Organization Science
Vol. 18, No. 1, January–February 2007, pp. 39–54
ISSN 1047-7039 | EISSN 1526-5455 | 07 | 1801 | 0039

informs®
DOI 10.1287/orsc.1060.0216

Two Faces of Search: Alternative Generation and Alternative Evaluation

Thorbjørn Knudsen
Department of Marketing, University of Southern Denmark, 55 Campusvej, Odense 5230, Denmark, tok@sem.sdu.dk

Daniel A. Levinthal
The Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania 19104, levinthal@wharton.upenn.edu

At its core, a behavioral theory of choice has two fundamental attributes that distinguish it from traditional economic models of decision making. One attribute is that choice sets are not available ex ante to actors, but must be constructed. This notion is well established in our models of learning and adaptation. The second fundamental postulate is that the evaluation of alternatives is likely to be imperfect. Despite the enshrinement of the notion of bounded rationality in the organizations literature, this second postulate has been largely ignored in our formal models of learning and adaptation. We develop a structure with which to capture the imperfect evaluation of alternatives at the individual level and then explore the implications of alternative organizational structures, comprising such individual actors, on organizational decision making.

Key words: organizational search; bounded rationality; organizational decision making

A central building block of the behavioral theory of the firm is the notion of bounded rationality (Simon 1955). In contrast to the optimizing agent of neoclassical economics, Simon offered the satisficing decision maker. The set of alternative actions are not presumed to be laid out in their entirety ex ante, but must be discovered or searched. This facet of the behavioral theory of the firm (March and Simon 1958, Cyert and March 1963, Nelson and Winter 1982) is by now well established. However, another critical facet of bounded rationality has been largely ignored in this tradition, and that is how alternatives, once identified, are to be evaluated.

Simon (1955) suggested that, rather than optimizing a utility function, individuals search for alternatives until they identify one that satisfies some minimum performance criteria—i.e., in his words, individuals engage in satisficing behavior. Central to this perspective is the view that choice alternatives are considered in a sequential manner and that the process of the sequential evaluation of alternatives stops well short of some latent optimal possible option. What is less salient, though considered in the original discussion, is how actors are to evaluate the proposed solutions or alternatives. How is an actor to know whether a given alternative in fact "satisfices" or not? Simon (1955) noted that there may be uncertainty as to whether a particular alternative yields a state of nature that is in the satisfactory set or not, but the text suggests that this indeterminacy may be resolved by identifying a new alternative that does not suffer this risk. However, this discussion points to an important lacuna in both this early and subsequent development of behavioral theories of individuals and firms.[1]

While ideas of search are central in behavioral theories of the firm (March and Simon 1958, Cyert and March 1963), the mechanisms by which these alternatives are evaluated are less clearly developed (Gavetti and Levinthal 2000). Models of adaptive search generally have the following characteristics. Some space of possible alternatives is sampled. The realization from this draw is then compared either to the current status quo action, or, in other cases, to an aspiration level (e.g., Levinthal and March 1981, Nelson and Winter 1982, Lant and Mezias 1990). When the space of alternatives constitutes attributes such as prices (Stigler 1961, Nelson 1970), assuming that quality attributes are well defined and equal, then the model does not seem to require any elaboration. However, consider other possible spaces of alternatives, such as the space of possible spouses or the set of possible new production technologies for a factory. When presented with a new alternative from one of these sorts of spaces, how is one to recognize a satisfactory solution when one is confronted with a proposed solution? When is one to stop searching for alternatives? Despite the neglect of uncertain evaluation in our models of adaptive search, it is an important characteristic of many task environments.

Another important gap in our formal models of search is that the work tends to be remarkably nonorganizational (but see March 1991, Lin and Carley 1997, Seshadri and Shapira 2003, Rivkin and Siggelkow 2003, and Siggelkow and Rivkin 2005 for important exceptions). While the label of organizations may be invoked, oftentimes the formal structure corresponds to a model of individual-level problem solving. We draw on Christensen and Knudsen's (2004) recent extension of the

Sah and Stiglitz (1986) characterization of organizational architectures to provide a framework with which we can consider the impact of alternative organizational structures on search processes.

As in the work initiated by Sah and Stiglitz (1986) and in the Theory of Teams (Marschak and Radner 1972), the notion of organizational form used in the present study is referring to a stylized representation of the flow of information among organization members. Thus, a *hierarchy* is a centralized information-processing structure and a *polyarchy* is a decentralized information-processing structure (Sah and Stiglitz 1986). Therefore, the terms *hierarchy* and (*de*)*centralization* are used here in a narrow sense of information-processing structures and do not fully incorporate issues of authority and control (Weber 1942) or, for instance, power and politics (Pfeffer 1981). While some progress has been made toward incorporating such considerations into formal models (cf. Dosi et al. 2003, Rivkin and Siggelkow 2003, and Siggelkow and Rivkin 2005), that is not the focus of the current work. Rather, we are concerned with the organization of individual evaluators and the nature of collective evaluation criteria.

At a basic level, evaluation of alternatives can suffer from two possible errors: *Type I errors* of rejecting a superior alternative and *Type II errors* of accepting an inferior alternative. As shown on the work on economic architecture (Sah and Stiglitz 1986, Christensen and Knudsen 2004), different organizational structures vary in their proclivity to make one type of error or the other. In particular, hierarchical structures, in which a proposal needs to be validated by successive ranks of the hierarchy in order to be approved, will tend to reduce the likelihood that an inferior alternative will be adopted—i.e., hierarchy reduces Type II errors. In contrast, what Sah and Stiglitz (1986) term polyarchies—a flat organizational structure in which approval by any one actor in a series of decision makers is sufficient for an alternative to be approved—will tend to minimize the probability of rejecting a superior alternative—i.e., polyarchy reduces Type I errors. Christensen and Knudsen (2004) provide a general graph-theoretic treatment of these structures that allows one to consider the full range of organizational architectures that range between these two extreme forms, and thereby allows one to specify structures that trade off these two types of errors as the relative degree of hierarchy and polyarchy shifts and, furthermore, to examine the change in the overall reliability of the organizational structure as the number of actors within the organization changes.

Using this analytical platform, we examine how alternative organizations of evaluators would move on a space of possible alternatives. In particular, we use the structure of fitness landscapes (Wright 1931, Kauffman 1993) to characterize a sense in which alternatives are more or less proximate to one another.[2] As characterized by Levinthal (1997), a process of local search is modeled as examining, at random, one of the adjacent points in the space of alternatives. The values of points in adjacent locations in fitness landscapes, as developed by Kauffman (1993), are correlated, with the degree of correlation being tuned by the intensity of the interdependencies among the N attributes that contribute to the fitness of a given alternative. Changing the level of interdependencies also has an impact on the overall structure of the landscape in that the number of local peaks increases with the degree of interdependencies (Kauffman 1993). The presence of local peaks poses particular challenges to a process of local search, because a decision maker at a local peak will be unable to identify superior alternatives that may be present on the broader landscape.

While the structure of fitness landscapes has been used to consider the issue of organizational adaptation (cf., Levinthal 1997, Rivkin 2000), as with much of the broader literature on search processes, the issue of how alternatives are to be evaluated has been underdeveloped. There have been some recent studies (e.g., Rivkin and Siggelkow 2003, Dosi et al. 2003, Ethiraj and Levinthal 2004) that examined how the allocation of decisions across actors within an organization and the degree to which actors make decisions based on the parochial concerns of their local subunit or the payoff to the broader organization affects the adaptive capabilities of an organization. However, in these analyses there is no uncertainty as to the payoff implications of the choices being made; rather, decision-making processes are affected by the perspective (local versus global, one subunit versus another) taken by the actors. Closer in spirit to our effort is Gavetti and Levinthal (2000), who contrasted evaluation that is *off-line*, in which assessment is done on the basis of actors' cognitive model of the fitness landscape, and *online*, in which the evaluation occurs subsequent to experience with the actual alternative.

We try to incorporate both faces of a search process: the sequential identification of alternatives and the uncertain evaluation of those alternatives that are identified. Whilst the first topic has been studied extensively in previous models of adaptive search, the second topic of uncertain evaluation has been rather neglected. We find that highly accurate evaluation of alternatives results in search processes being greatly influenced by the happenstance of the order in which alternatives are identified by the actor. As a consequence, highly accurate evaluators become trapped by their random starting positions in the landscape of alternative actions. Perfect evaluation leads to the rapid identification of a local peak and the persistence in that location across time. In contrast, imperfect evaluation leads to a more robust search process that is not as influenced by the happenstance of one's starting position in this landscape. Furthermore, we find that populations of moderately imperfect evaluators pay a surprisingly modest penalty in terms of the variability in their performance, either in a cross-sectional manner, or across time.[3]

We also consider, albeit in a rather stylized manner, how the structure of organizational evaluation of alternatives impacts these dynamics. Organizations that are hierarchical in structure, even if composed of imperfect evaluators, tend to replicate the conservatism of perfect evaluators and become trapped by local peaks. Hybrid forms, consisting of a mixture of polyarchy and hierarchy, effectively balance the dual imperatives of exploration and exploitation (Holland 1975, March 1991).

