Journal of Economic Perspectives—Volume 5, Number 2—Spring 1991—Pages 67–88

Fallibility in Human Organizations and Political Systems

Raaj K. Sah

Human judgments are fallible. This often contributes to an acceptance of bad projects, policies, and ideas, and a rejection of good ones. One aspect of fallibility is that an individual typically can extract only part of the decision-relevant information from the limited raw data available. My emphasis here is on the time and effort that an individual spends on decision making, including that spent on formulating and choosing, from the large number of possibilities, the information to be extracted. Another aspect of fallibility arises from the importance of the timeliness of decisions. A decision may become less fallible in some ways if more time is spent on it, but the additional delay may also reduce its value. Yet another aspect of fallibility is that individuals have biases which may not be apparent to themselves or to others.

If the fallibility of any one individual were the only relevant issue, then, at least in principle, perfect organizational decision making might be achievable. If any one individual has only a part of the total available information, then, by suitably structuring the exchange of information among many individuals, it might be possible to have, in aggregate, all of the information. Likewise, although any one individual has a limited capacity to extract information from the raw data, it might be possible to overcome this limitation by dividing and recombining the total task among many individuals.

One rough analogy here is to a network of computers, each with limited capacity, such that many of them together can fully accomplish the computational task at hand. Another rough analogy is to the design of reliable systems

■ *Raaj K. Sah is Professor of Public Policy, University of Chicago, Chicago, Illinois.*

using unreliable components. A celebrated result in this field, the Moore-Shannon theorem, shows that if a component has some capacity to perform, then specific structures of components can be devised, by using a sufficiently large number of components, in such a way that the system is completely reliable (Harrison, 1965, pp. 255–62). It turns out that this theorem's result has little to do with the specific structures. It hinges on having an infinite number of components (Sah and Stiglitz, 1988, p. 468), which, to an economist, implies an infinite cost. Consequently, the theorem is not of much organizational use.

A more fundamental reason why the above arguments cannot be translated to organizations is that human communication is also fallible. Faultless communication of facts as well as of biases and nuances of judgment is typically not possible. Fallible communication distorts and alters information. It also includes the possibility of a deleterious mutual contamination of information. Consequently, even if there were no personnel costs involved in hiring individuals, it may still not be possible to achieve perfect organizational decision making, no matter how many individuals work together or how they are organized. Personnel costs make this situation even worse.

Thus, in the picture drawn here, both the fallibility of an individual and that of an organization are largely reflections of the scarcity of such resources as time (including the importance of the timeliness of decisions), effort, the number and types of employees (which translate into personnel costs), and the intrinsic abilities of heterogeneous individuals for decision making and communication. This picture is consistent with an individual's economizing behavior. Moreover, the nature of the fallibility of an organization, or of any multi-person entity, depends not only on the nature of individuals' fallibility but also on the organization's "architecture": that is, on those features of the organization which influence how the decisions of individuals are aggregated.

This paper attempts to present a perspective on some organizational consequences of human fallibility. It may be easier to get a flavor of the relevant issues by examining the role of fallibility in specific settings, rather than through abstract arguments. So, in the next three sections, I consider several different settings: the question of diversification versus concentration of political authority, the managerial succession process in organizations, and the choice of ideas and projects (including innovation-oriented projects) in organizations.[1] In the last section, I highlight some aspects of the approach underlying the analyses of human fallibility, in particular: the premises concerning an individual decision maker, the potential association between the motivation of

[1]The role of fallibility has been considered in several other settings. See, for instance, Von Furstenberg and Jeong (1988) on the sources of macroeconomic uncertainty, Klevorick *et al.* (1984) on the decision-making process of juries in legal trials, and Rose-Ackerman (1986) on appellate procedures in public bureaucracies that disburse entitlements. Also, models resembling some of those employed in the study of human fallibility can be found in such literatures as the engineering literature on safety monitoring devices and relay circuits, and the computer-science literature on fault-tolerant computing, distributed processing and error-correcting codes.

an organization's employees and their fallibility, and the nature and the aims of the analysis of organizations.

While focusing on human fallibility, I abstract in this paper from most other considerations. Apart from reasons of length, this is to complement the other papers in this symposium which deal with several other considerations. However, since some aspects of individual behavior emphasized in this paper may remind readers of the literature on bounded rationality (see Simon, 1979, for a lucid overview and references), I present in the last section some brief remarks on this literature.

Diversification versus Concentration of Political Authority

> The recognition of the insuperable limits to his knowledge ought indeed to teach the student of society a lesson in humility which should guard him against becoming an accomplice in men's fatal striving to control society . . . .
>
> Friedrich Hayek (1978, p. 34)

Plato's ranking of political systems, from best to worst, was: aristocracy, timocracy (governance by a soldier caste), oligarchy, democracy and tyranny (Bloom, 1968, Book VIII). The debate on the merits of alternative political systems that Plato initiated, though often polemical and tendentious, has been one of the longest running intellectual debates in history. The changes in Eastern Europe and the Soviet Union in the late 1980s have led some to suggest that such debates have been concluded.[2] However, this scenario is unlikely. For instance, an archetypal comparison in the last four decades has been between a political-economic system based on despotic party bosses versus one based on greedy capitalists and vile politicians. In contrast, one of the archetypal comparisons in the coming decades may be between a market economy that is closely guided by the government or by large corporations (for example, the post-War Japanese economy has been viewed as having been guided by elite government officials) versus one that is comparatively unguided (as many would say is the case in the United States).[3]

At a somewhat more basic level, an implicit debate is between those who postulate the pivotal role of "preceptors" in the governance of a society versus those who do not. (Here, "preceptor" is a shorthand for an individual who

[2]Fukuyama (1989) derives such a view from an unorthodox use of Hegelian dialectics. See Islam (1990) for a different view.

[3]Analogous comparisons and questions concerning the degree of centralization of government institutions have been important. For example: Is a regionally decentralized federal system, as in the U.S., better or worse than a more centralized one, as in France? Is it better to have several government organizations, rather than just one, provide the same services or deal with the same problems, even though this entails duplication and perhaps internecine competition?

exercises significant societal authority. This category includes elite bureaucrats, political leaders, and planners.) What has changed at different times in different societies is the backdrop of this debate and the actors who are claimants to the roles of preceptors. This debate is likely to continue.

Opposing views on this debate have been exposited and criticized from many perspectives. Popper's (1966) magisterial exposition and critique of the theories and prescriptions of Plato, Hegel and Marx is a leading example. However, the role of human fallibility does not appear to have been studied in this context. This section presents some arguments concerning this issue; see Sah (1989) for models and additional results.

