Rationality as Process and as Product of Thought
Author(s): Herbert A. Simon
Source: *The American Economic Review*, Vol. 68, No. 2, Papers and Proceedings of the Ninetieth Annual Meeting of the American Economic Association (May, 1978), pp. 1-16
Published by: American Economic Association
Stable URL: http://www.jstor.org/stable/1816653
Accessed: 10-10-2017 12:41 UTC

RICHARD T. ELY LECTURE

Rationality as Process and as Product of Thought

By HERBERT A. SIMON*

This opportunity to deliver the Richard T. Ely Lecture affords me some very personal satisfactions. Ely, unbeknownst to him, bore a great responsibility for my economic education, and even for my choice of profession. The example of my uncle, Harold Merkel, who was a student of Commons and Ely at Wisconsin before World War I, taught me that human behavior was a fit subject for scientific study, and directed me to economics and political science instead of high energy physics or molecular biology. Some would refer to this as satisficing, for I had never heard of high energy physics or molecular biology, and hence was spared an agonizing weighing of alternative utiles. I simply picked the first profession that sounded fascinating.

Ely's influence went much further than that. My older brother's copy of his *Outlines of Economics*—the 1930 edition—was on our bookshelves when I prepared for high school debates on tariffs versus free trade, and on the Single Tax of Henry George. It provided me with a sufficiently good grounding in principles that I was later able to take Henry Simons' intermediate theory course at the University of Chicago, and the graduate courses of Frank Knight and Henry Schultz without additional preparation.

The Ely textbook, in its generation, held the place of Samuelson or Bach in ours. If it would not sound as though I were denying any progress in economics over the past half century, I might suggest that Ely's textbook could be substituted for any of our current ones at a substantial reduction in weight, and without students or teacher being more than dimly aware of the replacement. Of course they would not hear from Ely about marginal propensities to do this or that, nor about the late lamented Phillips curve. But monetarists could rejoice in Ely's uncompromising statement of the quantity theory (p. 298, italics), and in his assertion that "the solution of the problem of unemployment depends largely upon indirect measures, such as monetary and banking reform"—Ely does go on to say, however, that "we shall recognize that society must offer a willing and able man an opportunity to work" (p. 528).

I. Rationality in and out of Economics

I have more than personal reasons for directing your attention to Ely's textbook. On page 4, we find a definition of economics that is, I think, wholly characteristic of books contemporary with his. "Economics," he says, "is the science which treats of those social phenomena that are due to the wealth-getting and wealth-using activities of man." Economics, that is to say, concerns itself with a particular subset of man's behaviors—those having to do with the production, exchange, and consumption of goods and services.

Many, perhaps most, economists today would regard that view as too limiting. They would prefer the definition proposed in the *International Encyclopedia of the Social Sciences*: "Economics . . . is the study of the allocation of scarce resources among unlimited and competing uses" (vol. 4, p. 472). If beefsteak is scarce, they would say, so are votes, and the tools of economic analysis can be used as readily to analyze the allocation of the one as of the other. This point of view has launched economics into many excursions and incursions into political science and her other sister social sciences, and has generated a certain amount of hubris in the profession with respect to its broader civilizing mission. I

*Carnegie-Mellon University.

would suppose that the program of this meeting, with its emphasis upon the relations between economics and the other social sciences, is at least partly a reflection of that hubris.

A. *Rationality in Economics*

The topic of allocating scarce resources can be approached from either its normative or its positive side. Fundamental to the approach from either side are assumptions about the adaptation of means to ends, of actions to goals and situations. Economics, whether normative or positive, has not simply been the study of the allocation of scarce resources, it has been the study of the *rational* allocation of scarce resources.

Moreover, the term "rational" has long had in economics a much more specific meaning than its general dictionary signification of "agreeable to reason; not absurd, preposterous, extravagant, foolish, fanciful, or the like; intelligent, sensible." As is well known, the rational man of economics is a maximizer, who will settle for nothing less than the best. Even his expectations, we have learned in the past few years, are rational (see John Muth, 1961).[1] And his rationality extends as far as the bedroom for, as Gary Becker tells us, "he would read in bed at night only if the value of reading exceeded the value (to him) of the loss in sleep suffered by his wife" (1974, p. 1078).

It is this concept of rationality that is economics' main export commodity in its trade with the other social sciences. It is no novelty in those sciences to propose that people behave rationally—if that term is taken in its broader dictionary sense. Assumptions of rationality are essential components of virtually all the sociological, psychological, political, and anthropological theories with which I am familiar. What economics has to export, then, is not rationality, but a very particular and special form of it—the rationality of the utility maximizer, and a pretty smart one at that. But international flows have to be balanced. If the program of this meeting aims at more active intercourse between economics and her sister social sciences, then we must ask not only what economics will export, but also what she will receive in payment. An economist might well be tempted to murmur the lines of the tentmaker: "I wonder often what the Vintners buy—One half as precious as the stuff they sell."

My paper will be much concerned with that question, and before I proceed, it may be well to sketch in outline the path I propose to follow in answering it. The argument has three major steps.

First, I would like to expand on the theme that almost all human behavior has a large rational component, but only in terms of the broader everyday sense of rationality, not the economists' more specialized sense of maximization.

Second, I should like to show that economics itself has not by any means limited itself to the narrower definition of rationality. Much economic literature (for example, the literature of comparative institutional analysis) uses weaker definitions of rationality extensively; and that literature would not be greatly, if at all, improved by substituting the stronger definition for the weaker one.[2] To the extent that the weaker definition is adequate for purposes of analysis, economics will find that there is indeed much that is importable from the other social sciences.

