uncertain about what we "ought" to be trying to do. Many of these problems concern new social developments for which cultural tradition has not established any generally accepted guidance. For example, in the case of drugs whose main function is psychological, that is, the relief of pain or of emotional distress or the creation of sensations of temporary euphoria, what criteria should we use to decide which applications are desirable and which are undesirable? In the case of equality of educational opportunity how do we allocate resources among handicapped, normal, and specially gifted children? In the case of government programs whose main effect is the redistribution of income how do we decide what is desirable? In the case of school desegregation how should we balance the objectives of community pride, independence, and solidarity against the conflicting objective of educational exposure to a wide diversity of ideas and values? In the case of medical policy what are our objectives with regard to abortion, euthanasia, and suicide? These are but a small sample of important issues in social policy where the central problem is not so much the achievement as the appropriate selection of the objective. It seems likely that in many of these situations a more formal understanding of the origins of human values may provide useful insight in the selection of more appropriate objectives.

To understand the relationship between value theory and the traditional, human decision process we must consider the underlying mental processes through which people decide intuitively which social objectives to advocate or support. In a theoretical sense Campbell's description of the process is incomplete because it does not address the behavioral mechanisms that underlie the intuitive selection of objectives. According to value theory the value criteria that go into these selections (as in the case of all other decisions) can be traced ultimately to the individual's experience with the environment and with his own innate valuative sensations. It is from these sensations that he forms his objectives both for himself and for those he loves. And it is out of these objectives that he forms his policy preferences with regard to broad social issues such as avoidance of nuclear war, survival of the human race, avoidance of pollution and overpopulation, etc.

Obviously a philosophical question on whether that is how he "ought" to decide such issues can be raised. From the perspective of the theory of values the question is like asking whether the sun "ought" to set in the evening. The issue is simply not a matter for rational choice. Fundamentally the individual has no choice but to use his intuitive valuative criteria as his ultimate source of values, for that is how he is designed as a value-driven decision system. In specific cases he may select objectives well or poorly relative to these criteria, but he has no choice in what ultimate criteria to use.

It is my personal belief that objectives are often poorly chosen and that a formal theory of values which models the essential features of social choice can lead to better insight and substantially improved, social decision processes.

NOTES

1. George Edgin Pugh, *The Biological Origin of Human Values* (New York: Basic Books, 1977).
2. E. O. Wilson, *Sociobiology: The New Synthesis* (Cambridge, Mass.: Harvard University Press, 1975).
3. Aryeh Routtenberg, "The Reward System of the Brain," *Scientific American* 239 (November 1978): 154-64; Charles Laughlin, Jr., and Eugene G. d'Aquili, *Biogenetic Structuralism* (New York: Columbia University Press, 1974).
4. For a discussion of the concept of evolutionary stable strategies (ESS) see John Maynard Smith, "The Evolution of Behavior," *Scientific American* 239 (September 1978): 176-92, and Richard Dawkins, *The Selfish Gene* (London: Oxford University Press, 1976).
5. Donald T. Campbell, "On the Conflicts between Biological and Social Evolution and between Psychology and Moral Tradition," *American Psychologist* 30 (1975): 1103-26 (reprinted in *Zygon* 11 [1976]: 167-208).
6. David P. Barash, "Sociobiology: Evolution as a Paradigm for Behavior" (paper read at the California Symposium on Science and Values, San Francisco State University, June 14, 1977).
7. Abraham H. Maslow, *Motivation and Personality* (New York: Harper & Row, 1970); see also K. B. Madsen, *Modern Theories of Motivation* (New York: John Wiley & Sons, 1974).
8. E.g., see F. A. Hayek's otherwise excellent "The Three Sources of Human Values" (paper read at the London School of Economics and Political Science, May 17, 1978).
9. See n. 1 above.
10. See n. 4 above.
11. Of course these mechanisms, like any genetic inheritance, are imperfect and subject to genetic variability. Within any society one is likely to find specific instances of excessive altruism as well as individuals whose behavior seems almost totally selfish.
12. Campbell.
13. The decisive importance of such secondary values in chess is illustrated by chess-playing computer programs which often are so designed as to use such heuristic values as their sole criteria of decision. In such an automated system the heuristic values serve as the "primary values" for the decision system.
14. See n. 1 above.
15. Routtenberg.
16. See n. 1 above.
17. Campbell.