Consider the following stylized framework. Total societal authority relating to such issues as which government policies are to be pursued, and the nature and the degree of government intervention, is divided among preceptors. A society is called "more centralized" or "less centralized" depending on whether there are few or many preceptors. (Within reasonable bounds, the precise numbers are unimportant here.) All preceptors are fallible, but some are more fallible than others. Also, before they are placed in positions of authority, the relevant differences in the abilities of preceptors can be observed only to a limited degree. The aggregate of the performances of all the preceptors in a period is a key determinant of the societal capital (that is, physical, human and other forms of capital) at the end of the period. This capital, along with the aggregate performance of all the preceptors in the next period, determines the societal capital at the end of the next period, and so on.

It follows, then, that a less centralized society has the advantage of a greater diversification of its performance across a larger number of preceptors. This is because diversification here dilutes the impact of the ability, or the lack thereof, of each preceptor on the aggregate societal performance. An analogy based on financial investment may be useful in bringing this point out. Think of the current societal capital as the initial wealth that is divided equally among a given number of investment managers (preceptors) who act independently of one another. An investment manager achieves a random return, but it is more likely to be higher if the investment manager is more capable. For the moment, assume that the investment managers are randomly selected from a given pool. Also, assume for now that there are no economies or diseconomies of scale in societal decision making. In the analogy of financial investment, this last assumption is represented as follows. The probabilities of different returns on the total wealth do not change if one investment manager is replaced by two with the same ability as the former, but with each managing half as much wealth.

Then, the wealth in any future period will have the same mean but a higher volatility under greater centralization. The difference in volatility will be more dramatic if there are large differences among the abilities of the investment managers (as perhaps is the case for preceptors), or if the difference in the number of investment managers in the two situations being compared is

large. Thus, for political systems, the largest difference in the volatility of performance will be between an absolute dictatorship with a single preceptor and a system in which political authority is widely distributed.

It is not suggested here that highly centralized societies cannot have very good performances. Such a society may get a preceptor like Lee Kwan Yu of Singapore or the late Chung Hee Park of South Korea, who have been viewed as having made substantial contributions to their societies. By the same token, such a society may get a preceptor like Idi Amin of Uganda, with correspondingly opposite consequences. Nor is it claimed here that the mean performance of more centralized and less centralized societies will be similar. What is suggested here is that, setting aside a number of considerations, an effect of human fallibility is that more centralized societies will have more volatile performances.

Gains from Coordination and Economies of Scale in Societal Decision Making

The possibility of such gains has been among the most enduring arguments in favor of centralized societies. These were, for instance, central to Lenin's (1932, p. 84) hope that the "whole of society would become a single office and a single factory." A key technical-economic argument in favor of the collectivization of Soviet agriculture was that a single authority making all decisions for production and for resource allocation, such as through machine tractor stations, would reap the otherwise unavailable advantages of coordination and of economies of scale, and that these advantages would more than offset the losses from weakened incentives.

While the presence of economies of scale in societal decision making may benefit a centralized society, there are also related drawbacks. Suppose that the future performance of an economy drops precipitously if its current societal capital falls below some threshold level. An example is the sequence of deleterious events that ensue from a substantial lowering of a country's credit rating in the international capital markets. These often include projects currently in the pipeline being shelved, credit becoming exorbitantly expensive for many future periods, and the domestic industrial and financial infrastructure being severely damaged. Such features will lead to a worsening, over time, of the relative performance of a more centralized society, because the probability of a very bad outcome is higher in this case. One bad dictator can, in a few periods, ruin the society to such a degree that the resulting lack of a threshold level of societal capital impairs performance for many future periods, even if there are significant economies of scale in societal decision making.

Likewise, while the potential for coordination may benefit a centralized society, there are countervailing effects as well. Since coordination is not possible without communication, and communication is fallible (a central premise in this paper), coordination beyond some point will be detrimental. Moreover, coordination almost always introduces delays. If each preceptor has to coordinate with several others, then many decisions must wait until other

decisions have been made and communicated. Such delays are perhaps costlier the more rapid the change in the economic environment.[4]

Merit-based Selection of Preceptors

Another argument in favor of centralized societies that has been historically important goes like this. In a more centralized society, more hinges on the ability of fewer individuals. Hence, a greater effort will be made to ensure that these individuals are of higher ability. This will improve the society's performance.

Such an emphasis on "merit-based selection" was a foundation of Chinese bureaucracy, which lasted more than two millennia, originating in the Western Han dynasty and enduring until its formal abolition in the early 20th century. This bureaucracy invented elaborate methods for selecting a very small group of individuals who were to exercise vast administrative authority. At times, emperors themselves administered and supervised the examination of candidates (Franke, 1960). A similar emphasis on merit-based selection provided a basis of the British cadres or "civil services" designated for colonial rule. The members of the elite civil services, always very small in number, were to be rigorously selected and then fully empowered to rule the colonies. Domestic variants and successors of these civil services, deriving their legitimacy largely from a merit-based selection, can be found today in most of the former colonies. The members of these services exercise far greater authority than their unelected counterparts in most modern democracies.

A system of merit-based selection of preceptors is obviously better than one based on bribery and nepotism. However, there are several reasons, apart from the direct public and private costs of such a system, that limit the extent to which it can improve the performance of centralized societies. Once a system has been in place for some time, and is thus well understood, a larger proportion of those who qualify are likely to do so less because of their intrinsic abilities and more because of the resources they spent on the coaching they needed to master the techniques necessary to qualify. More important, selection systems do not change as rapidly as does the mix of the characteristics of the preceptors that is most useful to a society. This may partly reflect inertia. It may also reflect the fact that no one has the ability to fully foresee the characteristics of preceptors that will be societally most valuable in the future. Again, this is a reflection of human fallibility. A less centralized society is partially insured against this problem because the larger number of preceptors are more likely to span a wider range of characteristics. Diversification here not only reduces risk in a standard, static sense, but it also has a deeper, longer-term advantage in an environment that is changing and is essentially unpredictable.

[4]These and other adverse consequences of coordination were largely ignored in the post-War theory and practice of economic planning in socialist and mixed economies, as well as in the literature on economic growth.

Some Empirical Implications

The perspective presented in this section can help interpret some aspects of the post-War development experience. For instance, the superior performance of a small number of countries with centralized political authority (like South Korea, Singapore, and, more recently, Indonesia) has often been presented as evidence that centralized political authority is conducive, if not essential, to rapid progress in the early phases of development. What is left out in this line of reasoning is that, during the last four decades, political authority has actually been highly centralized in the vast majority of developing countries (in most cases, dictatorships), and that, with a handful of exceptions of the type noted above, their performance has been poor. This overall pattern is in accord with the perspective presented here.

Moreover, this perspective may help explain some puzzles. For instance, since the October Revolution, there have been periods during which the performance of the Soviet economy was judged by outside observers to be remarkably good in comparison to historical and contemporary international standards. Given that the cumulative performance has turned out to be quite disappointing, it is not yet sufficiently understood why the Soviet economy performed as well as it did during particular periods. On the other hand, such a performance of a highly centralized economy, over time, is consistent with the perspective presented here.