Third, economics has largely been preoccupied with the *results* of rational choice rather than the *process* of choice. Yet as economic analysis acquires a broader concern with the dynamics of choice under uncertainty, it will become more and more essential to consider choice processes. In the past twenty years, there have been im-

[1]The term is ill-chosen, for rational expectations in the sense of Muth are profit-maximizing expectations only under very special circumstances (see below). Perhaps we would mislead ourselves and others less if we called them by the less alluring phrase, "consistent expectations."

[2]For an interesting argument in support of this proposition from a surprising source, see Becker (1962). What Becker calls "irrationality" in his article would be called "bounded rationality" here.

portant advances in our understanding of procedural rationality, particularly as a result of research in artificial intelligence and cognitive psychology. The importation of these theories of the processes of choice into economics could provide immense help in deepening our understanding of the dynamics of rationality, and of the influences upon choice of the institutional structure within which it takes place.

We begin, then, by looking at the broader concept of rationality to which I have referred, and its social science applications.

B. *Rationality in the Other Social Sciences: Functional Analysis*

Let me provide some examples how rationality typically enters into social science theories. Consider first so-called "social exchange" theories (see, for example, George Homans). The central idea here is that when two or more people interact, each expects to get something from the interaction that is valuable to him, and is thereby motivated to give something up that is valuable to the others. Social exchange, in the form of the "inducements-contributions balance" of Chester I. Barnard and the author (1947), has played an important role in organization theory, and in even earlier times (see, for example, George Simmel) was a central ingredient in sociological theories. Much of the theorizing and empirical work on the topic has been concerned with determining what constitutes a significant inducement or contribution in particular classes of exchange situations—that is, with the actual shape and substance of the "utility function." Clearly, the man of social exchange theory is a rational man, even if he is never asked to equate things at the margin.

It is perhaps more surprising to discover how pervasive assumptions of rationality are in psychoanalytic theory—confirming the suspicion that there is indeed method in madness. In his *Five Lectures* Sigmund Freud has this to say about neurotic illnesses:

> We see that human beings fall ill when, as a result of external obstacles or of an internal lack of adaptation, the satisfaction of their erotic needs *in reality* is frustrated. We see that they then take flight into *illness* in order that by its help they may find a satisfaction to take the place of what has been frustrated . . . We suspect that our patients' resistance to recovery is no simple one, but compounded of several motives. Not only does the patient's ego rebel against giving up the repressions by means of which it has risen above its original disposition, but the sexual instincts are unwilling to renounce their substitutive satisfaction so long as it is uncertain whether reality will offer them anything better.

Almost all explanations of pathological behavior in the psychoanalytic literature take this form: they explain the patient's illness in terms of the functions it performs for him.

The quotation from Freud is illustrative of a kind of functional reasoning that goes far beyond psychoanalysis and is widely used throughout the social sciences, and especially anthropology and sociology. Behaviors are functional if they contribute to certain goals, where these goals may be the pleasure or satisfaction of an individual or the guarantee of food or shelter for the members of a society. Functional analysis in this sense is concerned with explaining how "major social patterns operate to maintain the integration or adaptation of the larger system" (see Frank Cancian). Institutions are functional if reasonable men might create and maintain them in order to meet social needs or achieve social goals.

It is not necessary or implied that the adaptation of institutions or behavior patterns to goals be conscious or intended. When awareness and intention are present, the function is usually called *manifest*, otherwise it is a *latent* function. The function, whether it be manifest or latent, provides the grounds for the reasonableness or rationality of the institution or behavior pattern. As in economics, evolutionary arguments are often adduced to explain the persistence and survival of functional patterns, and to avoid assumptions of deliberate calculation in explaining them.

In practice, it is very rarely that the existence or character of institutions are *deduced* from the functions that must be performed for system survival. In almost all cases it is the other way round; it is empirical observation of the behavior pattern that raises the question of why it persists—what function it performs. Perhaps, in an appropriate axiomatic formulation, it would be possible to *deduce* that every society must have food-gathering institutions. In point of fact, such institutions can be *observed* in every society, and their existence is then rationalized by the argument that obtaining food is a functional requisite for all societies. This kind of argument may demonstrate the sufficiency of a particular pattern for performing an essential function, but cannot demonstrate its necessity—cannot show that there may not be alternative, functionally equivalent, behavior patterns that would satisfy the same need.

The point may be stated more formally. Functional arguments are arguments about the movements of systems toward stable self-maintaining equilibria. But without further specification, there is no reason to suppose that the attained equilibria that are reached will be global maxima or minima of some function rather than local, relative maxima or minima. In fact, we know that the conditions that every local maximum of a system be a global maximum are very strong (usually some kind of "convexity" conditions).

Further, when the system is complex and its environment continually changing (that is, in the conditions under which biological and social evolution actually take place), there is no assurance that the system's momentary position will lie anywhere near a point of equilibrium, whether local or global. Hence, all that can be concluded from a functional argument is that certain characteristics (the satisfaction of certain functional requirements in a particular way) are consistent with the survival and further development of the system, not that these same requirements could not be satisfied in some other way. Thus, for example, societies can satisfy their functional needs for food by hunting or fishing activities, by agriculture, or by predatory exploitation of other societies.

C. *Functional Analysis in Economics*

Functional analysis of exactly this kind, though with a different vocabulary, is commonly employed by economists, especially when they seek to use economic tools to "explain" institutions and behaviors that lie outside the traditional domains of production and distribution. Moreover, it occurs within those domains. As an example, the fact is observed that individuals frequently insure against certain kinds of contingencies. Attitudes are then postulated (for example, risk aversion) for which buying insurance is a functional and reasonable action. If some people are observed to insure, and others not, then this difference in behavior can be explained by a difference between them in risk aversion.

