

BOUNDED RATIONALITY, AMBIGUITY, AND THE ENGINEERING OF CHOICE

JAMES G. MARCH

THE ENGINEERING OF CHOICE AND ORDINARY CHOICE BEHAVIOR

Recently I gave a lecture on elementary decision theory, an introduction to rational theories of choice. After the lecture, a student asked whether it was conceivable that the practical procedures for decision making implicit in theories of choice might make actual human decisions worse rather than better. What is the empirical evidence, he asked, that human choice is improved by knowledge of decision theory or by application of the various engineering forms of rational choice? I answered, I think correctly, that the case for the usefulness of decision engineering rested primarily not on the kind of direct empirical confirmation that he sought, but on two other things: on a set of theorems proving the superiority of particular procedures in particular situations if the situations are correctly specified and the procedures correctly applied, and on the willingness of clients to purchase the services of experts with skills in decision sciences.

The answer may not have been reasonable, but the question clearly was. It articulated a classical challenge to the practice of rational choice, the possibility that processes of rationality might combine with properties of human beings to produce decisions that are less sensible than the unsystematized actions of an intelligent person, or at least that the way in which we might use rational procedures intelligently is not self-evident. Camus (1951) argued, in effect, that man was not smart enough to be rational, a point made in a different way at about the same time by Herbert A. Simon (1957). Twenty years later, tales of horror have become contemporary clichés of studies of rational analysis in organizations (Wildavsky, 1971; Wildavsky and Pressman, 1973; Warwick, 1975).

I do not share the view of some of my colleagues that microeconomics, decision science, management science, operations analysis, and the other forms of rational decision engineering are mostly manufacturers of massive mischief when they are put into practice. It seems to me likely that these modern technologies of reason have, on balance, done more good than harm, and that

students of organizations, politics, and history have been overtly gleeful in their compilations of disasters. But I think there is good sense in asking how the practical implementation of theories of choice combines with the ways people behave when they make decisions, and whether our ideas about the engineering of choice might be improved by greater attention to our descriptions of choice behavior.

At first blush, pure models of rational choice seem obviously appropriate as guides to intelligent action, but more problematic for predicting behavior. In practice, the converse seems closer to the truth for much of economics. So long as we use individual choice models to predict the behavior of relatively large numbers of individuals or organizations, some potential problems are avoided by the familiar advantages of aggregation. Even a small signal stands out in a noisy message. On the other hand, if we choose to predict about small numbers of individuals or organizations or give advice to a single individual or organization, the saving graces of aggregation are mostly lost. The engineering of choice depends on a relatively close articulation between choice as it is comprehended in the assumptions of the model and choice as it is made comprehensible to individual actors.

This relation is reflected in the historical development of the field. According to conventional dogma, there are two kinds of theories of human behavior: descriptive (or behavioral) theories that purport to describe actual behavior of individuals or social institutions, and prescriptive (or normative) theories that purport to prescribe optimal behavior. In many ways, the distinction leads to an intelligent and fruitful division of labor in social science, reflecting differences in techniques, objectives, and professional cultures. For a variety of historical and intellectual reasons, however, such a division has not characterized the development of the theory of choice. Whether one considers ideas about choice in economics, psychology, political science, sociology, or philosophy, behavioral and normative theories have developed as a dialectic rather than as separate domains. Most modern behavioral theories of choice take as their starting point some simple ideas about rational human behavior. As a result, new developments in normative theories of choice have quickly affected behavioral theories. Contemplate, for example, the impact of game theory, statistical decision theory, and information theory on behavioral theories of human problem-solving, political decision making, bargaining, and organizational behavior (Rapoport, 1960; Vroom, 1964; Binkley, Bronaugh, and Marras, 1971; Tversky and Kahneman, 1974; Mayhew, 1974). It is equally obvious that prescriptive theories of choice have been affected by efforts to understand actual choice behavior. Engineers of artificial intelligence have modified their perceptions of efficient problem solving procedures by studying the actual behavior of human problem solvers (Simon, 1969; Newell and Simon, 1972). Engineers of organizational decision making have modified their models of rationality on the basis of studies of actual organizational behavior (Charnes and Cooper, 1963; Keen, 1977).

Modern students of human choice frequently assume, at least implicitly, that actual human choice behavior in some way or other is likely to make sense. It can be understood as being the behavior of an intelligent being or a group of intelligent beings. Much theoretical work searches for the intelligence in apparently anomalous human behavior. This process of discovering sense in human behavior is conservative with respect to the concept of rational man and to behavioral change. It preserves the axiom of rationality; and it preserves the idea that human behavior is intelligent, even when it is not obviously so. But it is not conservative with respect to prescriptive models of choice. For, if there is sense in the choice behavior of individuals acting contrary to standard engineering procedures for rationality, then it seems reasonable to suspect that there may be something inadequate about our normative theory of choice or the procedures by which it is implemented.

Rational choice involves two kinds of guesses: guesses about future consequence of current actions and guesses about future preference for those consequences (Savage, 1954; Thompson, 1967). We try to imagine what will happen in the future as a result of our actions and we try to imagine how we shall evaluate what will happen. Neither guess is necessarily easy. Anticipating future consequences of present decisions is often subject to substantial error. Anticipating future preferences is often confusing. Theories of rational choice are primarily theories of these two guesses and how we deal with their complications. Theories of choice under uncertainty emphasize the complications of guessing future consequences. Theories of choice under conflict or ambiguity emphasize the complications of guessing future preferences.

Students of decision making under uncertainty have identified a number of ways in which a classical model of how alternatives are assessed in terms of their consequences is neither descriptive of behavior nor a good guide in choice situations. As a result of these efforts, some of our ideas about how the first guess is made and how it ought to be made have changed. Since the early writings of Herbert A. Simon (1957), for example, bounded rationality has come to be recognized widely, though not universally, both as an accurate portrayal of much choice behavior and as a normatively sensible adjustment to the costs and character of information gathering and processing by human beings (Radner, 1975a, b; Radner and Rothschild, 1975; Connolly, 1977).

The second guess has been less considered. For the most part, theories of choice have assumed that future preferences are exogenous, stable, and known with adequate precision to make decisions unambiguous. The assumptions are obviously subject to question. In the case of collective decision making, there is the problem of conflicting objectives representing the values of different participants (March, 1962; Olson, 1965; M. Taylor, 1975; Pfeffer, 1977). In addition, individual preferences often appear to be fuzzy and inconsistent, and preferences appear to change over time, at least in part as a consequence of actions taken. Recently, some students of choice have been examining the ways individuals and organizations confront the second guess under conditions of

ambiguity (i.e., where goals are vague, problematic, inconsistent, or unstable) (Cohen and March, 1974; Weick, 1976; March and Olsen, 1976; Crozier and Friedberg, 1977). Those efforts are fragmentary, but they suggest that ignoring the ambiguities involved in guessing future preferences leads both to misinterpreting choice behavior and to mistating the normative problem facing a decision maker. The doubts are not novel; John Stuart Mill (1838) expressed many of them in his essay on Bentham. They are not devastating; the theory of choice is probably robust enough to cope with them. They are not esoteric; Hegel (1832) is relevant, but may not be absolutely essential.

BOUNDED RATIONALITY

There is a history. A little over twenty years ago, Simon published two papers that became a basis for two decades of development in the theory of choice (1955, 1956). The first of these examined the informational and computational limits on rationality by human beings. The paper suggested a focus on step-function utility functions and a process of information gathering that began with a desired outcome and worked back to a set of antecedent actions sufficient to produce it. The second paper explored the consequences of simple payoff functions and search rules in an uncertain environment. The two papers argued explicitly that descriptions of decision making in terms of such ideas conformed more to actual human behavior than did descriptions built upon classical rationality, that available evidence designed to test such models against classical ones tended to support the alternative ideas.