The Succession of Managers in Organizations

The current set of managers in any organization must, at some stage, be replaced. A part of human fallibility is that the choices of successors are fallible. Thus, an organization's succession process has dynamic consequences for its performance; for example, the quality of the current managers is influenced by that of the past managers, and the current managers will affect the quality of the future managers. A natural question, then, is: how does the degree of centralization, in regard to the decision-making authority to appoint successor managers within an organization, affect the quality of the managers actually appointed?

Consider, first, an organization consisting of two managers. If one of these two managers chooses the successor for himself or herself as well as for the other manager, then, arguably, this organization is more centralized than one in which each manager chooses his or her own successor. Extending this further, consider an organization consisting of many managers, only some of whom appoint successors. The number of slots for which a particular manager appoints successors varies across managers. The total number of slots in the organizations under consideration is fixed. Then, one organization could be viewed as being more centralized than another if, in the former, a smaller

fraction of those who appoint successors have the authority to appoint a larger fraction of the total number of successors.

Now, suppose that individuals differ in their abilities, including their ability to choose able successors, and that these differences are ascertainable in advance only to a limited degree. Suppose that there are two types of managers: those with high and those with low abilities to select successor managers. Call them "good" and "bad" managers, respectively. Both types of managers may choose a bad successor for a given slot, but a good manager is less likely to do so than a bad manager. Then, an indicator of the managerial quality in an organization is its steady-state distribution of managerial abilities; that is, the probability density of having different numbers of good managers in the organization, in the hypothetical situation in which this density does not change over time.

This indicator shows that there is a greater dispersion in the distribution of managerial abilities in a more centralized organization (Sah and Stiglitz, 1990). The intuition behind this result is as follows. While a good manager choosing a large number of successors is likely to have a significant beneficial effect on the quality of managers, a bad manager placed in the same position has a correspondingly large adverse effect. Therefore, greater centralization induces a greater dispersion in managerial quality.

The relationship between the quality of managers and the output or the performance of an organization depends in part on how the managers are organized and what tasks they perform. Suppose that managers who are good or bad in choosing successors are also, respectively, good or bad in producing output. Suppose the output of an organization depends only on the number of good managers; that is, bad managers are an irrelevant but not a harmful input. Suppose that the extra output due to an additional good manager is positive, and that it decreases if the number of good managers increases. Then, an effect of the greater dispersion in the number of good managers in a more centralized organization is to depress its expected output. The expected value of the output is also depressed if there is some risk aversion.

The Choice of Projects and Ideas in Organizations

In this section, I discuss some of the trade-offs that arise from human fallibility in the choice of projects and ideas in organizations. The background for this discussion is provided by some stylized settings. I first consider a committee with fallible members. I then compare hierarchies and polyarchies, to be defined below, focusing on how they might accept or reject innovation-oriented projects.

Committees

Consider a committee with n members who must collectively decide to accept or reject various projects in a portfolio. The word "project" is a

shorthand here for such objects of decisions as investment opportunities and ideas to be pursued and implemented. It does not, however, refer to the choice of personnel, which was discussed earlier. The projects are of two types: "good" and "bad." Parts of the limited total available information about any particular project are randomly distributed across members. Communication among members is limited. Each member makes a yes-no decision on each project.

A committee member can make two types of errors: the error of rejecting a good project and that of accepting a bad project. Suppose that a member has some ability, though not perfect, to identity a project's quality. Thus, a member is more likely to accept a good project than a bad project. In addition, suppose that the committee members have identical abilities and that the chance that a member will make a particular type of error for a particular project is independent across members and has the same probability for all members.

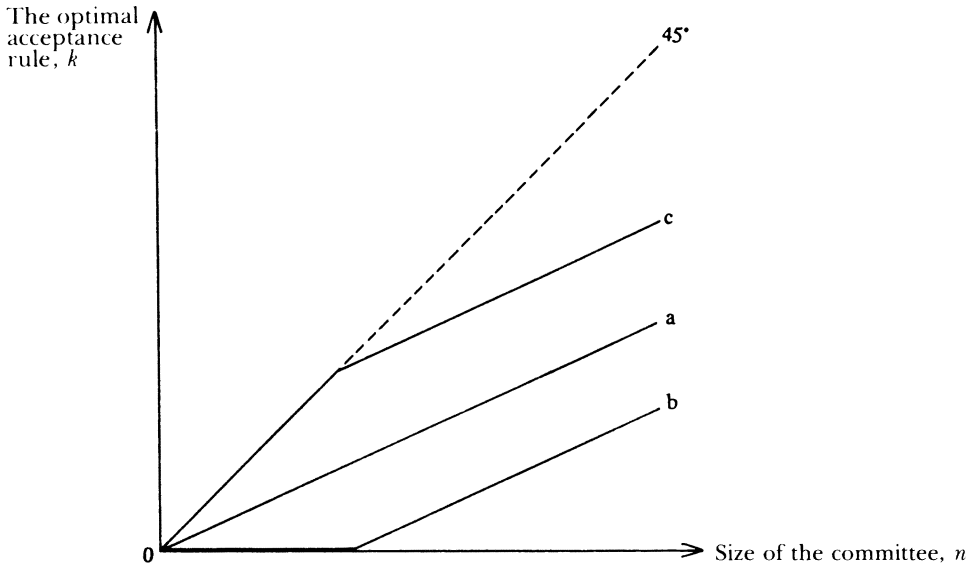
The quality of the portfolio of projects available to the committee is summarized as follows. The quality of a portfolio is called "average" if its expected value would be zero if all of the projects were accepted. An example of such a portfolio is one in which half the projects are good, half bad, and the gain from a good project is the same as the loss from a bad project. The quality of a portfolio is called "below average" or "above average" depending on whether its expected value would be negative or positive if all of the projects were accepted.

The committee accepts a project if k or more members accept it, and rejects it otherwise. Thus, the acceptance rule $k = n/2$ describes the majority rule. For a given n, a larger k implies a tighter acceptance rule. This decreases the committee's probability of accepting a bad project, but also increases the probability of rejecting a good project. A relevant question, then, is: which acceptance rule maximizes the expected value of the committee's decisions? This question has a crisp answer. It turns out that the optimal acceptance rule can be stated as an explicit closed-form formula expressed in terms of the parameters (Sah, 1990).[5] This formula also allows us to examine how the optimal acceptance rule will differ in different environments.