To take a second example, George Stigler and Becker wish to explain the fact (if it is a fact—their empiricism is very casual) that as people hear more music, they want to hear still more. They invent a commodity, "music appreciation" (not to be confused with time spent in listening to music), and suggest that listening to music might produce not only immediate enjoyment but also an investment in *capacity* for appreciating music (i.e., in amount of enjoyment produced per listening hour). Once these assumptions are granted, various conclusions can be drawn about the demand for music appreciation. However, only weak conclusions follow about listening time unless additional strong postulates are introduced about the elasticity of demand for appreciation.

A rough "sociological" translation of the Stigler-Becker argument would be that listening to music is functional both in producing pleasure and in enhancing the pleasure of subsequent listening—a typical functional argument. It is quite unclear what is gained by dressing it in the garb of marginalism. We might be willing to grant that people would be inclined to invest more in musical appreciation early in life than later in life (because they would have a longer time in which to amortize the investment) without insisting that costs and returns were being equated at the margin, and without gaining any new insights into the situation from making the latter assumption.

A sense of fairness compels me to take a third example from my own work. In my 1951 paper, I defined the characteristics of an employment contract that distinguish it from an ordinary sales contract, and then showed why reasonable men might prefer the former to the latter as the basis for establishing an employment relation. My argument requires a theorem and fifteen numbered equations, and assumes that both employer and employee maximize their utilities. Actually, the underlying functional argument is very simple. An employee who didn't care very much which of several alternative tasks he performed would not require a large inducement to accept the authority of an employer—that is, to permit the employer to make the choice among them. The employer in turn would be willing to provide the necessary inducement in order to acquire the right to postpone his decisions about the employee's agenda, and in this way to postpone some of his decisions whose outcomes are contingent on future uncertain events.[3] The rigorous economic argument, involving the idea of maximizing behavior by employer and employee, is readily translatable into a simple qualitative argument that an employment contract may be a functional ("reasonable") way of dealing with certain kinds of uncertainty. The argument then explains why employment relations are so widely used in our society.

The translation of these examples of economic reasoning into the language of functional analysis could be paralleled by examples of translation scholarship which run in the opposite direction. Political scientists, for example, long ago observed that under certain circumstances institutions of representative democracy spawned a multiplicity of political parties, while under other circumstances, the votes were divided in equilibrium between two major parties. These contrasting equilibria could readily be shown by functional arguments to result from rational voting decisions under different rules of the electoral game, as was observed by Maurice Duverger, in his classic work on political parties, as well as by a number of political scientists who preceded him. In recent years, these same results have been rederived more rigorously by economists and game theorists, employing much stronger assumptions of utility maximization by the voters; it was hard to see that the maximization assumptions have produced any new predictions of behavior.[4]

D. *Summary*

Perhaps these examples suffice to show that there is no such gap as is commonly supposed between the view of man espoused by economics and the view found in the other social sciences. The view of man as rational is not peculiar to economics, but is endemic, and even ubiquitous, throughout the social sciences. Economics tends to emphasize a particular

[3]Recently, Oliver Williamson has pointed out that I would have to introduce slightly stronger assumptions to justify the employment contract as rational if one of the alternatives to it were what he calls a "contingent claims" contract, but the point of my example is not affected. To exclude the contingent claims contract as a viable alternative, we need merely take account of the large transaction costs it would entail under real world conditions.

[4]For an introduction to this literature, see William H. Riker and Peter C. Ordeshook, and Riker. Anthony Downs' book belongs to an intermediate genre. While it employs the language of economics, it limits itself to verbal, nonrigorous reasoning which certainly does not make any essential use of maximizing assumptions (as contrasted with rationality assumptions in the broader sense), and which largely translates into the economic vocabulary generalizations that were already part of the science and folklore of politics. In the next section, other examples of this kind of informal use of rationality principles are examined to analyze institutions and their behavior.

form of rationality—maximizing behavior—as its preferred engine of explanation, but the differences are often differences in vocabulary more than in substance. We shall see in a moment that in much economic discussion the notion of maximization is used in a loose sense that is very close to the common sense notions of rationality used elsewhere in the social sciences.

One conclusion we may draw is that economists might well exercise a certain amount of circumspection in their endeavors to export economic analysis to the other social sciences. They may discover that they are sometimes offering commodities that are already in generous supply, and which can therefore be disposed of only at a ruinously low price. On the other side of the trade, they may find that there is more of interest in the modes and results of inquiry of their fellow social scientists than they have generally been aware.

II. On Applying the Principle of Rationality

What is characteristic of the examples of functional analysis cited in the last section, whether they be drawn from economics or from the other social sciences, is that they are not focused on, or even much concerned with, how variables are equated at the margin, or how equilibrium is altered by marginal shifts in conditions (for example, shifts in a supply or demand schedule). Rather, they are focused on qualitative and structural questions, typically, on the choice among a small number of discrete institutional alternatives:

Not "how much flood insurance will a man buy?" but "what are the structural conditions that make buying insurance rational or attractive?"

Not "at what levels will wages be fixed?" but "when will work be performed under an employment contract rather than a sales contract?"

If we want a natural science analogy to this kind of theorizing, we can find it in geology. A geologist notices deep scratches in rock; he notices that certain hills of gravel are elongated along a north-south axis, and that the boulders embedded in them are not as smooth as those usually found on beaches. To explain these facts, he evokes a structural, and not at all quantitative, hypothesis: that these phenomena were produced by the process of glaciation.

In the first instance, he does not try to explain the depth of the glacial till, or estimate the weight of the ice that produced it, but simply to identify the basic causative process. He wants to explain the role of glaciation, of erosion, of vulcanization, of sedimentation in producing the land forms that he observes. His explanations, morever, are after-the-fact, and not predictive.

A. *Toward Qualitative Analysis*

As economics expands beyond its central core of price theory, and its central concern with quantities of commodities and money, we observe in it this same shift from a highly quantitative analysis, in which equilibration at the margin plays a central role, to a much more qualitative institutional analysis, in which discrete structural alternatives are compared.