More generally, this work fits into a tradition of an information-based processing approach to contingency theory (Galbraith 1973, Burton and Obel 1984). Our analysis points to a three-way contingency among agents' screening ability, the nature of the problem environment as defined by the degree of interaction in the task environment, and the structure of decision making within the organization. Screening ability and organizational structure display an important degree of complementarity. The less able (or, conversely, the more able) individual evaluators are, the more attractive are organizational forms that tend toward hierarchy (polyarchy) as the hierarchical structure tends to compensate for the high error rates of less able individual evaluators (or, conversely, the variance induced by the polyarchy forms tends to compensate for the overly precise judgments of more able evaluators).

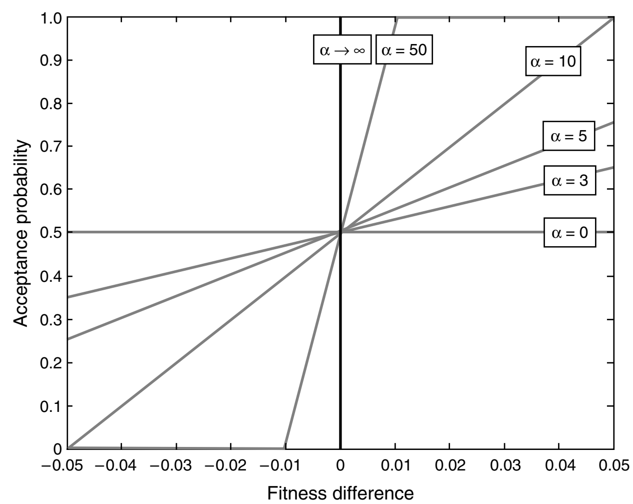
Model Structure

The model structure has three basic elements: the characterization of individual evaluation of alternatives, how individual evaluators are aggregated into an organizational form, and the specification of the task environment or the space of alternatives.

Individual Evaluation of Alternatives

Individual evaluators are characterized as being able to distinguish between a proposed action alternative and the status quo with more or less reliability. A perfect evaluator would, with certainty, distinguish between inferior and superior alternatives no matter how small the value differences are between two proposals. However, decision makers are unlikely to conform to such high standards. Actors are likely to make errors in identifying which of a pair of alternatives is, in fact, superior. However, one would expect that the likelihood of making a false classification is a decreasing function of the actual differences in value between the alternatives. That is, one may frequently misclassify pairs of alternatives that vary in payoff by only a small amount. In contrast, if the payoff to the two alternatives is substantially different, then the probability of making a misclassification, while not zero, would certainly be less than in the former case. Thus, misclassification is not a random process.

Figure 1 Six Levels of Screening Ability for an Evaluator, Ranging from Completely Random Screening ($\alpha = 0$) **to Perfect Screening** ($\alpha \to \infty$)



All evaluators are presumed to be intelligent in that they are more likely to favor superior alternatives; rather, they are simply assumed to vary in the precision with which they do this.

These properties are reflected in the screening functions represented in Figure 1. The horizontal axis indicates the actual difference in payoffs between a currently held alternative and a proposed alternative (current fitness minus new fitness), ranging from large negative differences in value to large positive differences in values. The vertical axis indicates the probability that an evaluator would accept the proposed alternative. Obviously, an intelligent screening function should have an upward slope, such that superior alternatives are more likely to be accepted than are inferior alternatives. In the extreme, with a perfect evaluator, the curve would have a point of discontinuity at zero, such that proposed alternatives with a payoff less than the current alternative (yielding a negative fitness difference) would be rejected with probability one and those with higher payoff (positive fitness difference) accepted with certainty.

We specify a family of screening functions $f(x)$ that takes the difference, x, of current fitness minus new fitness as an argument. The particular functional form used in the present work is a linear screening function, $f(x) = \alpha x + \beta$. The slope of the line, indicated by the variable α, can be interpreted as the screening capability of the evaluator. A steeper slope, or higher value of α, implies that the probability of accepting a proposal is more sensitive to changes in its actual value. The cutoff of the line, indicated by the variable β, can be interpreted as the bias of the evaluator's error (a higher value of β favors Type II errors of accepting an inferior alternative, while a lower β favors Type I errors rejecting a superior alternative). Within the class of linear screening functions, we restrict our attention to those that are unbiased in that if there is no difference in payoff between the proposed alternative and the current action (the value of fitness difference is zero), then the actor is equally likely to accept or reject the proposed alternative.[4] Thus, we have symmetric errors $\beta = 0$ and the screening function becomes $f(x) = \alpha x$. As α becomes arbitrarily large ($\alpha \to \infty$), the screening function approximates that of a perfect evaluator.

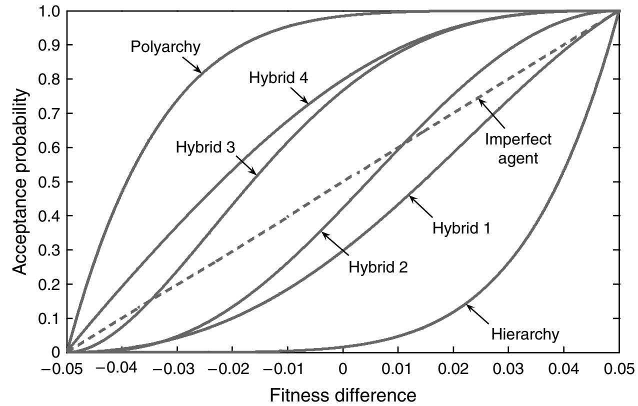
Aggregation of Individual Evaluators into an Organizational Form

Individual evaluators can be aggregated into more complex organizational forms. In particular, organizations can be characterized by the number of evaluators within them, but also more subtly by the nature of decision authority within them. Following Sah and Stiglitz (1986), we focus on whether or not a given actor has the authority to approve or reject a proposed alternative, or is merely authorized to pass the proposed initiative along within a broader chain of command. In particular, consider the flow of decisions in six distinct organizational forms shown in the appendix. In what we term a hierarchy, a proposal is initially considered by the evaluator at the far left in Figure A1. If the proposal is rejected by that evaluator, it is eliminated from further consideration (indicated in the figure by the dashed vertical line from that decision node). Alternatively, if the proposal is approved by that individual, then it is passed rightward to the next individual in the chain of command. A proposal is acted on only if it is positively screened by all six evaluators. The evaluator to the far right, effectively sitting at the top of this hierarchy, views only proposals that have been successfully vetted by the lower-level actors and has final say as to whether the organization adopts those proposed initiatives that reach his or her attention. In a hierarchical evaluation of a stack of job applications, for example, applications are eliminated at each level in the hierarchy and a diminished stack moves up to the next level. The evaluator at the top of the hierarchy will only see a very small stack of applications.

At the other extreme is the polyarchy structure. Here, proposed alternatives can be adopted by any of the six decision makers; an alternative is only dismissed if all decision makers in succession reject it. It is this contrast between the conservative (rejection-oriented) hierarchical structure and the acceptance-oriented polyarchy structure that Sah and Stiglitz (1986) considered. Christensen and Knudsen (2004) developed an analytical structure that allowed them to consider a wide range of hybrid forms that lie intermediate to these two extremes.[5] The appendix illustrates four intermediate forms that range from nearly hierarchical (Hybrids 1 and 2) to nearly polyarchical (Hybrids 3 and 4).

In analyzing the role of alternative organizational forms, we wish to distinguish between the effect of individual differences in screening ability and the impact

Figure 2 **Ability of Imperfect Evaluator Compared to Six Organizational Forms, Each Built of Six Identical Imperfect Evaluators**



Notes. We represent the case where $n = 6$ and $\alpha = 10$. The ability of the six hybrid organizational forms were derived according to the procedure shown in the appendix.

of the structure of the relationship among evaluators within the organization. Therefore, we treat organizations as being homogeneous in the screening ability of the individual evaluators that comprise the organization, although we examine the impact of varying this homogeneous level.

Figure 2 indicates the effective screening properties of six alternative organizational forms, all comprised of 6 evaluators with an α value of 10. It compares the screening function of an individual evaluator (indicated by the straight diagonal line) with the screening functions of six different organizational forms that are each built from a number (six) of such individual agents. Using methods outlined in Christensen and Knudsen (2004), we derive an organizational-level screening function, F, which is a mathematical representation of the flow of decisions in an organizational form (as shown in the appendix). In order to examine the effect of changing organizational structure, we assume that all members in an organization have identical abilities. That is, we assume that the individual-level screening function, $f(x)$, the probability that an individual accepts an alternative, is the same for all members of an organization. In this case, the organizational-level screening function, F, is a polynomial in the individual-level screening function, $f(x)$, i.e., F takes $f(x)$ as an argument.

The particular functional form of F represents an aggregation of information flows in the organization under consideration. The appendix shows the organizational-level screening function, F, for the six organizations studied in the present work. Notably, the organizational-level screening function, F, of the six hybrid organizations shown here will change in the same way relative to any screening function $f(x)$ at the individual level. Thus, the aggregation of an organizational form does not depend on the underlying individual-level screening function. Any given individual-level screening function will be changed in the way shown in Figure 2.