For example, Figure 1 depicts the case in which a committee member is as likely to reject a good project as to accept a bad project. The precise magnitude

[5]The conclusions presented in the text can be obtained directly from the formula described below; see Sah (1990) for its derivation and for additional results. Denote the fraction of good projects in the portfolio by α. Denote the net pay-offs from a good and a bad project by z_1 and $-z_2$ respectively, where z_1 and z_2 are positive. Define $e \equiv \alpha z_1 - (1 - \alpha)z_2$. The portfolio quality is called below average, average, or above average, depending on whether e is negative, zero, or positive. Define $\beta \equiv 1 - e/\alpha z_1$. Thus, $\beta \lesseqgtr 1$ if $e \gtreqless 0$. Let q_1 denote the probability that a committee member will reject a good project. Let q_2 denote the probability that a committee member will accept a bad project. Assume that $1 > 1 - q_1 > q_2 > 0$. Define $t \equiv (1 - q_1)/q_2$, $r \equiv q_1/(1 - q_2)$ and $K \equiv (\ln \beta - n \ln r)/\ln(t/r)$. Treat k and n as nonnegative integers (for brevity, this feature will be overlooked in Figures 1 and 2 presented later). Assume that $n \geq 1$. Then, $k = 0$ is optimal if and only if $\beta \leq r^n$, and $k = n$ is optimal if and only if $\beta \geq rt^{n-1}$. The interior values of optimal k are obtained as follows. If K is not an integer, then the only optimal value of k is the smallest integer larger than K. If K is an integer, then the only optimal values of k are K and $K + 1$.

Figure 1
The Optimal Acceptance Rule for a Committee in which a Committee Member Is as Likely to Reject a Good Project as to Accept a Bad Project



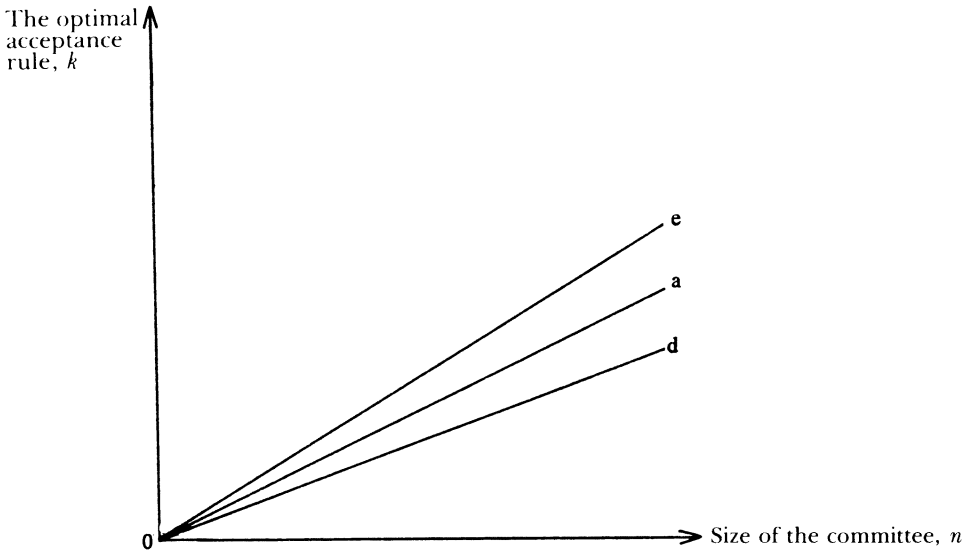
Note: Graphs a, b and c respectively depict the case in which the portfolio quality is average, above average and below average. The slope of line a is one-half.

of the probability of either error is irrelevant for the conclusions described here. If the portfolio quality is average, then, as line a shows, the optimal acceptance rule is half of the committee size. That is, the majority rule is optimal. If the portfolio quality is above average, then the line representing the acceptance rule shifts vertically downwards, as shown by line b. In this case, the optimal acceptance rule is less than half of the committee size. If the portfolio quality is very high and if the committee size is very small, then the optimal acceptance rule is at the extreme at which $k = 0$; that is, all projects should be accepted. If this is not the case, then the slope of the line representing the optimal acceptance rule remains unchanged at one-half. That is, the optimal acceptance rule increases by unity if the committee size increases by two.

Analogous conclusions hold for a below average portfolio, as shown by line c in Figure 1. The optimal acceptance rule is more than half of the committee size. If the portfolio quality is very low and if the committee size is very small, then the optimal acceptance rule is at the extreme at which $k = n$; that is, complete unanimity is optimal. Otherwise, the slope of the line representing the optimal acceptance rule is again one-half.

Figure 1 also shows that, unless an extreme acceptance rule is optimal, the optimal acceptance rule is slacker if the portfolio quality is better. This intuitive

Figure 2
The Optimal Acceptance Rule for a Committee Facing a Portfolio of Average Quality



Note: Graphs a, d and e respectively depict the case in which a committee member's probability of rejecting a good project is the same as, larger than and smaller than that of accepting a bad project. The slope of line e is less than one.

result holds even if the probabilities of a member's two types of errors have different magnitudes.

Next, compare two committees. In the first, a member is more likely to reject a good project than to accept a bad project. In the second, the situation is the opposite. We would expect the optimal acceptance rule to be slacker in the first committee. This is indeed the case if the portfolio quality is average. Figure 2 depicts this case. If a member is as likely to reject a good project as to accept a bad project, then, as before, the optimal acceptance rule is half of the committee size, as shown by line a. The line representing the optimal acceptance rule rotates clockwise, to line d, if a member is more likely to reject a good project than to accept a bad project. The opposite is the case if a member is less likely to reject a good project than to accept a bad project, as is shown by line e. Once again, these conclusions do not depend on the precise magnitudes of the probability of either error.

An interpretation of the optimal acceptance rule is as follows. Here, if a project is desirable at a given number of approvals, then it is also desirable at a higher number of approvals. For any given number of approvals that a project has received, we can calculate the posterior probability that this project is good; the corresponding prior probability is simply the fraction of good projects in

the portfolio. Thus, the optimal acceptance rule is the smallest number of approvals at which the posterior expected profit from a project (that is, the expected profit calculated using the posterior probability) is positive. A change in a parameter will, in general, affect this smallest number. For example, suppose one committee faces a better project portfolio than another. Then, the former committee's posterior expected profit will be larger at each possible number of approvals. Hence, the smallest number of approvals at which its posterior expected profit is positive cannot be larger. Consequently, its optimal acceptance rule cannot be tighter.

The trade-off on which the preceding analysis focused is that a tighter acceptance rule lowers the committee's probability of accepting a bad project, while raising that of rejecting a good project. Some of the other trade-offs that may be relevant are briefly noted here.

For example, in choosing the size of a committee, a trade-off arises between the higher personnel cost of a larger committee and the potentially improved quality of decisions of a larger committee. For a committee whose members choose what information to collect, as well a what individual decision rules to follow for accepting projects, there are trade-offs concerning how much individuals should spend on gathering information, what decision rules they should follow, and what acceptance rule the committee should use.[6]

Additional trade-offs arise in other, more complex settings. For instance, consider a multi-tier hierarchy in which each tier acts as a committee that decides which projects will be evaluated by the next higher tier. The final acceptance decision is made at the highest tier. A trade-off here involves how many tiers to have, of what sizes, and what acceptance rules to have for the various tiers. A larger number of tiers or a tighter acceptance rule at any tier implies an overall tighter screening of projects. Another trade-off in multi-tier organizations involves balancing the personnel cost of assigning individuals of different presumed levels of abilities to different tiers against the resulting quality of organizational decisions.