In these analyses aimed at explaining institutional structure, maximizing assumptions play a much less significant role than they typically do in the analysis of market equilibria. The rational man who sometimes prefers an employment contract to a sales contract need not be a maximizer. Even a satisficer will exhibit such a preference whenever the difference in rewards between the two arrangements is sufficiently large and evident.

For this same reason, such analyses can often be carried out without elaborate mathematical apparatus or marginal calculation. In general, much cruder and simpler arguments will suffice to demonstrate an inequality between two quantities than are required to show the conditions under which these quantities are equated at the margin. Thus, in the recent works of Janos Kornai, Williamson, and John Montias on economic organization, we find only rather modest and simple applications of mathematical analysis. In the ways in which they involve principles of rationality, the arguments of these authors resemble James March and the author's *Organizations* more closely than Paul Samuelson's *Foundations*.[5]

What is the predominant form of reasoning that we encounter in these theoretical treatments of social institutions? Do they contain arguments based on maximizing assumptions? Basically, they rest upon a very simple form of causal analysis. Particular institutional structures or practices are seen to entail certain undesirable (for example, costly) or desirable (for example, value-producing) consequences. *Ceteris paribus*, situations and practices will be preferred when important favorable consequences are associated with them, and avoided when important unfavorable consequences are associated with them. A shift in the balance of consequences, or in awareness of them, may motivate a change in institutional arrangements.

Consider the following argument from Montias typical of this genre of analysis, which relates to the balance in organizations between centralization and decentralization.

> Decentralizing measures are generally aimed at remedying two shortcomings of an 'overcentralized' system structure. (1) Superordinates are overburdened with responsibility for the detailed direction and coordination of their subordinates' activities. (2) This 'petty tutelage' deprives subordinates of the opportunity to make decisions that might increase the payoff of the organization of which they are a part. . . . Why not loosen controls . . . ? . . . When controls are loosened, unless the incentive system is modified to bring about greater harmony between the goals of supervisors and supervisees, it may induce producers to shift their input and output mix in directions that . . . vitiate any benefits that might be reaped by the organization as a whole from the exercise of greater initiative at lower tiers. [p. 215]

Here two costs or disadvantages of centralization (burden on supervisors, restriction of choice-set of subordinates) are set off against a disadvantage of decentralization (goals of subordinates divergent from organization goals).

What can we learn about organization from an argument like this? Certainly little or nothing about the optimal balance point between centralization and decentralization in any particular organization. Rather, we might derive conclusions of these kinds:

1. That increasing awareness of one of the predicted consequences may cause an organization to move in the direction of centralization or decentralization. (For example, an egregious case of "suboptimizing" by a subordinate may cause additional centralized controls to be instituted.)

2. That new technical devices may tilt the balance between centralization and decentralization. For example, invention and adoption of divisionalized profit and loss statements led toward decentralization of many large American business firms in the 1950's; while reduction in information costs through computerization led at a later date to centralization of inventory control decisions in those same firms.

Of course Montias' conclusions could also be derived from a more formal optimization analysis—in fact he presents

[5]A notable exception to this generalization about the economic literature on organizations is the work of Jacob Marschak and Roy Radner on the theory of teams. These authors chose the strategy of detailed, precise analysis of the implications of maximizing assumptions for the transmission of information in organizations. The price they paid for this rigor was to find themselves limited to the highly simplified situations where solutions could be found for the mathematical problems they posed. We need not, of course, make an either-or choice between these two modes of inquiry. While it may be difficult or impossible to extend the formal analysis of the theory of teams to problems of real world complexity, the rigorous microtheory may illuminate the workings of important component mechanisms in the complex macrosituations. The methodological issues in choosing between analytic tractability and realism are quite parallel to those involved in the choice between laboratory and field methods for gathering empirical information about social phenomena. Neither one by itself marks the exclusive path toward truth.

such an analysis on the two pages following the passage quoted above. But it is not clear that anything new is added by the formalization, since the parameters imputed to the system are largely unmeasured and unmeasurable.

There is something to be said for an Ockham's Razor that, eschewing assumptions of optimization, provides an explanation of behavior that is consistent with *either* optimizing or satisficing procedures on the part of the human agents. Parsimony recommends that we prefer the postulate that men are reasonable to the postulate that they are supremely rational when either one of the two assumptions will do our work of inference as well as the other.[6]

B. *Procedural Rationality*

The kind of qualitative analysis I have been describing has another virtue. In complex situations there is likely to be a considerable gap between the real environment of a decision (the world as God or some other omniscient observer sees it) and the environment as the actors perceive it. The analysis can then address itself either to normative questions—the whole range of consequences that *should* enter into decisions in such situations—or descriptive questions, including the questions of which components of the situation are likely to be taken into account by the actors, and how the actors are likely to represent the situation as a whole.

In the precomputer era, for example, it was very difficult for managers in business organizations to pay attention to all the major variables affected by their decisions. Company treasurers frequently made decisions about working capital with little or no attention to their impact on inventory levels, while production and marketing executives made decisions about inventory without taking into account impacts on liquidity. The introduction of computers changed the ways in which executives were able to reach decisions; they could now view them in terms of a much wider set of interrelated consequences than before. The perception of the environment of a decision is a function of—among other things—the information sources and computational capabilities of the executives who make it.

Learning phenomena are also readily handled within this framework. A number of the changes introduced into planning and control procedures in eastern European countries during the 1960's were instituted when the governments in question learned by experience of some of the dysfunctional consequences of trying to control production by means of crude aggregates of physical quantities. An initial distrust of prices and market mechanisms was gradually and partially overcome after direct experience of the disadvantages of some of the alternative mechanisms. These learning experiences could be paralleled with experiences of American steel companies, for example, that experimented with tonnage incentives for mill department superintendents.