Because subsequent developments were extensive, it is well to recall that the original argument was a narrow one. It started from the proposition that all intendedly rational behavior is behavior within constraints. Simon added the idea that the list of technical constraints on choice should include some properties of human beings as processors of information and as problem solvers. The limitations were limitations of computational capability, the organization and utilization of memory, and the like. He suggested that human beings develop decision procedures that are sensible, given the constraints, even though they might not be sensible if the constraints were removed. As a shorthand label for such procedures, he coined the term "satisficing."

Developments in the field over the past twenty years have expanded and distorted Simon's original formulation. But they have retained some considerable flavor of his original tone. He emphasized the theoretical difficulty posed by self-evident empirical truths. He obscured a distinction one might make between individual and organizational decision making, proposing for the most part the same general ideas for both. He obscured a possible distinction between behavioral and normative theories of choice, preferring to view differences between perfect rationality and bounded rationality as explicable consequences of constraints. Few of the individual scholars who followed had precisely the same interests or commitments as Simon, but the field has generally maintained

the same tone. Theoretical puzzlement with respect to the simplicity of decision behavior has been extended to puzzlement with respect to decision inconsistencies and instabilities, and the extent to which individuals and organizations do things without apparent reason (March and Olsen, 1976). Recent books on decision making move freely from studies of organizations to studies of individuals (Janis and Mann, 1977). And recent books on normative decision making accept many standard forms of organizational behavior as sensible (Keen, 1977).

Twenty years later, it is clear that we do not have a single, widely accepted, precise behavioral theory of choice. But I think it can be argued that the empirical and theoretical efforts of the past twenty years have brought us closer to understanding decision processes. The understanding is organized in a set of conceptual vignettes rather than a single, coherent structure; and the connections among the vignettes are tenuous. In effect, the effort has identified major aspects of some key processes that appear to be reflected in decision making; but the ecology of those processes is not well captured by any current theory. For much of this development, Simon bears substantial intellectual responsibility.

Simon's contributions have been honored by subsumption, extension, elaboration, and transformation. Some writers have felt it important to show that aspiration level goals and goal-directed search can be viewed as special cases of other ideas, most commonly classical notions about rational behavior (Riker and Ordeshook, 1974). Others have taken ideas about individual human behavior and extended them to organizations (both business firms and public bureaucracies) and to other institutions, for example, universities (Bower, 1968; Allison, 1969; Steinbruner, 1974; Williamson, 1975). Simon's original precise commentary on specific difficulties in rational models has been expanded to a more general consideration of problems in the assumptions of rationality, particularly the problems of subjective understanding, perception, and conflict of interest (Cyert and March, 1963; Porat and Haas, 1969; Carter, 1971; R. N. Taylor, 1975; Slovic, Fischhoff, and Lichtenstein, 1977). The original articles suggested small modifications in a theory of economic behavior, the substitution of bounded rationality for omniscient rationality. But the ideas ultimately have led to an examination of the extent to which theories of choice might subordinate the idea of rationality altogether to less intentional conceptions of the causal determinants of action (March and Olsen, 1976).

ALTERNATIVE RATIONALITIES

The search for intelligence in decision making is an effort to rationalize apparent anomalies in behavior. In a general way, that effort imputes either calculated or systemic rationality to observed choice behavior. Action is presumed to follow either from explicit calculation of its consequences in terms of objectives, or from rules of behavior that have evolved through processes that are sensible but which

obscure from present knowledge full information on the rational justification for any specific rule.

Most efforts to rationalize observed behavior have attempted to place that behavior within a framework of calculated rationality. The usual argument is that a naive rational model is inadequate either because it focuses on the wrong unit of analysis, or because it uses an inaccurate characterization of the preferences involved. As a result, we have developed ideas of limited rationality, contextual rationality, game rationality, and process rationality.

Ideas of *limited rationality* emphasize the extent to which individuals and groups simplify a decision problem because of the difficulties of anticipating or considering all alternatives and all information (March and Simon, 1958; Lindblom, 1959; 1965; Radner, 1975a, b). They introduce, as reasonable responses, such things as step-function tastes, simple search rules, working backward, organizational slack, incrementalism and muddling through, uncertainty avoidance, and the host of elaborations of such ideas that are familiar to students of organizational choice and human problem solving.

Ideas of *contextual rationality* emphasize the extent to which choice behavior is embedded in a complex of other claims on the attention of actors and other structures of social and cognitive relations (Long, 1958; Schelling, 1971; Cohen, March, and Olsen, 1972; Weiner, 1976; Sproull, Weiner, and Wolf, 1978). They focus on the way in which choice behavior in a particular situation is affected by the opportunity costs of attending to that situation and by the apparent tendency for people, problems, solutions, and choices to be joined by the relatively arbitrary accidents of their simultaneity rather than by their *prima facie* relevance to each other.

Ideas of *game rationality* emphasize the extent to which organizations and other social institutions consist of individuals who act in relation to each other intelligently to pursue individual objectives by means of individual calculations of self-interest (Farquharson, 1969; Harsanyi and Selten, 1972; Brams, 1975). The decision outcomes of the collectivity in some sense amalgamate those calculations, but they do so without imputing a super-goal to the collectivity or invoking collective rationality. These theories find reason in the process of coalition formation, sequential attention to goals, information bias, and interpersonal gaming, and the development of mutual incentives.

Ideas of *process rationality* emphasize the extent to which decisions find their sense in attributes of the decision process, rather than in attributes of decision outcomes (Edelman, 1960; Cohen and March, 1974; Kreiner, 1976; Christensen, 1976). They explore those significant human pleasures (and pains) found in the ways we act while making decisions, and in the symbolic content of the idea and procedures of choice. Explicit outcomes are viewed as secondary and decision making becomes sensible through the intelligence of the way it is orchestrated.

All of these kinds of ideas are theories of intelligent individuals making calculations of the consequences of actions for objectives, and acting sensibly to achieve those objectives. Action is presumed to be consequential, to be connected

consciously and meaningfully to knowledge about personal goals and future outcomes, to be controlled by personal intention.

Although models of calculated rationality continue to be a dominant style, students of choice have also shown considerable interest in a quite different kind of intelligence, systemic rather than calculated. Suppose we imagine that knowledge, in the form of precepts of behavior, evolves over time within a system and accumulates across time, people, and organizations without complete current consciousness of its history. The sensible action is taken by actors without comprehension of its full justification. This characterizes models of adaptive rationality, selected rationality, and posterior rationality.

Ideas of *adaptive rationality* emphasize experiential learning by individuals or collectivities (Cyert and March, 1963; Day and Groves, 1975). Most adaptive models have the property that, if the world and preferences are stable and the experience prolonged enough, behavior will approach the behavior that would be chosen rationally on the basis of perfect knowledge. Moreover, the postulated learning functions normally have properties that permit sensible adaptation to drifts in environmental or taste attributes. By strong information on past experiences in some simple behavioral predictions, adaptive rationality permits the efficient management of considerable experiential information; but it is in a form that is not explicitly retrievable — particularly across individuals or long periods of time. As a result, it is a form of intelligence that tends to separate current reasons from current actions.

Ideas of *selected rationality* emphasize the process of selection among individuals or organizations through survival or growth (Winter, 1964, 1971, 1975; Nelson and Winter, 1973). Rules of behavior achieve intelligence not by virtue of conscious calculation of their rationality by current role players but by virtue of the survival and growth of social institutions in which such rules are followed and such roles are performed. Selection theories focus on the extent to which choice is dominated by standard operating procedures and the social regulation of social roles.

Ideas of *posterior rationality* emphasize the discovery of intentions as an interpretation of action rather than as a prior position (Hirschman, 1967; Weick, 1969; March, 1973). Actions are seen as being exogenous and as producing experiences that are organized into an evaluation after the fact. The valuation is in terms of preferences generated by the action and its consequences, and choices are justified by virtue of their posterior consistency with goals that have themselves been developed through a critical interpretation of the choice. Posterior rationality models maintain the idea that action should be consistent with preferences, but they conceive action as being antecedent to goals.