The Choice of Innovation-Oriented Projects

Technological innovation has been a principal source of rapid progress since the Industrial Revolution (see Mokyr, 1990, and references therein). The explanation of the differences in the pace of such innovation in different societies remains a fundamental question. Also, researchers have tentatively identified patterns of innovation that are noteworthy; for instance, the share of significant industrial innovations produced by small firms in the U.S. appears

[6]In typical committees, members vote more or less simultaneously. An alternative to this is sequential decision making in which each decision maker has some information concerning the antecedent stages at which decisions were made. This alternative may involve a different delay in the final decision, as well as a different set of errors in communication. Abstracting from such issues of comparative organizational analysis, Meyer (1989) has studied sequential decision making in which the organizational task is to choose one of two objects, given limited information and communication.

to be larger than their shares of sales, employment, or R&D spending in the industrial sector (Scherer, 1984, pp. 222–37). It is thus important to study decision making concerning the choice of innovation-oriented projects.

A feature of innovation-oriented projects (like R&D investments, and ideas for new products and technologies) is that the possible returns from them are much more dispersed than those from routine projects (like projects for manufacturing standardized products using standardized technologies). Truly extraordinary returns are possible from some innovation-oriented projects but not from virtually any routine project. On the other hand, the entire investment might be lost for a much larger proportion of innovation-oriented projects than for routine projects. Another feature of innovation-oriented projects might be called "unfamiliarity." For example, it is more difficult for two persons, even if they are experts, to communicate their information about an innovation-oriented project than about a routine project. This is partly because adequate past experiences and examples do not exist for most innovation-oriented projects against which individuals can easily calibrate their own assessments or using which they can easily communicate their views. Moreover, it is possible that individuals' biases play a greater role in their decisions if a project is more unfamiliar.

Hierarchies and Polyarchies. Consider two stylized architectures, a "hierarchy" and a "polyarchy," each consisting of two managers. In the hierarchy, each project is first evaluated by the lower-down manager; if approved, the project is evaluated by the higher-up manager, who finally accepts or rejects the project. In the polyarchy, each manager can undertake projects independently of the other, and a manager does not know whether a project under consideration has been rejected or even evaluated by the other manager.

This dichotomy attempts to capture the idea that, compared to a hierarchy, a polyarchy entails a greater independence in decision making, in the sense that a constituent unit can more easily accept a project without needing the approval of others. A related idea is that a project has more opportunities of being independently evaluated in a polyarchy. Of course, the details of the dichotomy can be altered. For instance, the constituent units of a hierarchy or a polyarchy may be organizations themselves, instead of managers. (Therefore, in the rest of this section, I will refer to a constituent unit simply as a "screen.") Or, there may be many screens in a hierarchy or a polyarchy, instead of just two. Also, it might be possible to assess the relative importance of hierarchical and polyarchical attributes in particular real-world institutions. For instance, a large corporation with a unified governance is more hierarchical than one that is divided into several independent divisions that can deal with the same product groups and can mutually compete. Similarly, an economy with very few venture capitalists is more hierarchical than one with many.

Now, consider the performance of the simple two-unit hierarchy and polyarchy described above (Sah and Stiglitz, 1986, 1988, offer models and results). First, suppose that a screen makes the same set of errors (that is, its probability of accepting a given project is the same) whether it is a constituent

unit of a hierarchy or a polyarchy. Then, a polyarchy will accept a larger fraction of all types of projects. This is because the two screens in a polyarchy together have a larger probability of accepting any given project than either of them alone, whereas the situation is the opposite in a hierarchy. Whether the expected profit is larger in a hierarchy or in a polyarchy will depend on the nature of the project portfolio and that of the errors of the individual screens. For instance, if there are two types of projects in the project portfolio, then, for a given portfolio quality and a given probability that a screen will accept a good project, a polyarchy has a higher expected profit, provided that the probability that a screen will accept a bad project is below a certain level. The rough underlying intuition is that if a screen's approval of undesirable projects is a less frequent problem than its rejection of desirable projects, then giving a second chance to a project (which is what happens in a polyarchy) is better than introducing a second hurdle (which is what happens in a hierarchy).

Next, consider the decision rules that the individual screens might use for accepting projects. Suppose a screen observes the net profit from the projects plus a scalar noise that has zero mean. Then, a screen in a polyarchy will typically use a more restrictive decision rule. For instance, the lowest observed profit level at which a screen in a polyarchy will accept a project to maximize its expected profit will typically be higher than what a hierarchy will set for its screens to maximize its expected profit. This is intuitive. A screen in a polyarchy knows that, unlike in a hierarchy, the project it approves has no further hurdle to cross. A screen in a polyarchy also knows that it faces a project portfolio that is worse than the original project portfolio because some of the projects it contains are those rejected by the other screen. Taking these respective decision rules into account, which will perform better, a hierarchy or a polyarchy? Some particular cases of this problem have been analyzed; in these, the expected profit is higher in a polyarchy if the observations are very noisy (Sah and Stiglitz, 1986).

The performance of different architectures when the unfamiliarity feature of innovation-oriented projects is taken into account is, I believe, an important issue. It seems reasonable to conjecture that the presence of unfamiliarity may induce individual decision-making units to exhibit very high probabilities of rejection of very good projects, and, as a result, a highly hierarchical choice of innovation-oriented projects may hinder innovation. For instance, almost all of the major companies in the office equipment business when reprographic technology became feasible failed to see an attractive market potential for it; Xerox was an exception in this regard. Another example, closer to economists, is that of the watershed paper of Black and Scholes (1973) on the pricing of derivative financial assets. This paper was originally rejected by the *Review of Economics and Statistics* and the *Journal of Political Economy*; it was eventually accepted by the latter. Many cases like these can be found. In most such cases, the unfamiliarity feature of the projects may very well have played a central role in the fallibility of individual decision making units, and, if so, highly

hierarchical architectures (for example, only one company in the office equipment business or only one economics journal) would have hindered innovation.

Some Related Issues. Several issues in this area, including the following, are potentially important but have yet to be investigated. First, a fine division of labor within organizations, often increasing over time, is a primary vehicle of productivity in routine production. However, the division of labor in the evaluation of innovation-oriented projects may produce adverse effects that are not well recognized. For instance, the evaluation of projects is often broken down into functional components (such as technical, marketing and financial), and also into components based on whether the evaluation is more quantitative, which is usually done at the lower tiers of an organization, or more qualitative, which is usually done at the higher tiers. If the decision making process is highly hierarchical (e.g., if each functional group has veto power to reject a project), then the gates that an innovation has to pass through may be undesirably narrow.[7]

Second, since the architecture determines the acceptance probabilities of various types of projects, it may in the longer run also affect the distribution of project types that is generated within the economy. If unfamiliar innovation-oriented projects are almost surely rejected in an economy, then inventors and research scientists are unlikely to come up with many such projects, no matter how large the potential for profit. This is partly for pecuniary reasons, but also because the implementation of an idea is by itself often an important part of the reward for an inventor.[8]

Some Characterizations of Human Fallibility

> [*S*]*cience is fallible, because science is human.*
> But . . . it does not follow that the choice between theories is arbitrary, or non-rational: . . . that our knowledge cannot grow.
> Karl Popper (1966, Vol. II, p. 375)

This section highlights some aspects of the approach underlying the analyses of human fallibility. It describes the premises concerning an individual decision maker, and then discusses the potential association between the motivation of an organization's employees and their fallibility. The section concludes with some brief remarks on the nature and the aims of the analysis of organizations.