The organizational-level screening function of an n-member hierarchy is $F = f(x)^n$, and the organizational-level screening function of an n-member polyarchy is $F = 1 - (1 - f(x))^n$. For example, accepting an alternative in the six-member hierarchy requires that all six members accept the alternative. Therefore, the organizational-level screening function of the hierarchy (shown in Figure 2) is given by $F = f(x)^6$, which is the probability that this structure accepts the alternative in question. In a similar way, it is easy to see that the organizational-level screening function of the polyarchy (shown in Figure 2) is given by $F = 1 - (1 - f(x))^6$, i.e., the probability that at least one out of the six polyarchy members accepts an alternative. The screening functions of the four hybrids shown in the appendix are derived in a similar way.[6] As a point of reference, in Figure 2 we also include for comparison the evaluation function of a single perfect evaluator.[7]

The set of possible organizational forms is enormous, with the number of members ranging from $n = 2$ to infinity. We chose a set of structures with six organizational members—a number sufficient to generate a rich set of possible forms (there is a total of 223 topologically distinct graphs representing hybrid organizations with six members),[8] yet small enough to allow for an explicit representation of the individual organizational forms (see Figure A1 in the appendix). Among these, we picked the four hybrid structures that best spanned the topological (and functional) difference between the hierarchy and polyarchy. We see that the screening function of the polyarchy lies everywhere above the screening functions of alternative forms. This implies that the polyarchy for any given fitness difference is, as we suggested above, more prone to accept alternatives, even those with a negative value—a Type II error. That is, the polyarchy reduces Type I error at the expense of increasing Type II error. Conversely, hierarchies are very unlikely to mistakenly accept an inferior alternative (a Type II error) with the probability of acceptance being near zero for alternatives with a negative fitness difference. However, that same caution causes the hierarchical structure to reject many superior alternatives, i.e., alternatives with a positive fitness difference—a Type I error. Hierarchy reduces Type II error and increases Type I error. Interestingly, the screening function of hybrid forms will trade off the effects of polyarchy and hierarchy: Hybrids reduce both Type II and Type I errors (in particular, this effect is apparent in Hybrids 2 and 3).

Specification of the Task Environment

The final basic element of the model structure concerns the task environment in which evaluators (and organizations) operate. In prior work on the effect of alternative organizational forms on the effectiveness of screening alternatives, the process of alternative generation is treated as consisting of random draws from a fixed distribution of possibilities (Sah and Stiglitz 1986, Christensen and Knudsen 2004). However, research on organizational search processes (March and Simon 1958, Cyert and March 1963) has emphasized the spatial location of the set of possible alternatives, with the notions of local and distant search being central in theoretical (March and Simon 1958, Nelson and Winter 1982) and empirical (Stuart and Podolny 1996, Rosenkopf and Nerkar 2001) analyses. The imagery of spatial location is given clear expression in work on search in fitness landscapes (Levinthal 1997).

The generation of alternatives is not purely random, but is likely to reflect the availability of options in the neighborhood of the organization's current practices. Building on Levinthal (1997) and related work (Rivkin 2000, Rivkin and Siggelkow 2003, Dosi et al. 2003), the task environment of fitness landscapes (Kauffman 1993) is used to characterize a space of alternatives, where alternatives vary along any one of N dimensions and the correlation among distinct alternatives can be tuned by manipulating how interdependent these N elements are in determining the overall payoff.

If attributes of a policy contribute to performance in a relatively independent manner, the landscape of policy alternatives is relatively smooth; changing one attribute of the policy only affects the performance contribution of that attribute in isolation. In contrast, if policy attributes have a high degree of interdependence, then changing even just one attribute may have broader repercussions and affect the performance contributions of other attributes. As a result, a landscape of alternatives in which there is a high degree of interdependence will exhibit a relatively low level of correlation, with even modest shifts in attributes leading to a pronounced change in overall value. Related to this issue of degree of correlation among neighboring alternatives is the number of peaks in this performance landscape. With no interdependence among the policy attributes, there is a single peak in the landscape corresponding to the optimal setting of each of the individual policy attributes. As interdependence increases, the performance surface will exhibit local peaks of configurations of policy attributes that exhibit some degree of internal consistency (Kauffman 1993).

As regards alternative generation, we adopt the model used in much of the literature, i.e., randomly perturbing one of the N activities of the firm. This model provides a stylized representation of myopic local search by examining, at random, one of the adjacent points in the space of alternatives. While agents might perturb more activities, the properties of such longer jumps are fairly well known (Levinthal 1997). Thus, in order to focus on the issue of alternative evaluation, we limit our analysis of alternative generation to the simple model of myopic local search.

More formally, we specify alternatives as consisting of N attributes, $a_1, \ldots, a_N$. For simplicity, it is assumed that each attribute can take on two states. A performance landscape is a mapping of any possible vector of firm choices $A = (a_1, a_2, \ldots, a_N)$ to performance values $V(A)$. We create performance landscapes with a variant of the NK model (Kauffman 1993; see Sorenson 2002 for a review of these models in the organizations literature). The value of each individual attribute a_i is affected by both the state of that attribute itself and the states of a number of other attributes a_{-i}. Denote the value of attribute a_i by $c_i(a_i, a_{-i})$. For each landscape, the particular value of an attribute, c_i, is determined by drawing randomly from a uniform distribution over the unit interval, i.e., $c_i(a_i, a_{-i}) \sim u[0, 1]$. The value of a given set of alternatives A is then given by

$$V = [c_1(a_1, a_{-1}) + c_2(a_2, a_{-2}) + \cdots + c_N(a_N, a_{-N})].$$

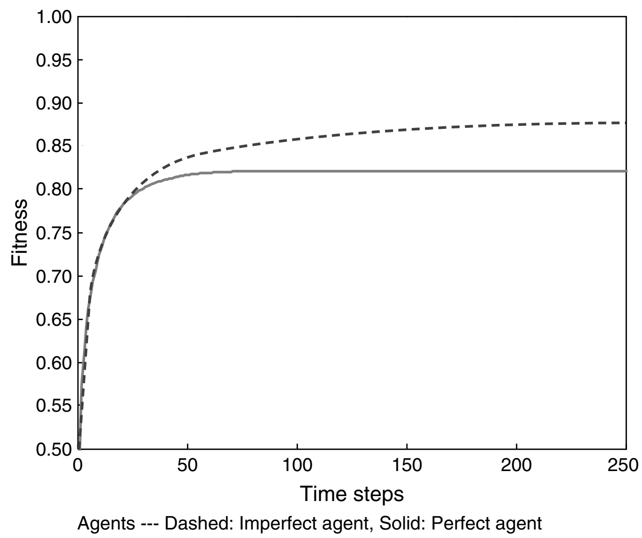
The identity of a_{-i}, i.e., the set of alternatives that affect each attribute a_i, is given by the interaction structure of the firm's decision problem (i.e., the variable K).

Analysis

To provide some initial understanding of the nature of the adaptive search process modeled here, we first consider the behavior of individual evaluators, and then, in the subsequent analysis, model the behavior of alternative organizational structures. All results reported here are based on simulations of landscapes with $N = 10$. We use $K = 3$ as a baseline case and, in a later section, expand the analysis to provide a comprehensive assessment of robustness by considering all levels of K $(0, 1, \ldots, N-1)$. Unless indicated otherwise, our results reflect the average of 100 entities searching on each of 100 distinct landscapes, resulting in 10,000 unique runs obtained from 100 different landscapes. Each of these landscapes has the same structure in terms of K, the degree of interdependence among attributes in contributing to performance, but represents a distinct draw on the common underlying probability-generating structure. At the beginning of each of the 10,000 runs, attribute sets are randomly assigned to the individual entities. In some simulations,[9] we analyze the behavior over a single, randomly specified performance landscape, but hold constant the number of unique runs at 10,000.

To enhance the comparison across these families of landscapes, we normalize the performance level on each surface so that average performance equals 0.5 and maximum performance equals one. That is, the crude fitness measure à la Kauffman is normalized in order to compare the results across different values of K. Using this normalized measure instead of the crude fitness measure does not alter the results in a qualitative sense.[10]

Figure 3 Fitness for Perfect Evaluator and Imperfect Evaluator



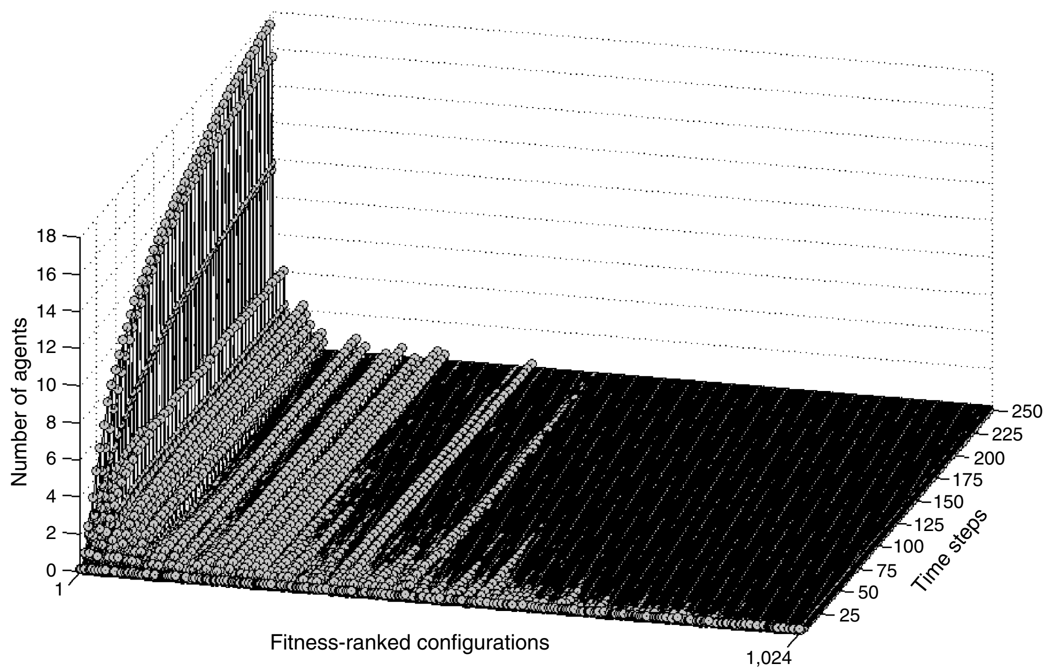
Notes. $K = 3$, $\alpha = 10$, 10,000 evaluators: 100 distinct landscapes with 100 evaluators on each.