These explorations into elements of systemic rationality have, of course, a strong base in economics and behavioral science (Wilson, 1975; Becker, 1976); but they pose special problems for decision engineering. On the one hand, systemic rationality is not intentional. That is, behavior is not understood as following from a calculation of consequences in terms of prior objectives. If such a

calculation is asserted, it is assumed to be an interpretation of the behavior but not a good predictor of it. On the other hand, these models claim, often explicitly, that there is intelligence in the suspension of calculation. Alternatively, they suggest that whatever sense there is in calculated rationality is attested not by its formal properties but by its survival as a social rule of behavior, or as an experientially verified personal propensity.

In a general way, these explanations of ordinary behavior as forms of rationality have considerably clarified and extended our understanding of choice. It is now routine to explore aspects of limited, contextual, game, process, adaptive, selected, and posterior rationality in the behavioral theory of choice. We use such ideas to discover and celebrate the intelligence of human behavior. At the same time, however, this discovery of intelligence in the ordinary behavior of individuals and social institutions is an implicit pressure for reconstruction of normative theories of choice, for much of the argument is not only that observed behavior is understandable as a human phenomenon, but that it is, in some important sense, intelligent. If behavior that apparently deviates from standard procedures of calculated rationality can be shown to be intelligent, then it can plausibly be argued that models of calculated rationality are deficient not only as descriptors of human behavior but also as guides to intelligent choice.

THE TREATMENT OF TASTES

Engineers of intelligent choice sensibly resist the imputation of intelligence to all human behavior. Traditionally, deviations of choice behavior from the style anticipated in classical models were treated normatively as errors, or correctable faults, as indeed many of them doubtless were. The objective was to transform subjective rationality into objective rationality by removing the needless informational, procedural, and judgmental constraints that limited the effectiveness of persons proceeding intelligently from false or incomplete informational premises (Ackoff and Sasieni, 1968). One of Simon's contributions to the theory of choice was his challenge of the self-evident proposition that choice behavior necessarily would be improved if it were made more like the normative model of rational choice. By asserting that certain limits on rationality stemmed from properties of the human organism, he emphasized the possibility that actual human choice behavior was more intelligent than it appeared.

Normative theories of choice have responded to the idea. Substantial parts of the economics of information and the economics of attention (or time) are tributes to the proposition that information gathering, information processing, and decision making impose demands on the scarce resources of a finite capacity human organism (Stigler, 1961; Becker, 1965; McGuire and Radner, 1972; Marschak and Radner, 1972; Rothschild and Stiglitz, 1976). Aspiration levels, signals, incrementalism, and satisficing rules for decision making have been described as sensible under fairly general circumstances (Hirschman and Lindblom, 1962; Spence, 1974; Radner, 1975a, b; Radner and Rothschild, 1975).

These developments in the theory of rational choice acknowledge important aspects of the behavioral critique of classical procedures for guessing the future consequences of present action. Normative response to behavioral discussions of the second guess, the estimation of future preferences, has been similarly conservative but perceptible. That standard theories of choice and the engineering procedures associated with them have a conception of preferences that differs from observations of preferences has long been noted (Johnson, 1968). As in the case of the informational constraints on rational choice, the first reaction within decision engineering was to treat deviations from well-defined, consistent preference functions as correctable faults. If individuals had deficient (i.e., inconsistent, incomplete) preference functions, they were to be induced to generate proper ones, perhaps through revealed preference techniques and education. If groups of organizations exhibited conflict, they were to be induced to resolve that conflict through prior discussion, prior side payments (e.g., an employment contract), or prior bargaining. If individuals or organizations exhibited instability in preferences over time, they were to be induced to minimize that instability by recognizing a more general specification of the preferences so that apparent changes became explicable as reflecting a single, unchanging function under changing conditions or changing resources.

Since the specific values involved in decision making are irrelevant to formal models of choice, both process rationality and contextual rationality are, from such a perspective, versions of simple calculated rationality. The criterion function is changed, but the theory treats the criterion function as any arbitrary set of well-ordered preferences. So long as the preferences associated with the process of choice or the preferences involved in the broader context are well defined and well behaved, there is no deep theoretical difficulty. But, in practice, such elements of human preference functions have not filtered significantly into the engineering of choice.

The record with respect to problems of goal conflict, multiple, lexicographic goals, and loosely coupled systems is similar. Students of bureaucracies have argued that a normative theory of choice within a modern bureaucratic structure must recognize explicitly the continuing conflict in preferences among various actors (Tullock, 1965; Downs, 1967; Allison and Halperin, 1972; Halperin, 1974). Within such systems "decisions" are probably better seen as strategic first-move interventions in a dynamic internal system than as choices in a classical sense. Decisions are not expected to be implemented, and actions that would be optimal if implemented are suboptimal as first moves. This links theories of choice to game-theoretic conceptions of politics, bargaining, and strategic actions in a productive way. Although in this way ideas about strategic choice in collectivities involving conflict of interest are well established in part of the choice of literature (Elster, 1977a), they have had little impact on such obvious applied domains as bureaucratic decision making or the design of organizational control systems. The engineering of choice has been explicitly concerned with multiple criteria decision procedures for dealing with multiple, lexicographic, or political

goals (Lee, 1972; Pattanaik, 1973). In some cases these efforts have considerably changed the spirit of decision analysis, moving it toward a role of exploring the implications of constraints and away from a conception of solution.

Behavioral inquiry into preferences has, however, gone beyond the problems of interpersonal conflict of interest in recent years and into the complications of ambiguity. The problems of ambiguity are partly problems of disagreement about goals among individuals, but they are more conspicuously problems of the relevance, priority, clarity, coherence, and stability of goals in both individual and organizational choice. Several recent treatments of organizational choice behavior record some major ways in which explicit goals seem neither particularly powerful predictors of outcomes nor particularly well represented as either stable, consistent preference orders or well-defined political constraints (Cohen and March, 1974; Weick, 1976; March and Olsen, 1976; Spruill, Weiner, and Wolf, 1978).

It is possible, of course, that such portrayals of behavior are perverse. They may be perverse because they systematically misrepresent the actual behavior of human beings or they may be perverse because the human beings they describe are, insofar as the description applies, stupid. But it is also possible that the description is accurate and the behavior is intelligent, that the ambiguous way human beings sometimes deal with tastes is, in fact, sensible. If such a thing can be imagined, then its corollary may also be imaginable: perhaps we treat tastes inadequately in our engineering of choice. When we start to discover intelligence in decision making where goals are unstable, ill-defined, or apparently irrelevant, we are led to asking some different kinds of questions about our normative conceptions of choice and walk close not only to some issues in economics but also to some classical and modern questions in literature and ethics, particularly the role of clear prior purpose in the ordering of human affairs.

Consider the following properties of tastes as they appear in standard prescriptive theories of choice:

Tastes are *absolute*. Normative theories of choice assume a formal posture of moral relativism. The theories insist on morality of action in terms of tastes; but they recognize neither discrimination among alternative tastes, nor the possibility that a person reasonably might view his own preferences and actions based on them as normally distressing.

Tastes are *relevant*. Normative theories of choice require that action be taken in terms of tastes, that decisions be consistent with preferences in the light of information about the probable consequences of alternatives for valued outcomes. Action is willful.

Tastes are *stable*. With few exceptions, normative theories of choice require that tastes be stable. Current action is taken in terms of current tastes. The implicit assumption is that tastes will be unchanged when the outcomes of current actions are realized.

Tastes are *consistent*. Normative theories of choice allow mutually

inconsistent tastes only insofar as they can be made irrelevant by the absence of scarcity or reconcilable by the specification of trade-offs.

Tastes are *precise*. Normative theories of choice eliminate ambiguity about the extent to which a particular outcome will satisfy tastes, at least insofar as possible resolutions of that ambiguity might affect the choice.