A general proposition that might be asserted about organizations is that the number of considerations that are potentially relevant to the effectiveness of an organization design is so large that only a few of the more salient of these lie within the circle of awareness at any given time, that the membership of this subset changes continually as new situations (produced by external or internal events) arise, and that "learning" in the form of reaction to perceived consequences is the dominant way in which rationality exhibits itself.

In a world where these kinds of adjustments are prominent, a theory of rational behavior must be quite as much concerned with the characteristics of the rational actors—the means they use to cope with uncertainty and cognitive complexity—as

[6]Ockham is usually invoked on behalf of the parsimony of optimizing assumptions, and against the additional *ad hoc* postulates that satisficing models are thought to require in order to guarantee uniqueness of solutions. But that argument only applies when we are trying to deduce unique equilibria, a task quite different from the one most institutional writers set for themselves. However, I have no urge to enlarge on this point. My intent here is not polemical, on behalf of satisficing postulates, but rather to show how large a plot of common ground is shared by optimizing and satisficing analysis. Again, compare Becker (1962).

with the characteristics of the objective environment in which they make their decisions. In such a world, we must give an account not only of *substantive rationality*—the extent to which appropriate courses of action are chosen—but also *procedural rationality*—the effectiveness, in light of human cognitive powers and limitations, of the *procedures* used to choose actions. As economics moves out toward situations of increasing cognitive complexity, it becomes increasingly concerned with the ability of actors to cope with the complexity, and hence with the procedural aspects of rationality. In the remainder of my talk, I would like to develop this concept of procedural rationality, and its implications for economic analysis.

III. Mind as the Scarce Resource

Until rather recently, such limited attention as was paid by economists to procedural, as distinct from substantive, rationality was mainly motivated by the problems of uncertainty and expectations. The simple notion of maximizing utility or profit could not be applied to situations where the optimum action depended on uncertain environmental events, or upon the actions of other rational agents (for example, imperfect competition).

The former difficulty was removed to some degree by replacing utility maximization with the maximization of subjective expected utility (*SEU*) as the criterion of rationality. In spite of its conceptual elegance, however, the *SEU* solution has some grave defects as either a normative or a descriptive formulation. In general, the optimal solution depends upon all of the moments of the frequency distributions of uncertain events. The exceptions are a small but important class of cases where the utility or profit function is quadratic and all constraints are in the form of equations rather than inequalities.[7] The empirical defect of the *SEU* formulation is that when it has been subjected to test in the laboratory or the real world, even in relatively simple situations, the behavior of human subjects has generally departed widely from it.

Some of the evidence has been surveyed by Ward Edwards, and more recently by Daniel Kahneman and Amos Tversky. They describe experimental situations in which estimates formed on the basis of initial information are not revised nearly as much by subsequent information as would be required by Bayes' Theorem. In other situations, subjects respond largely to the information received most recently, and take inadequate account of prior information.

Behavior that is radically inconsistent with the *SEU* framework occurs also in naturalistic settings. Howard Kunreuther et al. have recently carried out extensive studies of behavior and attitudes relating to the purchase of flood insurance by persons owning property in low-lying areas. They found that knowledge of the availability of insurance, or rates, and of objective risks was very imperfect, and that the actual decisions whether or not to insure were related much more to personal experience with floods than to any objective facts about the situation—or even to personal subjective beliefs about those facts. In the face of this evidence, it is hard to take *SEU* seriously as a theory of actual human behavior in the face of uncertainty.[8]

For situations where the rationality of an action depends upon what others (who are also striving to be rational) do again, no consensus has been reached as to what constitutes optimal behavior. This is one of the reasons I have elsewhere called imperfect competition "the permanent and ineradicable scandal of economic theory" (1976b, p. 140). The most imaginative and

[7]In this case the expected values of the environmental variables serve as certainty equivalents, so that *SEU* maximization requires only replacing the unknown true values by these expected values. See the author (1957).

[8]Kunreuther et al. point out that the theory cannot be "saved" by assuming utility to be radically nonlinear in money. In the flood insurance case, that interpretation of the data would work only if we were willing to assume that money has strongly *increasing* marginal utility, not a very plausible escape route for the theory.

ambitious attempt to resolve the difficulty was the von Neumann-Morgenstern theory of games, which is embarrassing in the wealth of alternative solutions it offers. While the theory of games reveals the potential richness of behavior when rational individuals are faced with conflict of interest, the capability of reacting to each other's actions (or expected actions), and possibilities for coalition, it has provided no unique and universally accepted criterion of rationality to generalize the *SEU* criterion and extend it to this broader range of situations.

The so-called "rational expectations" models, currently so popular (and due originally to Muth), pass over these problems rather than solving them. They ignore potential coalitions and attempted mutual outguessing behavior, and correspond to optimal solutions only when the losses are quadratic functions of the errors of estimate.[9] Hence they do not correspond to any classical criterion of rationality, and labeling them with that term, rather than the more neutral "consistent expectations," provides them with a rather unwarranted legitimation.

Finally, it should be remarked that the main motivation in economics for developing theories of uncertainty and mutual expectations has not been to replace substantive criteria of rationality with procedural criteria, but rather to find substantive criteria broad enough to extend the concept of rationality beyond the boundaries of static optimization under certainty. As with classical decision theory, the interest lies not in *how* decisions are made but in *what* decisions are made. (But see, contra, such analyses as Richard Cyert and Morris DeGroot.)

[9]That is, only under the conditions where the uncertainty equivalents of fn. 8 exist. Under other circumstances, a "rational" person would be well advised, if he knew that all others were following the "rational expectations" or "consistent expectations" rule, to recalculate his own optimal behavior on that assumption. Of course if others followed the same course, we would be back in the "outguessing" situation.