Individual Evaluators

Figure 3 indicates the performance of two types of evaluators who vary according to the accuracy of their evaluation function. For the sake of a baseline comparison, we model one as being a perfect evaluator; in this setting, only alternatives that enhance the actual payoff will be accepted. In contrast, the other evaluator $(\alpha = 10)$ exhibits some intelligence in evaluation (i.e., $\alpha > 0$), with the probability of accepting a more favorable alternative increasing as a linear function of the performance increases associated with that alternative; however, this evaluator will at times mistakenly accept alternatives that in fact offer inferior performance, and in other instances will reject alternatives that could enhance the organization's performance (i.e., α is finite). We see that the perfect evaluator quickly asymptotes in the performance that is achieved, while the imperfect evaluator not only outperforms the perfect evaluator, but, if additional periods are examined, continues to exhibit modest but steady performance improvement. Perfect evaluation leads to the rapid identification of a local peak and the perfect evaluation function will lead the actor to maintain that position for the remainder of the simulation, while imperfect evaluation leads to persistence in search.

We would expect, however, that imperfect evaluation would suffer from two possible downsides. First, it is natural to expect that an imperfect evaluator would experience a slower rate of ascent in initial performance gains because an imperfect evaluator, by definition, will at times make downward moves.[11] Even though the perfect evaluator converges faster to the local optimum than does the imperfect evaluator, the difference in the initial rate of progress between the imperfect and the perfect evaluator is too slight to be visible in the comparison shown in Figure 3. However, around Period 40, we start

Figure 4 Distribution of Imperfect Evaluators at Each Peak During the Entire Run



Note. $K = 3$, 10,000 evaluators each on a single randomly selected landscape.

to see a divergence in the two performance curves as the performance of the imperfect evaluator continues on an upward gradient, while that of the perfect evaluator begins to asymptote. With less ability of the imperfect evaluator or larger values of N, the faster convergence of the perfect evaluator to a local optimum becomes more pronounced.[12]

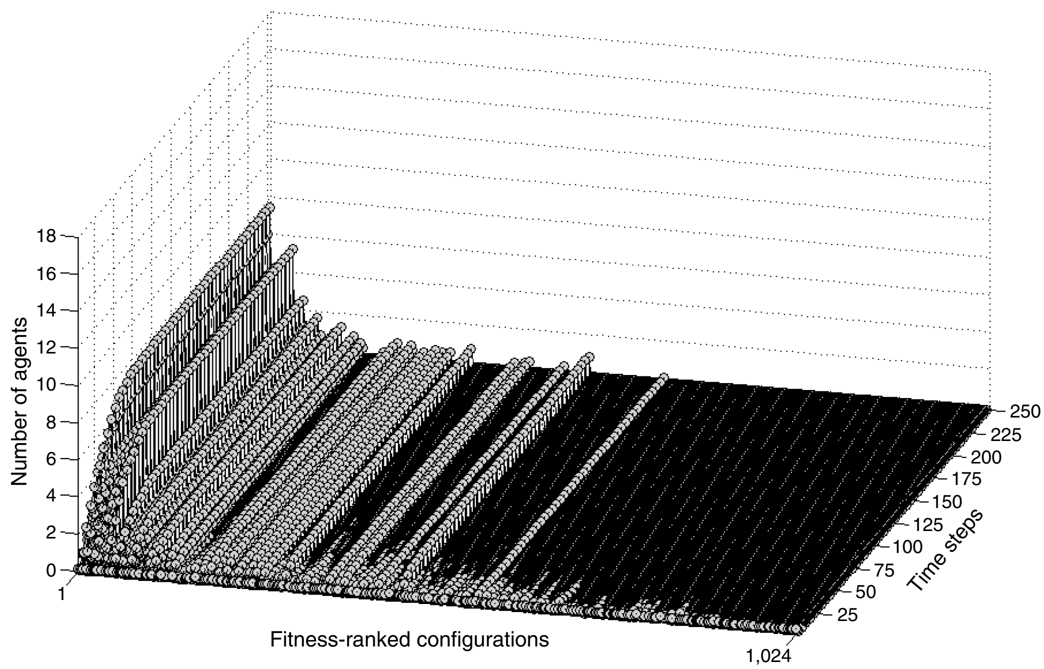
The other penalty that imperfect evaluation might exhibit is with respect to a limited ability to maintain, over extended periods of time, the attractive alternatives that have been identified. Given the noise in his or her evaluation process, even if a global peak is identified, there is a chance of mistakenly being seduced away from it by an alternative that appears superior. Figure 4 illustrates the emergence of the distribution of evaluators across the performance landscape, where the 1,024 distinct locations in the landscape are ranked ordered from 1 (the global peak) to 1,024 (the lowest value).[13] In the initial period, locations are randomly arrayed, and hence the distribution of evaluators across locations is quite flat. Rapidly, we see the emergence of clusters of evaluators aggregating on specific locations in the landscape. We see a particular massing of evaluators on the global peak, although given the imperfect evaluations, there is some dispersion of evaluators around this peak—a "cloud" of evaluators, as it were. In contrast, in Figure 5 with perfect evaluations, we see greater cross-sectional dispersion among evaluators as columns of evaluators aligned on distinct local peaks.

Figures 4 and 5 also provide more direct evidence regarding the ability of the two populations of evaluators to identify relatively attractive locations in the performance landscape. We see that the evaluators' locations across different performance levels are much more dispersed in the case of imperfect evaluators. Imperfect evaluators tend to mass at locales with the highest level of performance, but occasionally visit discrete locales associated with rather low levels of performance. While imperfect evaluators tend to mass among the highest-performing locations, there are fewer perfect evaluators at these locations. Perfect evaluators tend to spread out at intermediately performing locations.

Indeed, imperfect evaluators spread more unevenly than do perfect evaluators, with most locating near the highest-performing locations to a greater degree than do perfect evaluators, but with the right tail being thicker as well (i.e., there are more imperfect evaluators than perfect evaluators doing poorly). The tendency of imperfect evaluators to occasionally visit rather low levels of performance is a testimony to their very imperfection. The negative performance effect is offset, however, by a broadening of search that is also caused by the evaluator's imperfection. At the α value of 10 used in our baseline (and, more generally, at intermediate levels of α), the net effect is that imperfect evaluators, on average, tend to mass at the highest-performing locations, with occasional detours to low-performing places.

Building on this analysis, we can convey a more literal sense of the imagery of clouds of evaluators forming more or less tighter clusters of movement around different performance peaks in the landscape. Figures 4 and 5 indicate the number of evaluators in the two populations

Figure 5 Distribution of Perfect Evaluators at Each Peak During the Entire Run



Note. $K = 3$, 10,000 evaluators each on a single randomly selected landscape.