Tastes are *exogenous*. Normative theories of choice presume that tastes, by whatever process they may be created, are not themselves affected by the choices they control.

Each of these features of tastes seems inconsistent with observations of choice behavior among individuals and social institutions. Not always, but often enough to be troublesome. Individuals commonly find it possible to express both a taste for something and a recognition that the taste is something that is repugnant to moral standards they accept. Choices are often made without respect to tastes. Human decision makers routinely ignore their own, fully conscious, preferences in making decisions. They follow rules, traditions, hunches, and the advice or actions of others. Tastes change over time in such a way that predicting future tastes is often difficult. Tastes are inconsistent. Individuals and organizations are aware of the extent to which some of their preferences conflict with other of their preferences; yet they do nothing to resolve those inconsistencies. Many preferences are stated in forms that lack precision. It is difficult to make them reliably operational in evaluating possible outcomes. While tastes are used to choose among actions, it is often also true that actions and experience with their consequences affect tastes. Tastes are determined partly endogenously.

Such differences between tastes as they are portrayed by or models and tastes as they appear in our experience produce ordinary behavioral phenomena that are not always well accommodated within the structure of our prescriptions.

We manage our preferences. We select actions now partly in terms of expectations about the effect of those actions upon future preferences. We do things now to modify our future tastes. Thus, we know that if we engage in some particularly tasty, but immoral, activity, we are likely to come to like it more. We know that if we develop competence in a particular skill, we shall often come to favor it. So we choose to pursue the competence, or not, engage in an activity, or not, depending on whether we wish to increase or decrease our taste for the competence or activity.

We construct our preferences. We choose preferences and actions jointly, in part, to discover - or construct - new preferences that are currently unknown. We deliberately specify our objectives in vague terms to develop an understanding of what we might like to become. We elaborate our tastes as interpretations of our behavior.

We treat our preferences strategically. We specify goals that are different from the outcomes we wish to achieve. We adopt preferences and rules of actions that if followed literally would lead us to outcomes we do not wish, because we believe

that the final outcome will only partly reflect our initial intentions. In effect, we consider the choice of preferences as part of an infinite game with ourselves in which we attempt to deal with our propensities for acting badly by anticipating them and outsmarting ourselves. We use deadlines and make commitments.

We confound our preferences. Our deepest preferences tend often to be paired. We find the same outcome both attractive and repulsive, not in the sense that the two sentiments cancel each other and we remain indifferent, but precisely that we simultaneously want and do not want an outcome, experience it as both pleasure and pain, love and hate it (Catullus, 58 BC, 1.1).

We avoid our preferences. Our actions and our preferences are only partly linked. We are prepared to say that we want something, yet should not want it, or wish we did not want it. We are prepared to act in ways that are inconsistent with our preferences, and to maintain that inconsistency in the face of having it demonstrated. We do not believe that what we do must necessarily result from a desire to achieve preferred outcomes.

We expect change in our preferences. As we contemplate making choices that have consequences in the future, we know that our attitudes about possible outcomes will change in ways that are substantial but not entirely predictable. The subjective probability distribution over possible future preferences (like the subjective probability distribution over possible future consequences) increases as its variance as the horizon is stretched. As a result, we have a tendency to want to take actions now that maintain future options for acting when future preferences are clearer.

We suppress our preferences. Consequential argument, the explicit linking of actions to desires, is a form of argument in which some people are better than others. Individuals who are less competent at consequential rationalization try to avoid it with others who are more competent, particularly others who may have a stake in persuading them to act in a particular way. We resist an explicit formulation of consistent desires to avoid manipulation of our choices by persons cleverer than we at that special form of argument called consistent rationality.

It is possible, on considering this set of contrasts between decision making as we think it ought to occur and decision making as we think it does occur to trivialize the issue into a "definitional problem." By suitably manipulating the concept of tastes, one can save classical theories of choice as "explanations" of behavior in a formal sense, but probably only at the cost of stretching a good idea into a doubtful ideology (Stigler and Becker, 1977). More importantly from the present point of view, such a redefinition pays the cost of destroying the practical relevance of normative prescriptions for choice. For prescriptions are useful only if we see a difference between observed procedures and desirable procedures.

Alternatively, one can record all of the deviations from normative specifications as stupidity, errors that should be corrected, and undertake to transform the style of existing humans into the styles anticipated by the theory.

This has, for the most part, been the strategy of operations and management analysis for the past twenty years; and it has had its successes. But it has also had failures.

It is clear that the human behavior I have described may, in any individual case, be a symptom of ignorance, obtuseness, or deviousness. But the fact that such patterns of behavior are fairly common among individuals and institutions suggests that they might be sensible under some general kinds of conditions — that global ambiguity, like limited rationality, is not necessarily a fault in human choice to be corrected but often a form of intelligence to be refined by the technology of choice rather than ignored by it.

Uncertainty about future consequences and human limitations in dealing with them are relatively easily seen as intrinsic in the decision situation and the nature of the human organisms. It is much harder to see in what way ambiguous preferences are a necessary property of human behavior. It seems meaningful in ordinary terms to assert that human decision makers are driven to techniques of limited rationality by the exigencies of the situation in which they find themselves. But what drives them to ambiguous and changing goals? Part of the answer is directly analogous to the formulations of limited rationality. Limitations of memory organization and retrieval and of information capacity affect information processing about preferences just as they affect information processing about consequences (March and Simon, 1958; Cyert and March, 1963; Simon, 1973; March and Romelcer, 1976). Human beings have unstable, inconsistent, incompletely evoked, and imprecise goals at least in part because human abilities limit preference orderliness. If it were possible to be different at reasonable cost, we probably would want to be.

But viewing ambiguity as a necessary cost imposed by the information processing attributes of individuals fails to capture the extent to which similar styles in preferences would be sensible, even if the human organism were a more powerful computational system. We probably need to ask the more general question: Why might a person or institution intelligently choose to have ambiguous tastes? The answer, I believe, lies in several things, some related to ideas of bounded rationality, others more familiar to human understanding as it is portrayed in literature and philosophy than to our theories of choice.

First, human beings recognize in their behavior that there are limits to personal and institutional integration in tastes. They know that, no matter how much they may be pressured both by their own prejudices for integration and by the demands of others, they will be left with contradictory and intermittent desires partially ordered but imperfectly reconciled. As a result, they engage in activities designed to manage preferences or game preferences. These activities make little sense from the point of view of a conception of human choice that assumes people know what they want and will want, or a conception that assumes wants are morally equivalent. But ordinary human actors sense that they might come to want something that they should not, or that they might

make unwise or inappropriate choices under the influence of fleeting, but powerful desires if they do not act now either to control the development of tastes or to buffer action from tastes (Elster, 1977b).

Secondly, human beings recognize implicitly the limitations of acting rationally on current guesses. By insisting that action, to be justified, must follow preferences and be consistent both with those preferences and with estimates of future states, we considerably exaggerate the relative power of a choice based consistently upon two guesses compared to a choice that is itself a guess. Human beings are both proponents for preferences and observers of the process by which their preferences are developed and acted upon. As observers of the process by which their beliefs have been formed and consulted, they recognize the good sense in perceptual and moral modesty (Williams, 1973; Elster, 1977c).

Thirdly, human beings recognize the extent to which tastes are constructed, or developed, through a more or less constant confrontation between preferences and actions that are inconsistent with them, and among conflicting preferences. As a result, they appear to be comfortable with an extraordinary array of unreconciled sources of legitimate wants. They maintain a lack of coherence both within and among personal desires, social demands, and moral codes. Though they seek some consistency, they appear to see inconsistency as a normal, and necessary, aspect of the development and clarification of tastes (March, 1973).