[7]See Noyce (1978) for the experiences of a founder and later chairman of Intel with such outcomes and their consequences.

[8]Another issue is the effect of R&D and product market competition on the decision rules of the constituent units within an architecture. Due to these effects, a constituent unit has to take into account others' responses; see Stiglitz (1989). Bull and Ordover (1987) have studied a related problem, namely, the choice of the number of managers by fallible firms facing product market competition.

Some Premises Concerning an Individual Decision Maker

The ongoing lack of concrete descriptions and models of bounded rationality has often been noted by researchers (Kreps, 1990, p. 773; Lazear, this issue), even though this term has been in use for more than three decades. A possible source of this lacuna is that the literature on bounded rationality has often defined individual behavior by way of contrast, or opposition, to particular versions of neoclassical economics, rather than by defining it directly. For example, in Simon's (1987, p. 222) words: "The term 'bounded rationality' has been proposed to denote the whole range of limitations on human knowledge and human computation that prevent economic actors in the real world from behaving in ways that approximate the predictions of classical and neoclassical theory: including the absence of a complete and consistent utility function for ordering all possible choices, inability to generate more than a small fraction of the potentially relevant alternatives, and inability to foresee the consequences of choosing alternatives . . . ."

To avoid this lacuna, analyses of human fallibility have attempted to spell out directly the premises concerning an individual decision maker. For example, a premise here is that an individual can typically extract only a part of the decision-relevant information from the limited raw data available. This is partly because the individual must spend time and effort in formulating and choosing (usually implicitly), from the large number of possibilities, the models of causation and the hypotheses to be explored and tested. The emphasis here is on the scarcity of time and effort. This scarcity would remain important in most situations even if, contrary to reality, unlimited raw data were costlessly available to the individual. In contrast to the literature on bounded rationality (Simon, 1979), the emphasis here is not on the constraints on an individual's raw computational power. Nor is the emphasis on the constraints on an individual's data storage and retrieval capacity, as in some recent interpretations of bounded rationality; see Lazear (this issue) for some remarks on such an interpretation. These constraints will become increasingly less important, though not necessarily trivial, due to advances in information technology.

The time spent by an individual on a decision has several different features that contribute to its fallibility. First, time has a resource cost, as was noted earlier. Second, the object of decision itself may change if the decision is delayed, thereby reducing the value of the earlier effort. Third, a delayed decision usually means that the circumstances that the individual faces, including others' actions, change.

Biases. Another premise concerning individuals is that they have biases. Different individuals often look at different parts of the data, even though they share the same objective. For instance, Robert Solow and George Stigler, both extraordinarily able economists, may have markedly different opinions on even a relatively narrow question, although they both have access to the same data and knowledge. This is not easily explained in terms of differences in objectives. An individual may not even be aware of many of his or her biases. An example

of such a bias is the overestimation by many individuals of their own abilities; Adam Smith called this the "overweening conceit"; see Rosen (1986) for some thoughts on this.

The prevalence of decision-relevant biases has been widely recognized, though the etiology is still quite rudimentary; see Nisbett and Ross (1980) for an analysis and extensive references. There have also been some attempts, though far fewer than needed, to examine the effects that biases have on organizational and economic decisions. Lazear (this issue) has begun exploring the question of whether or not some common practices in business organizations (like the use of bonuses versus penalties in different situations) are in part responses to the nature of individuals' biases. Akerlof (1990) has analyzed the role of the "salience" of costs and benefits in individual decision making. George (1980, Ch. 3) has studied the role, often decisive, of the biases of American presidents and their advisors in foreign policy decisions.

Procedural Aspects of Decision Making. The work of Kahneman and Tversky, especially that which proposes some heuristic procedures as descriptions of individual decision making under uncertainty and which points out some resulting biases, is familiar to many economists (Kahneman, Slovic and Tversky, 1983). It turns out, however, that researchers on decision heuristics have found a large number of different models and heuristic procedures. Very little is known about whether or not there is a small number of basic heuristic procedures and, if not, which procedure is used for which task in which context and why. Also, the decision-making tasks that are presently considered in this research are typically much simpler than those faced by individuals in organizational settings (see Rutherford, 1988, for some references on this literature).

An understanding of the procedural fundamentals of decision making is important. At the same time, analyses of the consequences of human fallibility are not entirely contingent upon this understanding. For example, the perspective presented in an earlier section on the diversification of political authority is largely invariant with respect to which heuristic procedure an individual preceptor might be using. More generally, many important questions of organizational study and of positive economic analysis that relate to human fallibility can be fruitfully explored based on a descriptive understanding of the patterns of human decision making in different organizational contexts and of the important types of biases that arise in organizational decisions.[9]

Fallible Communication. A different premise concerning individuals, but now in an interpersonal setting, is that their communication is fallible. An individual's capacity for articulation and reception is limited. Communicating facts as well as biases and judgments, with full precision and comprehensiveness, is difficult and prohibitively expensive in any language.[10] The verbal and

[9]See, for instance, Janis (1989) for a description of the patterns of executive decision making during crises.

[10]The fact that all natural languages have only a limited capacity for precision, and a large scope for ambiguity, may itself reflect an economizing principle.

written methods and tools with which individuals can access each other's information and knowledge, themselves have marked effects on the messages, both as sent and as received. Moreover, some important forms of communication may be more revealing because they are more imprecise. For example, most managers insist on face-to-face meetings. At the heart of these meetings are such imprecise forms of communication as gestures, signs and symbols. Nevertheless, important "deals" are made or unmade on the outcome of face-to-face meetings among principals or senior managers, often overriding the detailed analyses of facts that their supporting staff might have prepared. Organizational economists have not adequately explained why this sort of communication is so important.

Fallible communication distorts and alters information. The original message may get replaced by an altogether different but no less plausible message, especially if the communication involves several intermediaries. Moreover, communication may lead to a deleterious mutual contamination of opinions. For instance, when individuals are asked to interpret some data and are allowed to communicate with one another, they tend to distort their own interpretations in order to conform to other group members. This may happen even when there are no pecuniary gains from conformity and conformity leads to faulty interpretations.[11] Further, the nature of the group that is communicating, and the context of the communication, may themselves have systematic effects on the nature of the mutual contamination; see Janis's (1982) analysis of "groupthink" in cohesive groups. The possibility of the unintended contamination of information due to communication is also recognized in other contexts. For instance, in American presidential elections, the results of exit polls on the east coast are believed to influence voting behavior on the west coast, leading many public interest groups to plead that the media should delay the announcement of these results.