A. *Search and Teams*

Decision procedures have been treated more explicitly in the small bodies of work that have grown up in economics on the theory of search and on the theory of teams. Both these bodies of theory are specifically concerned with the limits on the ability of the economic actor to discover or compute what behavior is optimal for him. Both aspire not only to *take account* of human bounded rationality, but to *bring it within the compass* of the rational calculus. Let me explain what I mean by that distinction.

Problems of search arise when not all the alternatives of action are presented to the rational actor *ab initio*, but must be sought through some kind of costly activity. In general, an action will be chosen before the search has revealed all possible alternatives. One example of this kind of problem is the sale of a house, or some other asset, when offers are received sequentially and remain open for only a limited time (see the author, 1955). Another example which has been widely cited is the purchase of an automobile involving travel to dealers' lots (see Stigler, 1961). In both these examples, the question is not how the search is carried out, but how it is decided when to terminate it—that is, the amount of search. The question is answered by postulating a cost that increases with the total amount of search. In an optimizing model, the correct point of termination is found by equating the marginal cost of search with the (expected) marginal improvement in the set of alternatives. In a satisficing model, search terminates when the best offer exceeds an aspiration level that itself adjusts gradually to the value of the offers received so far. In both cases, search becomes just another factor of production, and investment in search is determined by the same marginal principle as investment in any other factor. However cavalierly these theories treat the actual search process, they do recognize explicitly that information gathering is not a free activity, and that unlimited amounts of it are not available.

The theory of teams, as developed by Marschak and Radner, goes a step farther in specifying the procedure of decision. That theory, as is well known, is concerned with the improvement that may be realized in a team's decisions by interchange of information among the team members. But here the theory does not limit itself to determining the aggregate amount of information that should be transmitted, but seeks to calculate what messages should be exchanged, under what conditions, and at what cost. The content of the communication as well as the total amount of information becomes relevant to the theory.

In its attitude toward rationality, the theory of teams is as "classical," however, as is search theory. The bounds on the rationality of the team members are "externalized" and represented as costs of communication, so that they can be folded into the economic calculation along with the costs and benefits of outcomes.

B. *Rational Search Procedures*

To find theories that compare the merits of alternative search procedures, we must look largely outside the domain of economics. A number of such theories have been developed in the past thirty years, mainly by management scientists and researchers in the field of artificial intelligence. An important example is the body of work that has been done on integer programming.

Integer programming problems resemble linear programming problems (to maximize some quantity, subject to constraints in the form of linear equations and inequalities), with the added condition that certain variables can only take whole numbers as their values. The integer constraint makes inapplicable most of the powerful computational methods available for solving linear programming problems, with the result that integer programming problems are far less tractable, computationally, than linear programming problems having comparable numbers of variables.

Solution methods for integer programming problems use various forms of highly selective search—for example branch-and-bound methods that establish successively narrower limits for the value of the optimum, and hence permit a corresponding narrowing of search to promising regions of the space. It becomes a matter of considerable practical and theoretical interest to evaluate the relative computational efficiency of competing search procedures, and also to estimate how the cost of search will grow with the size of the problem posed. Until recently, most evaluation of search algorithms has been empirical: they have been tested on sample problems. Recently, however, a body of theory—called theory of computational complexity—has grown up that begins to answer some of these questions in a more systematic way.

I cannot give here an account of the theory of computational complexity, or all of its implications for procedural rationality. A good introduction will be found in Alfred Aho et al. One important set of results that comes out of the theory does require at least brief mention. These results have to do with the way in which the amount of computation required to solve problems of a given class grows with the size of the problems—with the number of variables, say.[10]

In a domain where computational requirements grow rapidly with problem size, we will be able to solve only small problems; in domains where the requirements grow slowly, we will be able to solve much larger problems. The problems that the real world presents to us are generally enormous compared with the problems that we can solve on even our largest computers. Hence, our computational models are always rough approximations to the reality, and we must hope that the approximation will not be too inexact to be useful.

[10]Most of the theorems in computational complexity have to do with the "worst case," that is, with the maximum amount of computation required to solve *any* problem of the given class. Very few results are available for the expected cost, averaged over all problems of the class.

We will be particularly concerned that computational costs not increase rapidly with problem size.

It is customary in the theory of computational complexity to regard problems of a given size as "tractable" if computations do not grow faster than at some fixed power of problem size. Such classes of problems are known as "polynomial complex." Problems that grow exponentially in complexity with size are not polynomial complex, since the rate of growth of computation comes to exceed any fixed power of their size.

A large and important class of problems which includes the general integer programming problem, as well as standard scheduling problems, all have been shown to have the same level of complexity—if one is polynomial complex, then all are; if one is not polynomial complex, then none are. These problems have been labeled "*NP*-complete." It is conjectured, but not yet proven, that the class of *NP*-complete problems is not polynomially complex, but probably exponentially complex.

The significance of these findings and conjectures is in showing that computational difficulties, and the need to approximate, are not just a minor annoying feature of our world to be dealt with by manufacturing larger computers or breeding smarter people. Complexity is deep in the nature of things, and discovering tolerable approximation procedures and heuristics that permit huge spaces to be searched very selectively lies at the heart of intelligence, whether human or artificial. A theory of rationality that does not give an account of problem solving in the face of complexity is sadly incomplete. It is worse than incomplete; it can be seriously misleading by providing "solutions" to economic questions that are without operational significance.

One interesting and important direction of research in computational complexity lies in showing how the complexity of problems might be decreased by weakening the requirements for solution—by requiring solutions only to approximate the optimum, or by replacing an optimality criterion by a satisficing criterion. Results are still fragmentary, but it is already known that there are some cases where such modifications reduce exponential or *NP*-complete problem classes to polynomial-complete classes.

The theory of heuristic search, cultivated in artificial intelligence and information processing psychology, is concerned with devising or identifying search procedures that will permit systems of limited computational capacity to make complex decisions and solve difficult problems. (For a general survey of the theory, see Nils Nilsson.) When a task environment has patterned structure, so that solutions to a search problem are not scattered randomly throughout it, but are located in ways related to the structure, then an intelligent system capable of detecting the pattern can exploit it in order to search for solutions in a highly selective way.