Fourthly, human beings are conscious of the importance of preferences as beliefs independent of their immediate action consequences. They appear to find it possible to say, in effect, that they believe something is more important to good action than they are able (or willing) to make it in a specific case. They act as though some aspects of their beliefs are important to life without necessarily being consistent with actions, and important to the long-run quality of choice behavior without controlling it completely in the short run. They accept a degree of personal and social wisdom in ordinary hypocrisy (Chomsky, 1968; March, 1973; Pondy and Olson, 1977).

Fifthly, human beings know that some people are better at rational argument than others, and that those skills are not particularly well correlated with either morality or sympathy. As a result, they recognize the political nature of argumentation more clearly, and more personally, than the theory of choice does. They are unwilling to gamble that God made clever people uniquely virtuous. They protect themselves from cleverness by obscuring the nature of their preferences: they exploit cleverness by asking others to construct reasons for actions they wish to take.

TASTES AND THE ENGINEERING OF CHOICE

These characteristics of preference processing by individual human beings and social institutions seem to me to make sense under rather general circumstances. As a result, it seems likely to me that our engineering of choice behavior does not

make so much sense as we sometimes attribute to it. The view of human tastes and their proper role in action that we exhibit in our normative theory of choice is at least as limiting to the engineering applicability of that theory as the perfect knowledge assumptions were to the original formulations.

Since it has taken us over twenty years to introduce modest elements of bounded rationality and conflict of interest into prescriptions about decision making, there is no particular reason to be sanguine about the speed with which our engineering of choice will accept and refine the intelligence of ambiguity. But there is hope. The reconstruction involved is not extraordinary, and in some respects has already begun. For the doubts I have expressed about engineering models of choice to be translated into significance changes, they will have to be formulated a bit more precisely in terms that are comprehensible within such theories, even though they may not be consistent with the present form of the theories or the questions the theories currently address. I cannot accomplish such a task in any kind of complete way, but I think it is possible to identify a few conceptual problems that might plausibly be addressed by choice theorists and a few optimization problems that might plausibly be addressed by choice engineers.

The conceptual problems involve discovering interesting ways to reformulate some assumptions about tastes, particularly about the stability of tastes, their exogenous character, their priority, and their internal consistency.

Consider the problem of *intertemporal comparison* of preferences (Strotz, 1956; Koopmans, 1964; Bailey and Olson, 1977; Sherrin and Thaler, 1977). Suppose we assume that the preferences that will be held at every relevant future point in time are known. Suppose further that those preferences change over time but are, at any given time, consistent. If action is to be taken now in terms of its consequences over a period of time during which preferences change, we are faced with having to make intertemporal comparisons. As long as the changes are exogenous, we can avoid the problem if we choose to do so. If we can imagine an individual making a complete and transitive ordering over possible outcomes over time, then intertemporal comparisons are implicit in the preference orderings and cause no particular difficulty beyond the heroic character of the assumption about human capabilities. If, on the other hand, we think of the individual as having a distinct, complete, and consistent preference relation defined over the outcomes realized in a particular time period, and we imagine that those preferences change over time, then the problem of intertemporal comparisons is more difficult. The problem is technically indistinguishable from the problem of interpersonal comparison of utilities. When we compare the changing preferences of a single person over time to make trade-offs across time, we are in the identical position to when we attempt to make comparisons across different individuals at a point in time. The fact that the problems are identical has the advantage of immediately bringing to bear on the problems of intertemporal comparisons the apparatus developed to deal with interpersonal comparisons (Muller, 1976). It has the disadvantage that that apparatus allows

a much weaker conception of solution than is possible within a single, unchanging set of preferences. We are left with the weak theorems of social welfare economics, but perhaps with a clearer recognition that there is no easy and useful way to escape the problem of incomparable preference functions by limiting our attention to a single individual, as long as tastes change over time and we think of tastes as being defined at a point in time.

Consider the problem of *endogenous change* in preferences (Von Weizsäcker, 1971; Olson, 1976). Suppose we know that future tastes will change in a predictable way as a consequence of actions taken now and the consequences of those actions realized over time. Then we are in the position of choosing now the preferences we shall have later. If there is risk involved, we are choosing now a probability distribution over future preferences. If we can imagine some "super goal," the problem becomes tractable. We evaluate alternative preferences in terms of their costs and benefits for the "super goal." Such a strategy preserves the main spirit of normal choice theory but allows only a modest extension into endogenous change. This is the essential strategy adopted in some of the engineering examples below. In such cases desirable preferences cannot always be deduced from the "super goal," but alternative preferences can be evaluated. In somewhat the same spirit, we can imagine adaptive preferences as a possible decision procedure and examine whether rules for a sequence of adaptations in tastes can be specified that lead to choice outcomes better in some easily recognized sense than those obtained through explicit calculated rationality at the start of the process. One possible place is the search for cooperative solutions in games in which calculated rationality is likely to lead to outcomes desired to no one (Cyert and de Groot, 1973, 1975). Also in the same general spirit, we might accept the strict morality position and attempt to select a strategy for choice that will minimize change in values. Or we might try to select a strategy that maximizes value change. All of these are possible explorations, but they are not fully attentive to the normative management of adaptation in tastes. The problem exceeds our present concepts: how do we act sensibly now to manage the development of preferences in the future when we do not have now a criterion for evaluating future tastes that will not itself be affected by our actions? There may be some kind of fixed-point theorem answer to such a problem, but I suspect that a real conceptual confrontation with endogenous preferences will involve some reintroduction of moral philosophy into our understanding of choice (Freidman, 1967; Williams, 1973; Beck, 1975).

Consider the problem of *posterior preferences* (Schutz, 1967; Hirschman, 1967; Weick, 1969; Elster, 1976). The theory of choice is built on the idea of prior intentions. Suppose we relax the requirement of priority, allow preferences to rationalize action after the fact in our theories as well as our behavior. How do we act in such a way that we conclude, after the fact, that the action was intelligent, and also are led to an elaboration of our preferences that we find fruitful? Such a formulation seems closer to a correct representation of choice problems in politics, for example, than is conventional social welfare theory. We find meaning

and merit in our actions after they are taken and the consequences are observed and interpreted. Deliberate efforts to manage posterior constructions of preferences are familiar to us. They include many elements of child rearing, psychotherapy, consciousness raising, and product advertising. The terms are somewhat different. We talk of development of character in child rearing, of insight in psychotherapy, of recognition of objective reality in political ethnic, or sexual consciousness raising, and of elaboration of personal needs in advertising. But the technologies are more similar than their ideologies. These techniques for the construction (or excavation) of tastes include both encouraging a reinterpretation of experience and attempting to induce current behavior that will facilitate posterior elaboration of a new understanding of personal preferences. I have tried elsewhere to indicate some of the possibilities this suggests for intelligent foolishness and the role of ambiguity in sensible action (March, 1973, 1977). The problem is in many ways indistinguishable from the problem of poetry and the criticism of poetry (or art and art criticism). The poet attempts to write a poem that has meanings intrinsic in the poem but not necessarily explicit at the moment of composition (Clardi, 1960). In this sense, at least, decisions, like poems, are open; and good decisions are those that enrich our preferences and their meanings. But to talk in such a manner is to talk the language of criticism and aesthetics, and it will probably be necessary for choice theory to engage that literature in some way (Eliot, 1933; Cavell, 1969; Steinberg, 1972; Rosenberg, 1975).