Motivation and Fallibility

A potential device for motivating an organization's employees consists of contracts through which individuals' compensations may be linked to the outcomes of their activities. The theory of such contracts for the principal-agent relationship has recently been extensively studied.[12] Existing studies have typically focused on characterizing optimal contracts, taking the structure of the relationship as given (mostly, but not always, one principal and one agent), rather than on comparing alternative structures or on endogenously determin-

[11]See Campbell (1958) and references therein on the early social psychology literature on errors in communication. See Coleman (1990, pp. 219–20) for a related discussion.

[12]See Hart and Holmstrom (1987) and Sappington (in this issue) for reviews. See Holmstrom (1989) for a discussion of the role and the limitations of principal-agent type contracts in the context of innovation. Koh (1989) examines some aspects of principal-agent type contracts in models that deal with human fallibility.

ing the optimal structure. The intra-organizational incentives that might result from such contracts are a part of what Williamson (1985) has called internal "low-powered incentives," in contrast to the "high-powered incentives" that arise in markets. An issue that the principal-agent literature is currently attempting to resolve is that, while the terms of contracts that existing models typically predict are not only complex but also extremely sensitive to the precise description of important informational variables, the terms of contracts that are observed in practice are usually simple (Hart and Holmstrom, 1987, pp. 105–6; Stiglitz, 1987, pp. 970–1). More important, many unresolved questions have been opened up by the recent empirical studies that show that the link between the performance of corporations and the rewards received, directly or indirectly, by their CEOs is extremely weak (Baker, Jensen and Murphy, 1988; Jensen and Murphy, 1990).

Suppose the organization's task under consideration is the evaluation and selection of innovation-oriented projects. Then, several features of such projects, in addition to those noted in the last section, may render typical principal-agent models less applicable than they would be in situations in which the organization's task is routine production (actually, many principal-agent models were initially developed to analyze a relatively routine task, namely, sharecropping). For example, a feature of innovation-oriented projects is that the outcome of a project undertaken by the organization is not known for a much longer period than that of routine production. Gestation lags of 15 years or more may arise for projects attempting to invent new drugs. By the time the outcome of an innovation-oriented project is known, many of those involved in the project's acceptance decision would have retired or left the organization, while other activities would have been undertaken that were not a part of the original project but that have impinged on the outcome of the project. These features increase the difficulty of inferring from the final outcome the levels of effort that were exerted by the individual members of the group that initially evaluated and approved the project. Thus, the motivation that can be created through outcome-based compensation contracts in these cases is likely to be lower-powered than what might be created in routine production.

These problems become even more acute for the projects that are rejected by the organization. The organization does not have in-house information on the quality of the project that it did not undertake. Other organizations also may not have this information because, for example, they may undertake the project in a modified version, or several periods later in changed circumstances. Moreover, whatever information other organizations have will typically be inaccesible even if it is potentially useful.

The preceding discussion suggests the importance of another organizational device, namely, the voluntarily agreed-upon supervision of one set of individuals by another, for modifying the nature and the intensity of the activities of employees who have different attributes and motivations. Though

this device is extensively used in practice, it has not received as much economic study as it deserves.[13] In particular, the strengths and weaknesses of supervision, compared to other organizational devices, in the context of innovation-oriented projects is a virtually unresearched topic.

Finally, it is also important to emphasize that human fallibility does not disappear even if there are no intra-organizational motivational problems. If it did, then self-employed individuals would be infallible in their decision making. The absence of intra-organizational motivational problems can improve the quality of effort associated with each unit of time that an individual spends on decision making. However, the individual still has to contend with such resource problems as the cost of time, the importance of the timeliness of decisions, and his or her limited intrinsic ability for decision making. The point here is not that motivation does not matter, but that, beyond some moderate level, additional motivation may have a weak effect on the nature of an individual's fallibility, even if it could be created within organizations at a reasonable cost.[14]

Some Remarks on the Analysis of Organizations

It should be apparent from the analysis in the preceding sections that the nature of the fallibility of an organization, or of any aggregate entity, depends not only on the nature of the fallibility of the individuals who constitute the organization but also on the organization's architecture. Which aspects of the architecture are important depends on the question at hand and the context. In general, such aspects may include coordination, communication and reporting relationships. In any event, it will usually not be appropriate to achieve the transition from an individual to an organization (say, a firm), as a unit of analysis, by attributing to the organization the same behavioral properties that are attributed, as primitive assumptions, to the individual. Instead, the organization's behavior will need to be derived as a consequence of its architecture.

The literature on bounded rationality has yielded and inspired many contributions (March, 1978, and Rutherford, 1988, provide several references). At the same time, some parts of this literature appear to be aimed at criticizing particular versions of the neoclassical theory (see, for example, Simon, 1979). In contrast, the analyses of human fallibility build on the contemporary economics of organization and economics of information. Descriptive models with different emphases and different degrees of detail are considered useful in exploring different questions, and the models are consistent with an individual's

[13]See, however, Stiglitz (1975) for some early notes, Holmstrom and Tirole (1989, pp. 110–3) for a brief review, and Rosen (1982) for a study in which the supervision needs within organizations are proposed as a possible explanation for some observed patterns of managerial earnings.

[14]The position that Nisbett and Ross (1980, p. 12) take on the limited efficacy of motivational factors in related contexts appears to be stronger than the one just stated: "We argue that many phenomena generally regarded as motivational... can be better understood as products of relatively passionless information-processing errors than of deep-seated motivational forces."

economizing behavior. The aim in this case is not criticism, but rather an attempt to present a perspective that may complement the existing frameworks and that may generate some insights, if not positive predictions. This does not reflect the view that criticism has little value; indeed, it is important. Instead, it reflects a bias that criticism, no matter how persuasive, is perhaps less valuable than even a limited attempt to extend the repertoire of approaches for understanding what is observed.

■ *I thank James Heckman, Sherwin Rosen, Alan Schwartz and Jitendra Singh for discussions and Jacques Cremer, Carl Shapiro, and Timothy Taylor for comments on an earlier version of this paper. Some parts of this paper are based on joint research with Joseph Stiglitz, to whom I am deeply indebted.*

References

Akerlof, George A., "Procrastination, Obedience and Authority," working paper, University of California, Berkeley, 1990.

Baker, George P., Michael C. Jensen, and Kevin J. Murphy, "Compensation and Incentives: Practice vs. Theory," *Journal of Finance*, July 1988, *43*, 593–616.

Black, Fischer, and Myron Scholes, "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy*, May/June 1973, *81*, 637–54.