One form, for example, of selective heuristic search, called best-first search, assigns to each node in the search space an estimate of the distance of that node from a solution. At each stage, the next increment of effort is expended in searching from the node, among those already reached, that has the smallest distance estimate (see, for example, the author and J.B. Kadane). As another example, when the task is to find a good or best solution, it may be possible to assign upper and lower bounds on the values of the solutions that can be obtained by searching a particular part of the space. If the upper bound on region *A* is lower than the lower bound on some other region, then region *A* does not need to be searched at all.

I will leave the topics of computational complexity and heuristic search with these sketchy remarks. What implications these developments in the theory of procedural rationality will have for economics defined as "the science which treats of the wealth-getting and wealth-using activities of man" remain to be seen. That they are an integral part of economics defined as "the science which treats of the allocation of scarce resources" is obvious. The scarce resource is computational capacity—the mind. The ability of man to solve complex problems, and the magnitude of the resources that have to be allocated to solving them, depend on the efficiency with which this resource, mind, is deployed.

C. *Attention as the Scarce Resource*

Finally, I would like to turn from the rather highly developed approaches to procedural rationality that I have been discussing back to the more qualitative kinds of institutional issues that were considered in the previous section of this paper. Many of the central issues of our time are questions of how we use limited information and limited computational capacity to deal with enormous problems whose shape we barely grasp.

For many purposes, a modern government can be regarded as a parallel computing device. While one part of its capability for rational problem solving is directed to fire protection, another is directed to paving highways, and another to collecting refuse. For other important purposes, a government, like a human being, is a serial processing system, capable of attending to only one thing at a time. When important new policies must be formulated, public and official attention must be focused on one or a few matters. Other concerns, no matter how pressing, must wait their turn on the agenda. When the agenda becomes crowded, public life begins to appear more and more as a succession of crises. When problems become interrelated, as energy and pollution problems have become, there is the constant danger that attention directed to a single facet of the web will spawn solutions that disregard vital consequences for the other facets. When oil is scarce, we return to coal, but forget that we must then deal with vastly increased quantities of sulfur oxides in our urban air. Or we outlaw nuclear power stations because of radiation hazards, but fail to make alternative provision to meet our energy needs. It is futile to talk of substantive rationality in public affairs without considering what procedural means are available to order issues on the public agenda in a rational way, and to insure attention to the indirect consequences of actions taken to reach specific goals or solve specific problems.

In a world where information is relatively scarce, and where problems for decision are few and simple, information is almost always a positive good. In a world where attention is a major scarce resource, information may be an expensive luxury, for it may turn our attention from what is important to what is unimportant. We cannot afford to attend to information simply because it is there. I am not aware that there has been any systematic development of a theory of information and communication that treats attention rather than information as the scarce resource.[11] Some of the practical consequences of attention scarcity have already been noticed in business and government, where early designs of so-called "management information systems" flooded executives with trivial data and, until they learned to ignore them, distracted their attention from more important matters. It is probably true of contemporary organizations that an automated information system that does not consume and digest vastly more information than it produces and distributes harms the performance of the organization in which it is incorporated.

The management of attention and tracing indirect consequences of action are two of the basic issues of procedural rationality that confront a modern society. There are others of comparable importance: what decision-making procedure is rational when the basic quantities for making marginal comparisons are simply not known? A few years ago, I served as chairman of a National Academy of Sciences (*NAS*) committee whose job it was to advise the Congress on the control of automobile emissions (see *NAS*, Coordinating Committee on Air Quality Studies). It is easy to formulate an *SEU* model to conceptualize the problem. There is a production function for automobiles that associates different costs with different levels of emissions. The laws govern-

[11]Some unsystematic remarks on the subject will be found in the author (1976a, chs. 13, 14).

ing the chemistry of the atmosphere determine the concentrations of polluting substances in the air as a function of the levels of emissions. Biomedical science tells us what effects on life and health can be expected from various concentrations of pollutants. All we need do is to attach a price tag to life and health, and we can calculate the optimum level of pollution control.

There is only one hitch—which will be apparent to all of you. None of the relevant parameters of the various "production functions" are known—except, within half an order of magnitude, the cost of reducing the emissions themselves. The physics and chemistry of the atmosphere presents a series of unsolved problems—particularly relating to the photochemical reactions affecting the oxides of nitrogen and ozone. Medical science is barely able to detect that there *are* health effects from pollutants, much less measure how large these effects are. The committee's deliberations led immediately to one conclusion—one that congressmen are accustomed to hearing from such committees: We need more research. But while the research is being done, what provisions should be incorporated in the Clean Air Act of 1977 (or the Acts of 1978 through 2000, for that matter)? For research won't give us clear answers then either. What constitutes procedural rationality in such circumstances?

"Reasonable men" reach "reasonable" conclusions in circumstances where they have no prospect of applying classical models of substantive rationality. We know only imperfectly how they do it. We know even less whether the procedures they use in place of the inapplicable models have any merit—although most of us would choose them in preference to drawing lots. The study of procedural rationality in circumstances where attention is scarce, where problems are immensely complex, and where crucial information is absent presents a host of challenging and fundamental research problems to anyone who is interested in the rational allocation of scarce resources.

IV. Conclusion

In histories of human civilization, the invention of writing and the invention of printing are always treated as key events. Perhaps in future histories the invention of electrical communication and the invention of the computer will receive comparable emphasis. What all of these developments have in common, and what makes them so important, is that they represent basic changes in man's equipment for making rational choices—in his computational capabilities. Problems that are impossible to handle with the head alone (multiplying large numbers together, for example) become trivial when they can be written down on paper. Interactions of energy and environment that almost defy conceptualization lend themselves to at least approximate modeling with modern computers.