Finally, consider the problem of *inconsistency* in preferences (Elster, 1977c). From the point of view of ordinary human ideas about choice, as well as many philosophical and behavioral conceptions of choice, the most surprising thing about formal theories of choice is the tendency to treat such terms as values, goals, preferences, tastes, wants, and the like as either equivalent or as reducible to a single objective function with properties of completeness and consistency. Suppose that instead of making such an assumption, we viewed the decision maker as confronted simultaneously with several orderings of outcomes. We could give them names, calling one a moral code, another a social role, another a personal taste, or whatever. From the present point of view what would be critical would be that the several orderings were independent and irreducible. That is, they could not be deduced from each other, and they could not be combined into a single order. Then, instead of taking the conventional step of imputing a preference order across these incomparables by some kind of revealed preference procedures, we treat them as truly incomparable and examine solutions to internal inconsistency that are more in the spirit of our efforts to provide intelligent guidance to collectivities in which we accept the incomparability of preferences across individuals. Then we could give better advice to individuals who want to treat their own preferences strategically, and perhaps move to a clearer recognition of the role of contradiction and paradox in human choice (Farber, 1976; Elster, 1977c). The strategic problems are amenable to relatively straightforward modifications of our views of choice under conflict

of interest; the other problems probably require a deeper understanding of contradiction as it appears in philosophy and literature (Elsner, 1977c).

Formulating the conceptual problems in these ways is deliberately conservative *vis-à-vis* the theory of choice. It assumes that thinking about human behavior in terms of choice on the basis of some conception of intention is useful, and that the tradition of struggle between normative theories of choice and behavioral theories of choice is a fruitful one. There are alternative paradigms for understanding human behavior that are in many situations likely to be more illuminating. But it is probably unwise to think that every paper should suggest a dramatic paradigm shift, particularly when the alternative is seen only dimly.

Such strictures become even more important when we turn to the engineering of choice. Choice theorists have often discussed complications in the usual abstract representation of tastes. But those concerns have had little impact on ideas about the engineering of choice, perhaps because they pose the problems at a level of philosophic complexity that is remote from decision engineering. Thus, although I think the challenges that ambiguity makes to our models of choice are rather fundamental, my engineering instincts are to sacrifice purity to secure tractability. I suspect we should ask the engineers of choice not initially to reconstruct a philosophy of tastes but to reexamine, within a familiar framework, some presumptions of our craft, and to try to make the use of ambiguity somewhat less of a mystery, somewhat more of a technology. Consider, for example, the following elementary problems in engineering.

The optimal ambition problem. The level of personal ambition is not a decision variable in most theories of choice; but, as a result of the work by Simon and others on satisficing, there has been interest in optimal levels of aspiration. These efforts consider an aspiration level as a trigger that either begins or ends the search for new alternatives. The optimization problem is one of balancing the expected costs of additional search with the expected improvements to be realized from the effort (March and Simon, 1958).

But there is another, rather different, way of looking at the optimum ambition problem. Individuals and organizations form aspirations, goals, targets, or ambitions for achievement. These ambitions are usually assumed to be connected to outcomes in at least two ways: they affect search (either directly or through some variable like motivation) and thereby performance; they affect (jointly with performance) satisfaction (March and Simon, 1958). Suppose we wish to maximize some function of satisfaction over time by selecting among alternative ambitions over time, alternative initial ambitions, or alternatives defined by some other decision variable that affects ambition. Examples of the latter might be division of income between consumption and savings, tax policies, or choice among alternative payment schemes. In effect, we wish to select a preference function for achievement that will, after the various behavioral consequences of that selection are accounted for, make us feel that we have selected the best ambition. It is a problem much more familiar to the real world of

personal and institutional choice than it is to the normative theory of choice, but it is something about which some things could be said.

The optimal clarity problem. Conventional notions about intelligent choice often begin with the presumption that good decisions require clear goals, and that improving the clarity of goals unambiguously improves the quality of decision making. In fact, greater precision in the statement of objectives and the measurement of performance with respect to them is often a mixed blessing. There are arguments for moderating an unrestrained enthusiasm for precise performance measures: Where contradiction and confusion are essential elements of the values, precision misrepresents them. The more precise the measure of performance, the greater the motivation to find ways of scoring well on the measurement index without regard to the underlying goals. And precision in objectives does not allow creative interpretation of what the goal might mean (March, 1978). Thus, the introduction of precision into the valuation of performance involves a trade-off between the gains in outcomes attributable to closer articulation between action and performance on an index of performance and the losses in outcomes attributable to misrepresentation of goals, reduced motivation to development of goals, and concentration of effort on irrelevant ways of beating the index. Whether one is considering developing a performance evaluation scheme for managers, a testing procedure for students, or an understanding of personal preferences, there is a problem of determining the optimum clarity in goals.

The optimal sin problem. Standard notions of intelligent choice are theories of strict morality. That is, they presume that a person should do what he believes right and believe that what he does is right. Values and actions are to be consistent. Contrast that perspective with a view, somewhat more consistent with our behavior (as well as some theology), that there is such a thing as sin, that individuals and institutions sometimes do things even while recognizing that what they do is not what they wish they did, and that saints are a luxury to be encouraged only in small numbers. Or contrast a theory of strict morality with a view drawn from Nietzsche (1918) or Freud (1927) (see also Jones, 1926) of the complicated contradiction between conscience and self-interest. Although the issues involved are too subtle for brief treatment, a reasonably strong case can be made against strict morality and in favor of at least some sin, and therefore hypocrisy. One of the most effective ways of maintaining morality is through the remorse exhibited and felt at immoral action. Even if we are confident that our moral codes are correct, we may want to recognize human complexities. There will be occasions on which humans will be tempted by desires that they recognize as evil. If we insist that they maintain consistency between ethics and actions, the ethics will often be more likely to change than the actions. Hypocrisy is a long-run investment in morality made at some cost (the chance that, in action might otherwise adjust to morals). To encourage people always to take responsibility

for their actions is to encourage them to deny that bad things are bad — to make evil acceptable. At the same time, sin is an experiment with an alternative morality. By recognizing sin, we make it easier for persons to experiment with the possibility of having different tastes. Moral systems need those experiments, and regularly grant licenses to experiment to drunks, lovers, students, or sinners. These gains from sin are purchased by its costs. Thus, the optimization problem.

The optimal rationality problem. Calculated rationality is a technique for making decisions. In standard versions of theories of choice it is the only legitimate form of intelligence, each with claims to legitimacy. Learned behavior, with its claim to summarize an irretrievable but relevant personal history, or conventional behavior and rules, with their claims to capture the intelligence of survival over long histories of experience more relevant than the susceptible to immediate calculation, are clear alternative contenders. There are others: revelation or intuition, by which we substitute one guess for two; or imitation, or expertise, by which we substitute the guess of someone else for our own. Among all of these, only calculated rationality really uses conscious preferences of a current actor as a major consideration in making decisions. It is easy to show that there exist situations in which any one of these alternative techniques will make better decisions than the independent calculation of rational behavior by ordinary individuals or institutions. The superiority of learned or conventional behavior depends, in general, on the amount of experience it summarizes and the similarity between the world in which the experience was accumulated and the current world. The superiority of imitation depends, in general, on the relative competence of actor and expert and the extent to which intelligent action is reproducible but not comprehensible. At the same time, each form of intelligence exposes an actor to the risks of corruption. Imitation risks a false confidence in the neutrality of the process of diffusion; calculated rationality risks a false confidence in the neutrality of rational argument; and so on. It is not hard to guess that the relative sizes of these risks vary from individual to individual, or institution to institution. What are harder to specify in any very precise way are the extent and occasions on which a sensible person would rely on calculated rationality rather than the alternatives.

A ROMANTIC VISION

Prescriptive theories of choice are dedicated to perfecting the intelligence of human action by imagining that action stems from reason and by improving the technology of decision. Descriptive theories of choice are dedicated to perfecting the understanding of human action by imagining that action makes sense. Not all behavior makes sense; some of it is unreasonable. Not all decision technology is intelligent; some of it is foolish. Over the past twenty years, the contradiction between the search for sense in behavior and the search for improvement in behavior has focused on our interpretation of the way information about future

consequences is gathered and processed. The effort built considerably on the idea of bounded rationality and a conception of human decision making as limited by the cognitive capabilities of human beings. Over the next twenty years, I suspect the contradiction will be increasingly concerned with an interpretation of how beliefs about future preferences are generated and utilized. The earlier confrontation led theories of choice to a slightly clearer understanding of information processing and to some modest links with the technologies of computing, inference, and subjective probability. So perhaps the newer confrontation will lead theories of choice to a slightly clearer understanding of the complexities of preference processing and to some modest links with the technologies of ethics, criticism, and aesthetics. The history of theories of choice and their engineering applications suggests that we might appropriately be pessimistic about immediate, major progress. The intelligent engineering of tastes involves questions that encourage despair over their difficulty (Savage, 1954). But though hope for minor progress is a romantic vision, it may not be entirely inappropriate for a theory built on a romantic view of human destiny.