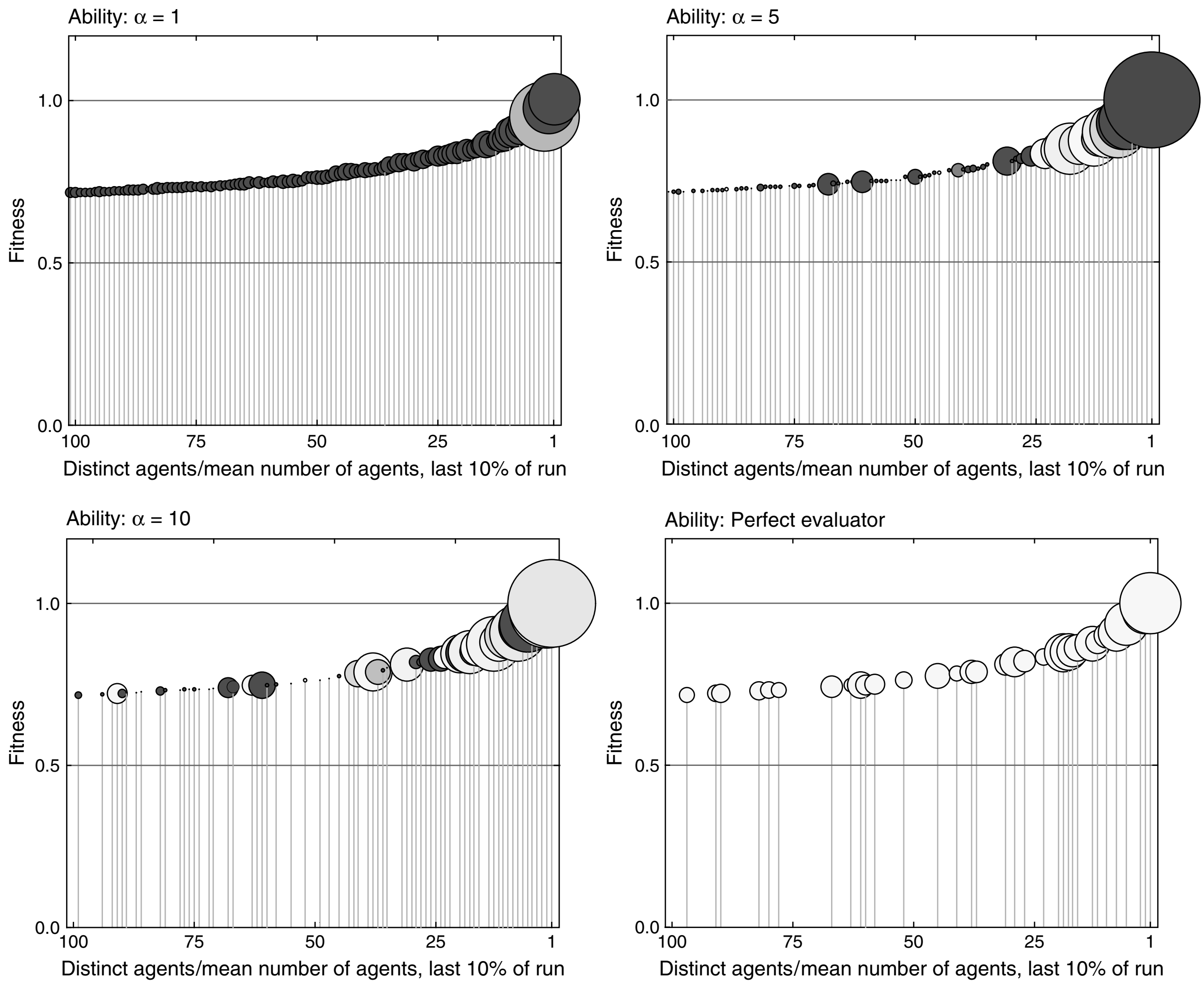
that mass at different locales, but these figures do not convey a sense of the instability or turbulence in the populations of imperfect evaluators. Imperfect evaluators are, on average, finding attractive locations in the space of alternatives, but are they wandering at some (possibly high) rate of velocity on the performance landscape?

Figure 6 creates a panel of images that conveys a sense of the dispersion of evaluators across locations in terms of the size of the clouds of evaluators massing at a location, and the flux in these constellations of evaluators. We measure flux, or turbulence, by the ratio of the number of different evaluators who visit a particular point in a given period of time (the last 10% of the run in this analysis) to the average number of evaluators at that location during that same time interval. In Figure 6, this measure of turbulence is indicated by color, with a darker color indicating a more turbulent setting. As the population of evaluators cools down and this ratio approaches 1, we use an increasingly lighter color. The size of the circle represents the number of evaluators at that location.

Highly imperfect evaluators (α equal to one) lead to both a very diffuse population (the circles are numerous in number and tend to be of modest size) and a very turbulent structure, with many different evaluators visiting a given locale (the clouds are very dark). Moderately able evaluators (α values of 3, 5, and 10) result in evaluators clustering on superior locations, with the evaluators forming a few large constellations around the superior points in the performance landscape. Clearly, the degree of turbulence diminishes as the evaluators' evaluation ability increases, with the color of the clouds shifting from dark to light. However, we see that with populations of evaluators who are highly accurate the clouds tend to be *frozen*, in that there is little or no turbulence. Such freezing tends to result in a number of distinct constellations of evaluators, many of which are not associated with particularly attractive points in the performance landscape.

Thus, imperfect evaluators appear not to wander too far away from the attractive peaks that they identify. That is, the clouds of evaluators around the peaks stay tightly clustered. We do not often see a situation in which slightly inferior alternatives are adopted and then, from this new lower base, even more inferior alternatives are mistakenly adopted. It is certainly possible for evaluators to take such a two-step "walk" from an attractive peak, and on occasion they will do so. However, the fact that the screening process, while imperfect, is nonetheless intelligent (in that more favorable alternatives are more likely to be accepted than are less favorable ones) implies that mistakes (walks away from superior alternatives) will tend to be self-correcting. After accepting an inferior alternative that takes him or her away from an attractive peak, it is more likely that the subsequent move will be back to this same peak rather than a move that takes the evaluator even farther away from this location. Ironically, the evaluators with an evaluation mechanism more prone to error exhibit less cross-sectional variability than does the population of perfect evaluators. While it is true that perfect evaluators will exhibit greater reliability in terms of period-to-period location

Figure 6 Fitness-Ranked Distribution of Evaluators Last 10% of Run



Notes. $K = 3$, $\alpha = \{1, 5, 10\}$, and a perfect evaluator for each level of α: 10,000 evaluators (100 distinct landscapes with 100 evaluators on each). The 100 most fit organizational forms are shown. If there are fewer than 100 distinct organizational forms (as when $\alpha \geq 5$), then all forms are reported. The number of evaluators at a peak is represented by the area of the circle (scaled by square root). The height of each circle represents the fitness of the peak. Darker colors of circles represent higher turbulence at the peak measured by the ratio of the number of distinct evaluators to the average number of evaluators at that peak. The last panel illustrates the results with a perfect evaluator.

and performance differences, we see from Figure 6 that the degree of instability in the behavior of the imperfect evaluator is rather modest. Furthermore, this population convergence among the imperfect evaluators occurs at the highest-performing locations in the landscape.

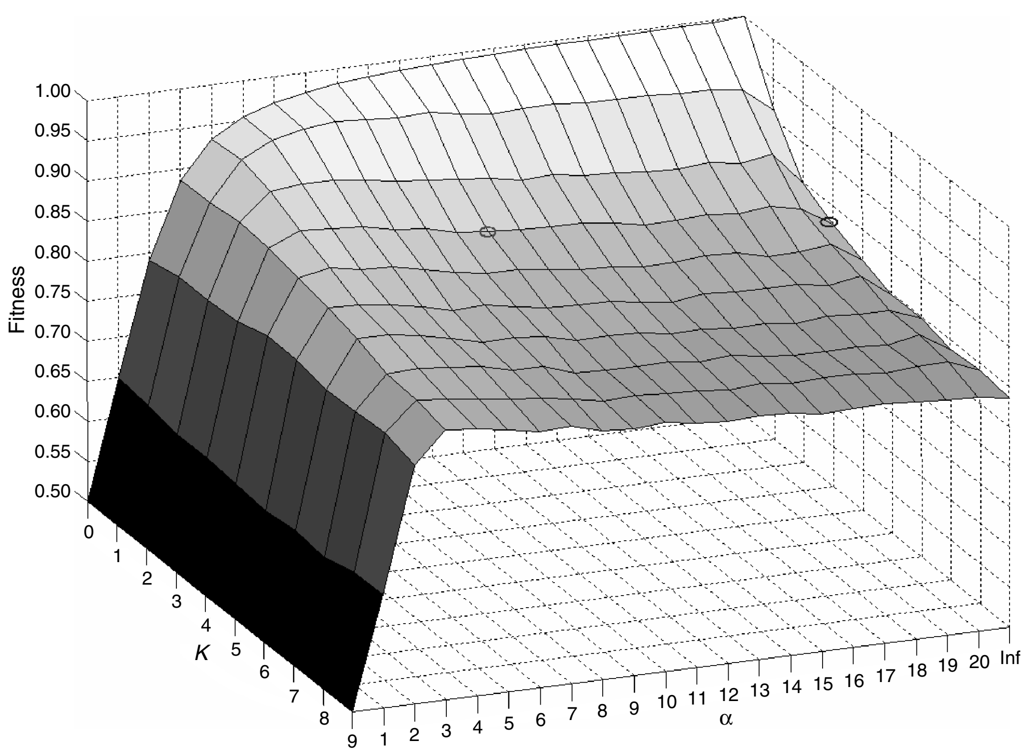
Our main result is driven by the fact that in high K worlds, firms who do local search (with perfect evaluation) will get stuck on one of the myriad local peaks. A general claim that noise (perturbations and mutations) is beneficial in complex landscapes is not novel. Scholars have long recognized that in complex environments some degree of perturbation or mutation leads to broader search and better outcomes. Our claim goes further, however, by considering what may be called intelligent noise. Mutation probabilities are (usually) identical for all of the alternatives and thus insensitive to fitness differences. In contrast, the screening function introduced in the present work is sensitive to the goodness of possible alternatives. An alternative that has much lower fitness than the current alternative will be accepted with a (very) low probability. Similarly, alternatives with much higher fitnesses will be accepted with a (very) high probability. Finally, alternatives that only differ marginally are accepted with a probability of about 1/2. Thus, imperfect evaluation introduces intelligent search in the sense that error probabilities depend on the performance of a new proposed alternative relative to the current.