The advances in man's capacity for procedural rationality are not limited to these obvious examples. The invention of algebra, of analytic geometry, of the calculus were such advances. So was the invention, if we may call it that, of the modern organization, which greatly increased man's capacity for coordinated parallel activity. Changes in the production function for information and decisions are central to any account of changes over the centuries of the human condition.

In the past, economics has largely ignored the processes that rational man uses in reaching his resource allocation decisions. This was possibly an acceptable strategy for explaining rational decision in static, relatively simple problem situations where it might be assumed that additional computational time or power could not change the outcome. The strategy does not work, however, when we are seeking to explain the decision maker's behavior in complex, dynamic circumstances that involve a great deal of uncertainty, and that make severe demands upon his attention.

As economics acquires aspirations to explain behavior under these typical conditions of modern organizational and public life, it will have to devote major energy to building a theory of procedural rationality to complement existing theories of substantive rationality. Some elements of such a theory can be borrowed from the neighboring disciplines of operations research, artificial intelligence, and cognitive psychology; but an enormous job remains to be done to extend this work and to apply it to specifically economic problems.

Jacob Marschak, throughout his long career, had a deep belief in and commitment to the interdependencies and complementarity of the several social sciences. I have shared that belief and commitment, without always agreeing with him in detail as to the precise route for exploiting it. The developments I have been describing strengthen greatly, it seems to me, the rational grounds for both belief and commitment. Whether we accept the more restricted definition of economics that I quoted from Ely's textbook, or the wider definition that is widely accepted today, we have every reason to try to communicate with the other social sciences, both to find out what we have to say that may be of interest to them, and to discover what they can teach us about the nature of procedural rationality.

REFERENCES

Alfred V. Aho et al., *The Design and Analysis of Computer Algorithms*, Reading 1974.

Chester I. Barnard, *The Functions of the Executive*, Cambridge 1938.

G. S. Becker, "Irrational Behavior and Economic Theory," *J. Polit. Econ.*, Feb. 1962, *70*, 1–13.

——, "A Theory of Social Interations," *J. Polit. Econ.*, Nov./Dec. 1974, *82*, 1063–93.

F. M. Cancian, "Functional Analysis," in *International Encyclopedia of the Social Sciences*, 1968, *6*, 29–42.

R. M. Cyert and M. H. Degrott, "Sequential Strategies in Dual Control," *Theory Decn.*, Apr. 1977, *8*, 173–92.

Anthony Downs, *An Economic Theory of Democracy*, New York 1957.

Maurice Duverger, *Political Parties*, rev. ed., New York 1959, (*Les Partis Politiques*, Paris 1951).

W. Edwards, "Conservation in Human Information Processing," in Benjamin Kleinmuntz, ed., *Formal Representation of Human Thought*, New York 1968.

Richard T. Ely, *Outlines of Economics*, rev. ed., New York 1930.

S. Freud, "Five Lectures on Psychoanalysis" (originally "The Origin and Development of Psychoanalysis" 1910) in *The Complete Psychological Works of Sigmund Freud*, Vol. 11, London 1957.

George Homans, *Social Behavior: Its Elementary Forms*, New York 1961.

D. Kahneman and A. Tversky, "On the Psychology of Prediction," *Psychol. Rev.*, July 1973, *80*, 237–51.

Janos Kornai, *Anti-Equilibrium*, Amsterdam 1971.

Howard Kunreuther et al., *Protecting Against High-Risk Hazards: Public Policy Lessons*, New York 1978.

James G. March and Herbert A. Simon, *Organizations*, New York 1958.

Jacob Marschak and Roy Radner, *Economic Theory of Teams*, New Haven 1972.

John M. Montias, *The Structure of Economic Systems*, New Haven 1976.

J. F. Muth, "Rational Expectations and the Theory of Price Movements," *Econometrica*, July 1961, *29*, 315–35.

Nils Nilsson, *Problem-Solving Methods in Artificial Intelligence*, New York 1971.

A. Rees, "Economics," in *International Encyclopedia of the Social Sciences*, 1968, *4*, 472.

William H. Riker, *The Theory of Political Coalitions*, New Haven 1962.

—— **and Peter C. Ordeshook**, *An Introduction to Positive Political Theory*, New Jersey 1973.

Paul Samuelson, *Foundations of Economic Analysis*, Cambridge 1947.

George Simmel, *Soziologie*, Berlin 1908.

Herbert A. Simon, "A Formal Theory of the Employment Relation," *Enconometrica*, July 1951, *19*, 293–305.

——, "A Behavioral Model of Rational

Choice," *Quart. J. Econ.*, Feb. 1955, *69*, 99–118.

———, "Dynamic Programming Under Uncertainty with a Quadratic Criterion Function," *Econometrica*, Jan. 1956, *24*, 74–81.

———, (1976a) *Administrative Behavior*, 3d ed., New York 1976.

———, (1976b) "From Substantive to Procedural Rationality," in Spiro J. Latsis, ed., *Method and Appraisal in Economics*, Cambridge 1976.

——— **and J. B. Kadane,** "Optimal Problem-Solving Search: All-or-None Solutions," *Artificial Intel.*, Fall 1975, *6*, 235–48.

G. J. Stigler, "The Economics of Information," *J. Polit. Econ.*, June 1961, *69*, 213–15.

——— **and G. S. Becker,** "De Gustibus non est Disputandum," *Amer. Econ. Rev.*, Mar. 1977, *67*, 76–90.

Oliver E. Williamson, *Markets and Hierarchies*, New York 1975.

National Academy of Sciences, (*NAS*) Coordinating Committee on Air Quality Studies, *Air Quality and Automobile Emission Control*, Vol. 1 summary rep., Washington 1974.