NOTE

Presented at a conference on the new industrial organization at Carnegie-Mellon University, October 14-15, 1977. The conference was organized to honor the contributions of Herbert A. Simon to economics, and his contribution to this paper is obvious. In addition, I profited from comments by Richard M. Cyert, Jon Elster, Alexander L. George, Elisabeth Hansen, Nannert O. Keohane, Robert O. Keohane, Tjalling Koopmans, Mancur Olson, Louis R. Pindy, Roy Radner, Giovanni Sartori, and Oliver E. Williamson. This research was supported by a grant from the Spencer Foundation. First published in *The Bell Journal of Economics*, 9 (1978), 587-608. Copyright © 1978, American Telephone and Telegraph Company.

REFERENCES

- Ackoff, R. L. and Sasieni, M. W. (1968). *Fundamentals of Operations Research*. New York: Wiley.
- Allison, G. T. (1969). *Essence of Decision: Explaining the Cuban Missile Crisis*. Boston: Little, Brown.
- Allison, G. T. and Halperin, M. H. (1972). "Bureaucratic politics: paradigm and some policy implications," in R. Tanter and R. H. Ullman, eds., *Theory and Policy in International Relations*, Princeton University Press.
- Bailey, M. J. and Olson, M. (1977). "Pure time preference, revealed marginal utility, and Friedman-Savage gambles." Unpublished manuscript.
- Beck, L. W. (1975). *The Actor and the Spectator*. New Haven: Yale University Press.
- Becker, G. S. (1965). "A theory of the allocation of time," *Economic Journal*, 75, 493-517.
- (1976). "Altruism, egoism, and genetic fitness: economics and sociobiology," *Journal of Economic Literature*, 14, 817-26.
- Binkley, R., Bronaugh, R., and Marras, A., eds. (1971). *Agent, Action, and Reason*. University of Toronto Press.

- Bower, J. L. (1968). "Descriptive decision theory from the administrative viewpoint," in R. A. Bauer and K. J. Gergen, eds., *The Study of Policy Formation*, New York: Free Press.
- Brams, S. J. (1975). *Game Theory and Politics*. New York: Free Press.
- Camus, A. (1951). *L'Homme révolté*. Paris: Gallimard. (Published in English as *The Rebel*.)
- Carter, E. E. (1971). "The behavioral theory of the firm and top-level corporate decisions," *Administrative Science Quarterly*, 16, 413-29.
- Catullus, G. V. (58 BC). *Carmina*, 85. Rome.
- Carroll, S. (1969). *Must We Mean What We Say?* New York: Scribner.
- Chames, A. and Cooper, W. W. (1963). "Deterministic equivalents for optimizing and satisficing under chance constraints," *Operations Research*, 11, 18-39.
- Christensen, S. (1976). "Decision making and socialization," in J. G. March and J. P. Olsen, eds., *Ambiguity and Choice in Organizations*. Bergen: Universitetsforlaget.
- Chomsky, N. (1968). *Language and Mind*. New York: Harcourt, Brace & World.
- Chiardi, J. (1960). *How Does a Poem Mean?* Cambridge: Houghton Mifflin.
- Cohen, M. D. and March, J. G. (1974). *Leadership and Ambiguity: The American College President*. New York: McGraw-Hill.
- Cohen, M. D., March, J. G., and Olsen, J. P. (1972). "A garbage can model of organizational choice," *Administrative Science Quarterly*, 17, 1-25.
- Connolly, T. (1977). "Information processing and decision making in organizations," in B. M. Staw and G. R. Salancik, eds., *New Directions in Organizational Behavior*, Chicago: St. Clair.
- Crozier, M. and Friedberg, E. (1977). *L'Acteur et le système*. Paris: Seuil.
- Cyert, R. M. and de Groot, M. H. (1973). "An analysis of cooperation and learning in a duopoly context," *The American Economic Review*, 63 (1), 24-37.
- (1975). "Adaptive utility," in R. H. Day and T. Groves, eds., *Adaptive Economic Models*. New York: Academic Press.
- Cyert, R. M. and March, J. G. (1963). *A Behavioral Theory of the Firm*. Englewood Cliffs, N.J.: Prentice-Hall.
- Day, R. H. and Groves, T., eds. (1975). *Adaptive Economic Models*. New York: Academic Press.
- Downs, A. (1967). *Inside Bureaucracy*. Boston: Little, Brown.
- Edelman, M. (1960). *The Symbolic Uses of Politics*. Champaign, Ill.: University of Illinois Press.
- Eliot, T. S. (1933). *The Use of Poetry and the Use of Criticism*. Cambridge: Harvard University Press.
- Elster, J. (1976). "A note on hysteresis in the social sciences," *Synthese*, 33, 371-91. (1977a). *Logic and Society*. London: Wiley.
- (1977b). "Ulysses and the sirens: a theory of imperfect rationality," *Social Science Information*, 16 (5), 469-526.
- (1977c). "Some unresolved problems in the theory of rational behavior." Unpublished manuscript.
- Farber, L. (1976). *Lying, Despair, Jealousy, Envy, Sex, Suicide, Drugs, and the Good Life*. New York: Basic Books.
- Farquharson, R. (1969). *Theory of Voting*. New Haven: Yale University Press.
- Freud, S. (1927). *The Ego and the Id*. London: Hogarth.
- Friedman, M. (1967). *To Deny Our Nothingness: Contemporary Images of Man*. New York: Delacorte.
- Halperin, M. H. (1974). *Bureaucratic Politics and Foreign Policy*. Washington, DC: The Brookings Institution.
- Haranyi, J. C. and Sellen, R. (1972). "A generalized Nash solution for two-person bargaining games with incomplete information," *Management Science*, 18, 80-106.
- Hegel, G. W. F. (1832). *G. W. F. Hegel's Werke*. Berlin: Duncker und Humblot.
- Hirschman, A. O. (1967). *Development Projects Observed*. Washington, DC: The Brookings Institution.
- Hirschman, A. and Lindblom, C. E. (1962). "Economic development, research and development, policy making: some converging views," *Behavioral Science*, 7, 211-22.
- Janis, I. L. and Mann, L. (1977). *Decision Making*. New York: Free Press.
- Johnson, E. (1968). *Studies in Multiplicative Decision Models*. Lund: Studentlitteratur.
- Jones, E. (1926). "The origin and structure of the superego," *International Journal of Psychoanalysis*, 7, 303-11.
- Keen, P. G. W. (1977). "The evolving concept of optimality," *TIMS Studies in the Management Sciences*, 6, 31-57.
- Koopmans, T. C. (1964). "On flexibility of future preferences," in M. W. Shelly and G. L. Bryan, eds., *Human Judgments and Optimality*, New York: Wiley.
- Kreiner, K. (1976). "Ideology and management in a garbage can situation," in J. G. March and J. P. Olsen, eds., *Ambiguity and Choice in Organizations*, Bergen: Universitetsforlaget.
- Lee, S. M. (1972). *Goal Programming for Decision Analysis*. Philadelphia: Auerbach.
- Lindblom, C. E. (1959). "The science of muddling through," *Public Administration Review*, 19, 79-88.
- (1965). *The Intelligence of Democracy*. New York: Macmillan.
- Long, N. E. (1958). "The local community as an ecology of games," *American Journal of Sociology*, 44, 251-61.
- McGuire, C. B. and Radner, R., eds. (1972). *Decision and Organization*. Amsterdam: North-Holland.
- March, J. G. (1962). "The business firm as a political coalition," *Journal of Politics*, 24, 662-78.
- (1973). "Model bias in social action," *Review of Educational Research*, 42, 413-29.
- (1977). "Administrative leadership in education." Unpublished manuscript.
- (1978). "American public school administration: a short analysis," *School Review*, 86, 217-50.
- March, J. G. and Olsen, J. P., eds. (1976). *Ambiguity and Choice in Organizations*. Bergen: Universitetsforlaget.
- March, J. G. and Romelager, P. (1976). "Position and presence in the drift of decisions," in J. G. March and J. P. Olsen, eds., *Ambiguity and Choice in Organizations*, Bergen: Universitetsforlaget.
- March, J. G. and Simon, H. A. (1958). *Organizations*. New York: Wiley.
- Marschak, J. and Radner, R. (1972). *Economic Theory of Teams*. New Haven: Yale University Press.
- Mayhew, D. R. (1974). *Congress: The Electoral Connection*. New Haven: Yale University Press.
- Mill, J. S. (1838). Reprinted in *Mill on Bentham and Coleridge*. London: Chatto and Windus, 1950.
- Mueller, D. C. (1976). "Public choice: a survey," *Journal of Economic Literature*, 14, 395-433.