Because the search effort of our imperfect evaluator (e.g., $\alpha = 10$, $N = 10$, $K = 3$) is characterized by intelligent noise, she can outperform perfect evaluators even if they benefit from (slight) random mutations.[14] The reason is that random mutations lead to broader search at the cost of occasional detours to inferior points in the fitness landscape. A screening function, by contrast, can be devised to favor broad search with a much lower probability of experiencing such detours. Even though we have used a linear screening function, our framework allows the screening function to take on any shape. This has two important implications. First, we can model a much larger family of disturbances that may influence the search process than is feasible with perturbations or mutations of bit-strings. Second, and related, we can capture any kind of deviations from perfect evaluation, including symmetric or nonsymmetric evaluation (as in prospect theory), and other kinds of misguided evaluation. This is an important property of our framework because it admits a straightforward way to model cognitive biases and evaluation errors emphasized in the behavioral literature (e.g., prospect theory).

Robustness

This result that the performance of imperfect evaluators can exceed that of a perfect evaluator is not a knife-edge property of the model. The critical factor regarding the robustness of the analysis relates to the structure of the task environment. We have examined in this baseline analysis a relatively small landscape ($N = 10$) with a moderate degree of interdependence ($K = 3$). In the limit, with no interdependencies (i.e., $K = 0$), the landscape would have just one peak corresponding to the globally maximum fitness level. In such a setting, the fact that perfect evaluators reliably identify their local peak and sustain their position on this peak over time can only serve to enhance performance. However, as long as K takes on a value of one or more, it is possible to identify an imperfect evaluation screening value that results in superior performance. Figure 7 displays the performance level reached in the final period ($t = 250$) for different screening abilities (α values) in landscapes that vary in their degree of ruggedness (i.e., their K value).[15] Not surprisingly, shifting from a zero-intelligence screener (α value of zero) to evaluators with some ability to discern superior from inferior alternatives (α values of one or more) leads to a marked improvement in performance. However, more surprisingly, as long as $K > 0$, we find that as the precision of the screeners is increased beyond some threshold level, the performance level that is achieved begins to decline. This threshold is realized at a lower value of α in more-complex environments; it is in more-complex, multipeak landscapes that the enhanced tendency for imperfect evaluators to search is most valuable. Having said that, there are sharply decreasing returns to error at high levels of error induced by high levels of imperfect evaluation (values of α of three or less in Figure 3). This observation highlights a notion of optimal imperfection,

Figure 7 Period 250 Fitness for Varying Evaluator Precision (α) and K (10,000 Evaluators: 100 Distinct Landscapes with 100 Evaluators on Each)



Note. Fitness of perfect evaluator and imperfect evaluator ($\alpha = 10$) marked at $K = 3$.

and also invites consideration of possible remedies that can help very imperfect evaluators. One obvious remedy would be to improve evaluation by training. When this is not a feasible option, it is conceivable that organization structures may have an important role in improving the overall (organization level) outcome even though the errors of individual evaluators cannot be remedied. This possibility is explored in more detail in the section below.

We have explored a relatively small landscape with $N = 10$. If the screening function is scaled appropriately, the results hold generally independent of the values of N. With higher N, the differences in fitnesses between points in the landscape shrink. This is because a larger set of points in the landscape maps onto the interval $[0, 1]$. If the slope of the screening function is scaled with a factor of sqrt$(12N)$, the results become invariant to the size of the landscape N, provided the ratio of K to N is held constant. For example, $N = 10$, $K = 1$ and slope α gives a result similar to $N = 100$, $K = 10$ and slope sqrt$(12N)\alpha$. With $N = 100$ and $K = 1$, the slope must be additionally increased for the imperfect screener to do better than the perfect screener (a scaling factor of about 5sqrt$(12N)$ has the desired property). All of these claims were supported by additional simulation results comparing $\{N = 100, K = 99, 90, 10, 1\}$, $\{N = 50, K = 49, 45, 5, 1\}$, and $\{N = 10, K = 9, 1\}$.[16]

Organizations of Evaluators

Search is not merely carried out by individuals in isolation, but such individual evaluation is typically embedded in a larger organizational context. One actor may endorse an initiative and pass it along to another actor, perhaps a hierarchical superior, for approval. Other actors may have sufficient authority to endorse or terminate an initiative on their own. We characterize an organization as consisting of a set of individuals who vary with respect to their authority to terminate proposals (terminating by their own evaluation or recommending termination to others) and endorsing proposals (authorizing the proposal on the basis of their own assessment or recommending acceptance to others).

As previously noted, we can characterize two extreme forms of organizational architectures: hierarchy and polyarchy. Hierarchy requires that for an alternative to be accepted it must pass through an approval process at each level of the organization. In this sense, hierarchy is very conservative and is unlikely to make Type II errors of accepting inferior alternatives. In contrast, polyarchy is a structure in which approval by any evaluator within the organization is sufficient for acceptance of an alternative. As a result, the polyarchy structure tends to be very proinnovation and prochange and tends not to make Type I errors of rejecting superior alternatives. Only one actor in the organization needs to see merit in the initiative for it to be adopted; however, this same property makes polyarchies prone to making Type II errors of adopting inferior alternatives.

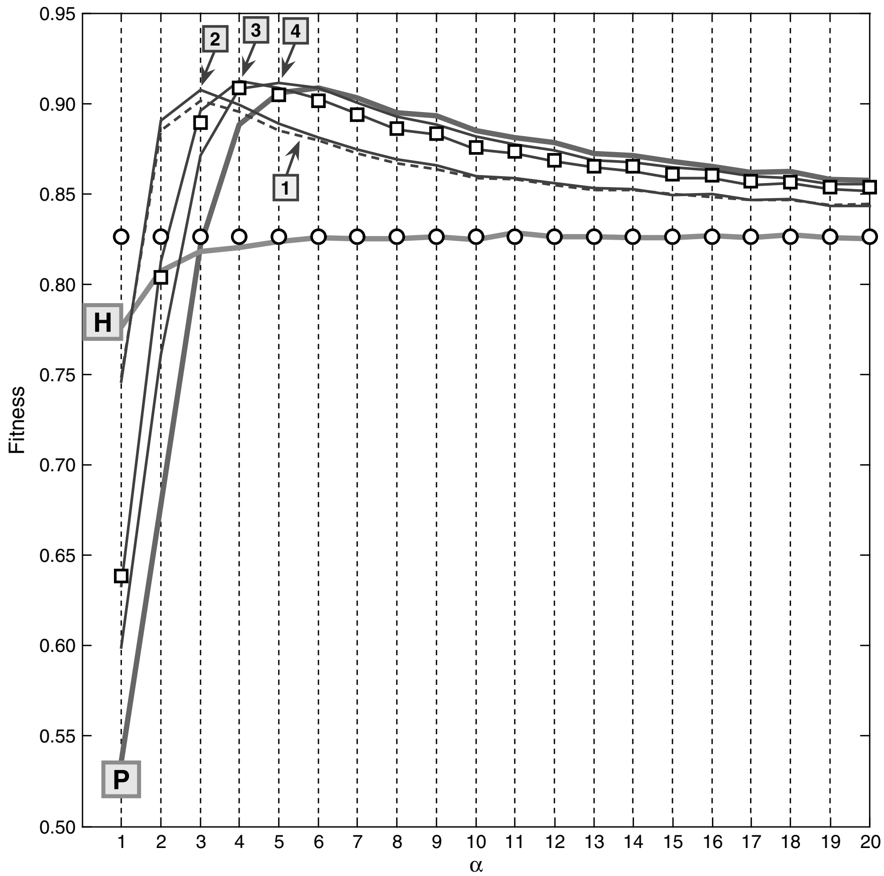
We model organizations of a fixed number of actors, six. These organizations include the two pure forms of hierarchy and polyarchy, as well as the four forms intermediate to them (see Figure 2, and Figure A1 in the appendix). Obviously, organizational structure is an interesting property for a population of imperfect evaluators; in contrast, with perfect evaluation the decision outcome would be invariant to structure. Thus, we take the same imperfect evaluation function previously examined in the individual actor analysis and examine organizations of six such actors arrayed according to the six alternative organizational forms.[17]

What is the effect of organizational architecture on search processes? We see that the hierarchical form has many of the properties of the perfect evaluator. Such organizations tend only to walk uphill, albeit slowly, because they tend only to accept new alternatives that do, in fact, lead to an increase in performance. Thus, as with perfect evaluators, they tend to be prisoners of their starting positions, identifying local peaks but not exploring more broadly in the landscape. Furthermore, we find that our intermediate forms can offer an effective mix of exploration and exploitation (Holland 1975, March 1991). Elements of polyarchy enhance the breadth of search, but some degree of hierarchy facilitates the organization's ability to reliably sustain an attractive position in the landscape, once such a position has been identified. However, given that even a population of single evaluators is able to cluster rather closely to the most attractive peaks in the landscape, only a modest degree of hierarchy is needed to reliably sustain an attractive position in the performance landscape.