Bloom, Allan, trans., *The Republic of Plato*. New York: Basic Books, 1968.

Bull, Clive, and Janusz A. Ordover, "Market Structure and Optimal Management Organizations," *RAND Journal of Economics*, Winter 1987, *18*, 480–91.

Campbell, Donald T., "Systematic Error on the Part of Human Links in Communication Systems," *Information and Control*, December 1958, *1*, 334–69.

Coleman, James S., *Foundations of Social Theory*. Cambridge: Harvard University Press, 1990.

Franke, Wolfgang, *The Reform and Abolition of the Traditional Chinese Examination System*. Center for East Asian Studies, Harvard University. Cambridge: Harvard University Press, 1960.

Fukuyama, Francis, "The End of History?" *National Interest*, Summer 1989, *16*, 4–18.

George, Alexander L., *Presidential Decisionmaking in Foreign Policy*. Boulder: Westview Press, 1980.

Harrison, M. A., *Introduction to Switching and Automata Theory*. New York: McGraw-Hill, 1965.

Hart, Oliver, and Bengt Holmstrom, "The Theory of Contracts." In Bewley, Truman, ed., *Advances in Economic Theory*. Cambridge: Cambridge University Press, 1987, pp. 71–155.

Hayek, Friedrich A., *New Studies in Philosophy, Politics, Economics and the History of Ideas*. London: Routledge and Kegan Paul, 1978.

Holmstrom, Bengt, "Agency Costs and Innovation," *Journal of Economic Behavior and Organization*, December 1989, *12*, 305–27.

Holmstrom, Bengt, and Jean Tirole, "The Theory of the Firm." In Schmalensee, Richard, and Robert D. Willig, eds., *Handbook of Industrial Organization*. Vol. I. Amsterdam: North-Holland, 1989, pp. 61–134.

Islam, Shafiqul, "Capitalism in Conflict," *Foreign Affairs*, 1990, *69*, 172–82.

Janis, Irving L., *Groupthink*. Boston: Houghton Mifflin, 1982.

Janis, Irving L., *Crucial Decisions: Leadership in Policymaking and Crisis Management*. New York: Free Press, 1989.

Jensen, Michael C., and Kevin J. Murphy, "Performance Pay and Top-Management In-

centives," *Journal of Political Economy*, April 1990, *98*, 225–64.

Kahneman, David, Paul Slovic, and Amos Tversky, eds., *Judgment under Uncertainty: Heuristics and Biases*. Cambridge: Cambridge University Press, 1983.

Klevorick, Alvin K., Michael Rothschild, and Christopher Winship, "Information Processing and Jury Decisionmaking," *Journal of Public Economics*, April 1984, *23*, 245–78.

Koh, Winston T. H., "Human Fallibility and Sequential Decision Making: Hierarchy versus Polyarchy," 1989. (Forthcoming, *Journal of Economic Behavior and Organization*.)

Kreps, David M., *A Course in Microeconomic Theory*. Princeton: Princeton University Press, 1990.

Lenin, Vladimir I., *State and Revolution*. New York: International Publishers, 1932.

March, James G., "Bounded Rationality, Ambiguity, and the Engineering of Choice," *Bell Journal of Economics*, Autumn 1978, *9*, 587–608.

Meyer, Margaret A., "Learning from Coarse Information: Biased Contests and Career Profile," 1989. (Forthcoming, *Review of Economic Studies*.)

Mokyr, Joel, *The Lever of Riches: Technological Creativity and Economic Progress*. New York: Oxford University Press, 1990.

Nisbett, Richard, and Lee Ross, *Human Inference: Strategies and Shortcomings of Social Judgments*. Englewood Cliffs: Prentice-Hall, 1980.

Noyce, Robert N., "Innovation: The Fruit of Success," *Technology Review* (Massachusetts Institute of Technology), February 1978, *80*, 24–7.

Popper, Karl R., *The Open Society and Its Enemies; Volume I—The Spell of Plato, Volume II—The High Tide of Prophecy: Hegel, Marx, and the Aftermath*. Princeton: Princeton University Press, 1966.

Rose-Ackerman, Susan, "Reforming Public Bureaucracies through Economic Incentives?" *Journal of Law, Economics and Organization*, Spring 1986, *2*, 131–61.

Rosen, Sherwin, "Authority, Control and the Distribution of Earnings," *Bell Journal of Economics*, October 1982, *13*, 311–23.

Rosen, Sherwin, "Prizes and Incentives in Elimination Tournaments," *American Economic Review*, September 1986, *76*, 701–715.

Rutherford, Malcolm, "Learning and Decision-making in Economics and Psychology: A Methodological Perspective." In Earl, Peter E., ed., *Psychological Economics*. Boston: Kluwer, 1988, pp. 35–54.

Sah, Raaj K., "How Capable Do the Planners (Preceptors) Have To Be?" working paper, Yale University, 1989.

Sah, Raaj K., "An Explicit Closed-form Formula for Profit-Maximizing *k*-out-of-*n* Systems Subject to Two Kinds of Failures," 1990. (Forthcoming, *Microelectronics and Reliability*.)

Sah, Raaj K., and Joseph E. Stiglitz, "The Architecture of Economic Systems: Hierarchies and Polyarchies," *American Economic Review*, September 1986, *76*, 716–27.

Sah, Raaj K., and Joseph E. Stiglitz, "Committees, Hierarchies and Polyarchies," *Economic Journal*, June 1988, *98*, 451–70.

Sah, Raaj K., and Joseph E. Stiglitz, "The Quality of Managers in Centralized versus Decentralized Organizations," 1990. (Forthcoming, *Quarterly Journal of Economics*.)

Scherer, F. M., *Innovation and Growth*. Cambridge: MIT Press, 1984.

Simon, Herbert A., "Rational Decision Making in Business Organizations," *American Economic Review*, September 1979, *69*, 493–513.

Simon, Herbert A., "Behavioral Economics." In Eatwell, John, Murray Milgate, and Peter Newman, eds., *The New Palgrave*. Vol. 1. London: Macmillan Press, 1987, pp. 221–5.

Stiglitz, Joseph E., "Incentives, Risk and Information: Notes Towards a Theory of Hierarchy," *Bell Journal of Economics*, Autumn 1975, *6*, 552–79.

Stiglitz, Joseph E., "Principal and Agent." In Eatwell, John, Murray Milgate, and Peter Newman, eds., *The New Palgrave*. Vol. 3. London: Macmillan Press, 1987, pp. 966–72.

Stiglitz, Joseph E., "Incentives, Information, and Organizational Design," *Empirica*, 1989, *16*, 3–29.

Von Furstenberg, George M., and Jin-Ho Jeong, "Owning up to Uncertainty in Macroeconomics," *Geneva Papers on Risk and Insurance*, January 1988, *13*, 12–90.

Williamson, Oliver, *The Economic Institutions of Capitalism*. New York: Free Press, 1985.