- Nelson, R. R. and Winter, S. G. (1973). "Towards an evolutionary theory of economic capabilities," *The American Economic Review*, 63, 440-9.
- Newell, A. and Simon, H. A. (1972). *Human Problem Solving*. Englewood Cliffs, N.J.: Prentice-Hall.
- Nietzsche, F. (1918). *The Genealogy of Morals*. New York: Boni and Liveright.
- Olson, M. (1965). *The Logic of Collective Action*. New York: Schocken.
- (1976). "Exchange, integration, and grants," in M. Pfaff, ed., *Essays in Honor of Kenneth Boulding*, Amsterdam: North-Holland.
- Pattanaik, P. K. (1973). "Group choice with lexicographic individual orderings," *Behavioral Science*, 18, 118-23.
- Pfeffer, J. (1977). "Power and resource allocation in organizations," in B. M. Staw and G. R. Salancik, eds., *New Directions in Organizational Behavior*, Chicago: St Clair.
- Pondy, L. R. and Olson, M. L. (1977). "Organization and performance." Unpublished manuscript.
- Porta, A. M. and Haas, J. A. (1969). "Information effects on decision making," *Behavioral Science*, 14, 98-104.
- Radner, R. (1975a). "A behavioral model of cost reduction," *The Bell Journal of Economics*, 6 (1), 196-215.
- (1975b). "Satisficing," *Journal of Mathematical Economics*, 2, 253-62.
- Radner, R. and Rothschild, M. (1975). "On the allocation of effort," *Journal of Economic Theory*, 10, 358-76.
- Rapoport, A. (1960). *Fights, Games, and Debates*. Ann Arbor: University of Michigan Press.
- Riker, W. and Ordeshook, P. (1974). *An Introduction to Positive Political Theory*. Englewood Cliffs, N.J.: Prentice-Hall.
- Rosenberg, H. (1975). *Art on the Edge: Creators and Situations*. New York: Macmillan.
- Rothschild, M. and Stiglitz, J. (1976). "Equilibrium in competitive insurance markets: an essay on the economics of imperfect information," *Quarterly Journal of Economics*, 90, 629-49.
- Savage, L. J. (1954). *Foundations of Statistics*. New York: Wiley.
- Schelling, T. (1971). "On the ecology of micro-motives," *Public Interest*, 25, 59-98.
- Schutz, A. (1967). *The Phenomenology of the Social World*. Evanston, Ill.: Northwestern.
- Shakespeare, W. (1623). *Hamlet, Prince of Denmark*. Stratford-upon-Avon.
- Shefrin, H. M. and Thaler, R. (1977). "An economic theory of self-control." Unpublished manuscript.
- Simon, H. A. (1955). "A behavioral model of rational choice," *Quarterly Journal of Economics*, 69, 99-118.
- (1956). "Rational choice and the structure of the environment," *Psychological Review*, 63, 129-38.
- (1957). *Models of Man*. New York: Wiley.
- (1969). *The Science of the Artificial*. Cambridge: MIT Press.
- (1973). "The structure of ill-structured problems," *Artificial Intelligence*, 4, 181-201.
- Slovic, P., Fischhoff, B., and Lichtenstein, S. (1977). "Behavioral decision theory," *Annual Review of Psychology*, 28, 1-39.
- Spence, A. M. (1974). *Market Signaling*. Cambridge: Harvard University Press.
- Sproull, L. S., Weiner, S. S., and Wolf, D. B. (1978). *Organizing an Anarchy*. University of Chicago Press.
- Rationality, ambiguity, and the engineering of choice
- Steinberg, L. (1972). *Other Criteria: Confrontations with Twentieth Century Art*. New York: Oxford University Press.
- Steinbruner, J. D. (1974). *The Cybernetic Theory of Decision*. Princeton University Press.
- Stigler, G. J. (1961). "The economics of information," *Journal of Political Economy*, 69, 213-25.
- Stigler, G. J. and Becker, G. S. (1977). "De gustibus non est disputandum," *The American Economic Review*, 67 (March), 76-90.
- Stroiz, R. H. (1956). "Myopia and inconsistency in dynamic utility maximization," *Review of Economic Studies*, 23.
- Taylor, M. (1975). "The theory of collective choice," in F. I. Greenstein and N. W. Polsky, eds., *Handbook of Political Science*, Vol. 3, Reading, Mass.: Addison-Wesley.
- Taylor, R. N. (1975). "Psychological determinants of bounded rationality: implications for decision-making strategies," *Decision Sciences*, 6, 409-29.
- Thompson, J. (1967). *Organizations in Action*. New York: McGraw-Hill.
- Tullock, G. (1965). *The Politics of Bureaucracy*. Washington, DC: Public Affairs.
- Tversky, A. and Kahneman, D. (1974). "Judgment under uncertainty: heuristics and biases," *Science*, 185, 1124-31.
- Von Weizsäcker, C. C. (1971). "Notes on endogenous change of taste," *Journal of Economic Theory*, 3, 345-72.
- Vroom, V. H. (1964). *Work and Motivation*. New York: Wiley.
- Warwick, D. P. (1975). *A Theory of Public Bureaucracy: Politics, Personality, and Organization in the State Department*. Cambridge: Harvard University Press.
- Weick, K. E. (1969). *The Social Psychology of Organizing*. Reading, Mass.: Addison-Wesley.
- (1976). "Educational organizations as loosely coupled systems," *Administrative Science Quarterly*, 21, 1-18.
- Weiner, S. S. (1976). "Participation, deadlines, and choice," in J. G. March and J. P. Olson, eds., *Ambiguity and Choice in Organizations*, Bergen: Universitetsforlaget.
- Wildavsky, A. (1971). *Revolt Against the Masses and Other Essays on Politics and Public Policy*. New York: Basic Books.
- Wildavsky, A. and Pressman, H. (1973). *Implementation*. Berkeley: University of California Press.
- Williams, B. A. O. (1973). *Problems of the Self*. Cambridge University Press.
- Williamson, O. E. (1975). *Markets and Hierarchies*. New York: Free Press.
- Wilson, E. O. (1975). *Sociobiology*. Cambridge: Harvard University Press.
- Winter, S. G. (1964). "Economic 'natural selection' and the theory of the firm," *Yale Economic Essays*, 4, 225-72.
- (1971). "Satisficing, selection, and the innovating remnant," *Quarterly Journal of Economics*, 85, 237-61.
- (1975). "Optimization and evolution in the theory of the firm," in R. H. Day and T. Groves, eds., *Adaptive Economic Models*, New York: Academic Press.