Reflecting these trade-offs between the search inducing polyarchical forms and the inertia generating hierarchical forms, we find an important complementarity between organizational form and the screening ability of the evaluators who comprise the organization. Figure 8 provides a cross-sectional slice of Figure 7 at a value of $K = 3$ while extending the results to encompass the six organizational forms. Thus, Figure 8 examines the performance implications of the six organizational forms in the final period ($t = 250$), with the screening ability of evaluators within the organization spanning levels from $\alpha = 1$ (very low ability) to $\alpha = 20$ (very high ability). We find, in this setting, that for α values of six or more, the polyarchy yields the highest performance among the six organizational forms.

However, it is also important to note that some of the six organizational forms would result in a lower performance than could be generated by an individual member of the organization (i.e., the imperfect evaluator shown with square markers). Organizations have the potential to compensate for weaknesses of individual screeners (hierarchy potentially helping to reduce the extensiveness of

Figure 8 Comparison of Period 250 Fitness for Perfect Evaluator, Imperfect Evaluator, and Six Organizational Forms for Each Level of α ($N = 10$, $K = 3$)



Notes. The perfect evaluator is illustrated with round markers (no line) and the imperfect evaluator with square markers (no line). The hierarchy (H) and the polyarchy (P) are illustrated with thick lines. The four hybrids (1–4) are illustrated with thin lines. Hybrid 1, in particular, is illustrated with a thin, dashed line.

search in the case of highly inaccurate screeners and polyarchy forms usefully enhancing the degree of search for screeners who are more accurate), but as indicated in the results in Figure 8, the inappropriate organizational form may exacerbate the pathologies associated with an individual evaluator.

Figure 8 further examines this interrelationship between organizational form and screening ability across a range of screening levels. For very imperfect screening ability (α values of one and two), we see that hierarchy yields a substantially higher level of performance relative to the polyarchy form. Hierarchy is a useful complement to very imperfect screeners. Individuals who evaluate alternatives with considerable imprecision naturally induce considerable breadth of search. Breadth of search has the virtuous quality of preventing organizations from locking in prematurely to inferior peaks. Sustained breadth of search, however, has the liability of generating a more dispersed cloud, or distribution, of organizations around the superior alternatives that come to be identified in the long run. Thus, highly imperfect evaluators, even if placed in a hierarchical structure, are likely to generate wide-ranging search, but the hierarchical form will enhance the ability of such organizations to retain the attractive solutions that are identified.

Conversely, polyarchy is a desired complement to organizations composed of highly accurate screeners (α values of six or more). Accurate screeners are likely to rapidly identify a local peak in the landscape. Polyarchy, a form that permits any individual within the organization to approve an alternative, only requires one of the six actors in the organization to view an alternative as favorable in order to result in its adoption. Thus, as long as the evaluation of the individual evaluators composing the organization has some possibility of error, polyarchy compounds the likelihood of accepting an alternative that results in an immediate decline in performance; at the same time, however, polyarchy offers the possibility of broadening search to new regions of the performance landscape. If actors are highly accurate in their individual screening, such organizations do not pay a significant price for their proacceptance bias of the polyarchy form in that, in the long run, the distribution of organizations is still tightly packed around the superior peaks in the performance landscape.

For modest α values of 2–5, some hybrid form dominates both the hierarchy and the polyarchy: Hybrid 2 dominates at α values of 2 and 3, Hybrid 3 dominates at α of 4, and Hybrid 4 dominates at α of 4. The effect of Hybrid 2 is to somewhat narrow the breadth of search while producing a sharper ability to discriminate between alternatives (as shown in Figure 2). This effect is critical at values of α that are modest, but not extremely low (α value of one). As the discriminatory ability sharpens at values of α above three, less hierarchical forms that broaden search are favored, i.e., Hybrid 3 and Hybrid 4, and then, with α values of 6 or more, the polyarchy. Interestingly, the very similar results produced by Hybrid 1 and Hybrid 2 reflect the trade-off between narrowing the breadth of search (Hybrid 1) and sharpening the discriminatory ability (Hybrid 2).

Indeed, these results suggest that organizational forms must be designed to fit the contingencies of the available workforce (screening ability) as well as the task environment (level of uncertain evaluation and interdependencies among policy attributes). Thus, in the same task environment, the more able are the individual evaluators composing the organization, the more that organizational form should shift toward the permissiveness of the polyarchy form. Very able evaluators need a structure that accepts and empowers the divergent views of organizational members. Conversely, evaluators who are less able and therefore less discriminating require the repeated checks on behavior that hierarchical elements provide. Note that, as evaluators become near-perfect screeners, performance becomes insensitive to the specification of organizational form. In the limit, with perfect screeners, evaluators would simply replicate each other's evaluation decision; thus, in the limit, performance is invariant to organizational form and the number of evaluators engaged in evaluation. Thus, a perfect evaluator would not benefit from being a member of an organization.

Conclusion

Much of our analysis of search processes has been very one-sided. We recognize that choice sets are not presented to decision makers, but must be identified through search processes. However, in considering these important issues of discovery, we have tended to treat the problem of evaluation as trivial or self-evident. However, the question of valuation is far from trivial, and indeed forms the crux of resource allocation processes (Bower 1970). If, as suggested by the Carnegie School view of the firm, organizations engage in problemistic search comparing small sets of alternatives to a status quo performance, then understanding the nature of that evaluation process is critical to understanding the dynamics of search processes and, ultimately, the pattern of firm adaptation that we observe.

While clearly a stylized and admittedly incomplete treatment of this question of evaluation, the work provides some useful redirection of the field's attention, as well as some initial results of interest. An unintended by-product of a precise evaluation mechanism is the short-circuiting of the search process. Highly accurate evaluation systems will rapidly identify one of the local peaks in the neighborhood of the location at which the search process commenced. Evaluation processes under a rather wide range of imprecision will yield superior performance. It is perhaps not surprising that evaluation that is less precise would tend to result in search processes persisting for longer periods of time. However, the relatively low variability in the range of behavior and performance that such populations experience is quite surprising. Indeed, in a cross-sectional sense, highly accurate evaluators generate greater variability in performance and behavior than do populations of less-precise evaluators. Under moderate ranges of evaluation ability, we observe clouds of evaluators clustering rather tightly around the superior peaks in the performance landscape. In contrast, with highly precise evaluation we observe distinct mass points of evaluators spread out evenly on a variety of local peaks that vary considerably in their performance value.

The precision of evaluation and the relative rates of Type I and Type II errors in the evaluation process are importantly affected by the structure of organizational decision processes. Hierarchical organizational forms, for a given screening ability of the evaluators who comprise it, tend to be cautious and are unlikely to mistakenly shift to less favorable alternatives. In contrast, flat forms that have a polyarchy quality tend to induce greater search as a result of the greater probability of making such errors in evaluation. Given that highly accurate screeners are likely to stop their search process prematurely, polyarchy as an organizational form is a useful complement to highly accurate screening ability. By the same token, elements of hierarchy can facilitate an organization of rather inaccurate screeners persisting on a favorable course of action, once identified.

This basic analytical structure that we have developed can be enriched and built on in a number of ways. In our current analysis, the generation of alternatives is specified exogenously and is determined by the structure of the performance landscape. As Nelson (1982) argued, a better cognitive understanding of one's task environment may allow for more intelligent identification of the alternatives to be sampled. We have treated the sampling process as defined by local search. While this is an important line of argument in the literature (from March and Simon 1958 onward), it is important to consider the intelligent identification of nonlocal options (Gavetti and Levinthal 2000).

A different form of endogeneity that would be interesting to consider is with respect to an actor's screening ability. There is a vast literature on experiential learning (Argote 1999) that suggests that skill at tasks increases with repeated trials. Therefore, it is reasonable to expect that screening ability may change with an actor's experience with a class of problems. Thus, actors may become quite skillful and accurate in evaluating one class of alternatives, but rather inaccurate in evaluating a different and—for them—novel set of alternatives. Experiential learning of this form should tend to acerbate the problem of competency traps previously identified in the literature (Levinthal and March 1981, Levitt and March 1988). By becoming more expert evaluators in the domain of the organization's current activities, actors are less likely to engage in further search. Consistent with work on organizational demography and innovation, and similar to March's (1991) model of exploration and exploitation, turnover in personnel and the introduction of novice actors may be necessary to facilitate search processes.

While there are many such avenues of further inquiry, we wish to reiterate the basic call with which we started. Search is not merely about generating and discovering alternatives. It is equally about judging the value of those alternatives with which one is presented. We hope at a broad level, to have redirected the conversation in the organization's literature to a more balanced consideration of search processes, as well as to have provided a particular structure and set of results to facilitate such consideration.

Acknowledgments

The authors thank Felipe Csaszar, Rich Burton, the associate editor, and the anonymous reviewers for offering very useful comments and advice.

Appendix

The organizational-level screening function, F, is a polynomial in the individual-level screening function, $f(x)$, under the assumption that all members of an organization have identical screening functions. Given the flow of decisions shown in Figure A1, we derive organizational-level screening functions (shown in Figure 2) for the six organizational forms. They are as follows:

Hierarchy: $F = f(x)^6$

Hybrid 1: $F = f(x)^6 - 3f(x)^5 + 2f(x)^4 + f(x)^2$

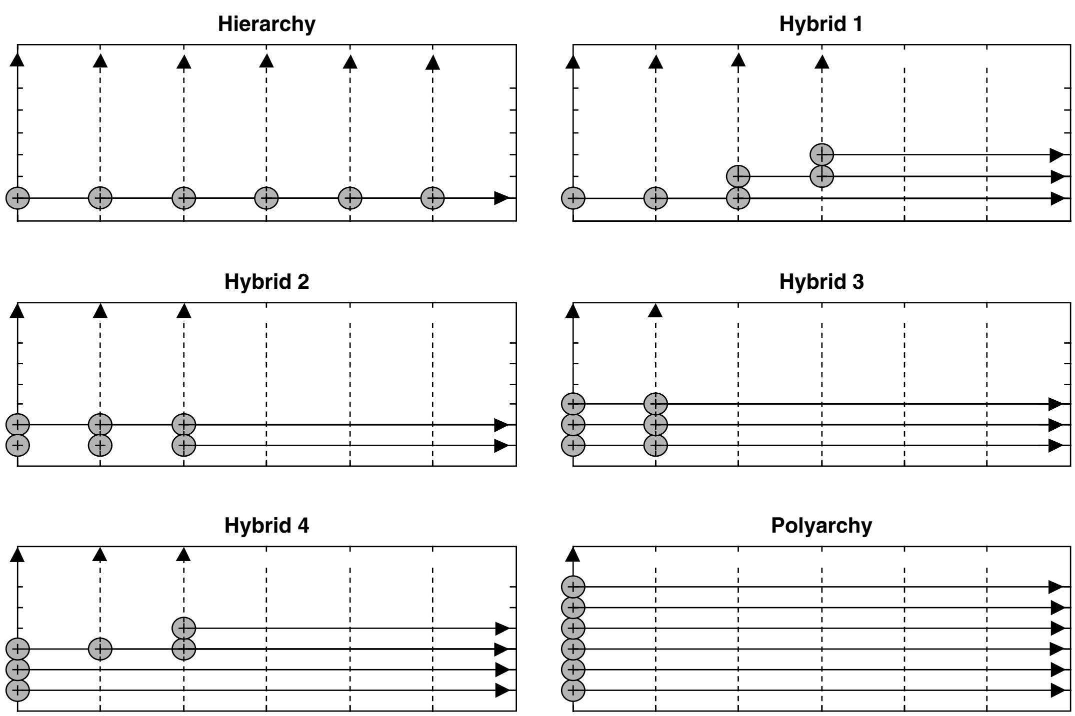
Hybrid 2: $F = -f(x)^6 + 6f(x)^5 - 12f(x)^4 + 8f(x)^3$

Hybrid 3: $F = f(x)^6 - 6f(x)^5 + 15f(x)^4 - 18f(x)^3 + 9f(x)^2$

Hybrid 4: $F = -f(x)^6 + 4f(x)^5 - 5f(x)^4 + 2f(x)^3 - f(x)^2 + 2f(x)$

Polyarchy. $F = 1 - (1 - f(x))^6 = -f(x)^6 + 6f(x)^5 - 15f(x)^4 + 20f(x)^3 - 15f(x)^2 + 6f(x)$

Figure A1 Flow of Decisions in Six Stylized Organizational Forms, Each with Six Members



Notes. In each of the six organizational forms shown in Figure A1, proposals enter with the actor in the lower left corner and flow from the left to the right. At each hierarchical level, all members must reject a proposal to avoid it being passed on to higher levels. Dashed lines show rejection of proposals and solid lines show acceptance. Proposals that exit to the right are adopted by the organization. (The number of organization members that may have a final say as to whether the organization adopts a proposal increases from one in the hierarchy to six in the polyarchy.)

Endnotes

[1]In contrast, the question of what constitutes an appropriate threshold, or aspiration level, has received considerable attention in the literature (cf., Greve 2003, Lant 1992). This boundary of what constitutes satisfactory or nonsatisfactory performance is defined by a comparison of current performance to prior performance, as well as to the performance of others who are viewed by the focal actor as belonging to his or her reference group.

[2]While it is common to refer to the value ascribed to the phenotype as fitness in this structure, the value may better be thought of as a kind of technical performance measure, as opposed to corresponding to a measure of birth and death rates as in the standard biological use of the term "fitness."

[3]It is important to note that the imperfect evaluation process that we model is not one of random error. Rather, imperfect evaluators are modeled as being intelligent and, in particular, being more likely to select superior alternatives to inferior alternatives. However, they do so with less-than-perfect reliability. Indeed, in separate analyses not reported here, we demonstrate that perfect evaluators subject to random errors perform less well than the intelligent, but imperfect, evaluators modeled here.

[4]One might imagine that actors have a status quo bias, in which case when faced with a new alternative that yields the same payoff as the current alternative, they would be inclined to reject the proposed alternative. We have analyzed screening functions with this property. Essentially, such a bias simply shifts the y intercept of these curves and generates qualitatively similar results to the results provided here. We examine the no-bias condition simply to eliminate the need to introduce another parameter in the subsequent analysis.

[5]The critical analytical challenge is to specify the implied organizational screening function that results from a set of individuals of a given ability (or individual screening function as in Figure 1) that are organized in a particular structure (as suggested by Figure A1 in the appendix). Christensen and Knudsen (2004) derived this mathematical relationship.

[6]For intermediate hybrid structures the organizational-level screening function is bounded by the number of members, n, in that the highest exponent in the polynomial F of any hybrid will be n. Intermediate structures are found by a systematic derivation of the set of feasible paths leading to acceptance and rejection (Christensen and Knudsen 2004).

[7]The issue of organizational form is not relevant in the case of perfect evaluators because each perfect evaluator in the organization would simply replicate the decision of others.

[8]We limit the study to organizations that have well-defined, noncyclic information flows as shown in the appendix. That is, there are 223 topologically distinct noncyclic, inward deterministic graphs representing hybrid organizations.

[9]See Figures 4 and 5.

[10]This claim was supported by additional simulations.

[11]This downward movement is masked by the results in Figure 2 because the figure provides the results for the average performance over a set of runs (10,000 agents: 100 distinct landscapes with 100 agents on each). Examination of individual runs does reveal instances of such downward walks.

[12]Confirmed through additional runs that are available from the authors.

[13]In Figures 4–7, we analyze the behavior over a single, randomly specified performance landscape (10,000 agents). This allows us to identify a fixed population of 2^N alternatives, or 1,024, given N has a value of 10 in our analysis. These figures were generated by assigning the average number of agents at each of the 1,024 locations for each of the 250 time steps.

[14]This conjecture was supported by an illustrative example, comparing the imperfect evaluator ($\alpha = 10$) with perfect evaluators that experience random mutations of (all) bits in a configuration with probability 0.001 ($N = 10$, $K = 3$). In this case, the imperfect evaluator ($\alpha = 10$) outperforms the perfect evaluators experiencing random mutations who, in turn, outperform the straight perfect evaluator. Perfect evaluators that experience random mutations of (all) bits in a configuration with probability 0.01 ($N = 10$, $K = 3$) are inferior to those experiencing a lower mutation probability as well as imperfect evalutors, but still superior to the perfect evaluator.

[15]The values of $K = 3$ and $\alpha = 10$ and perfect evaluation are indicated in this plot because these two points correspond to the values used in our prior baseline analysis.

[16]These results are available on request.

[17]Clearly, size is another facet of organizational form that could be varied. However, the set of possible hybrid forms grows exponentially with organizational size. Indeed, Christensen and Knudsen (2004) showed that with a sufficient number of agents, an organizational form can be specified that approximates arbitrarily closely a perfect screening function—i.e., a function that rejects inferior alternatives and accepts superior alternatives with certainty. However, note that per our analysis of perfect screening in the prior section, perfect screening need not be a desired property. To focus the attention on the role of changing organizational structure from structures that are more hierarchical to structures that are more polyarchical, we keep the number of actors within the organization fixed.

References

Argote, L. 1999. *Organizational Learning: Creating, Retaining and Transferring Knowledge*. Kluwer Academic Publishers, Boston, MA.

Bower, J. 1970. *Managing the Resource Allocation Process*. Harvard University Press, Boston, MA.

Burton, R., B. Obel. 1984. *Designing Efficient Organizations: Modeling and Experimentation*. North-Holland, Amsterdam, The Netherlands.

Christensen, M., T. Knudsen. 2004. The architecture of economic organization: Toward a general framework. LINK working paper.

Cyert, R. M., J. G. March. 1963. *A Behavioral Theory of the Firm*. Prentice Hall, Englewood Cliffs, NJ.

Dosi, G., D. Levinthal, L. Marengo. 2003. Bridging contested terrain: Linking incentive-based and learning perspectives on organizational evolution. *Indust. Corporate Change* **12** 413–436.

Ethiraj, S., D. Levinthal. 2004. Modularity and innovation in complex systems. *Management Sci.* **50** 159–174.

Galbraith, J. 1973. *Designing Complex Organizations*. Addison-Wesley, Reading, MA.

Gavetti, G., D. Levinthal. 2000. Looking forward and looking backward: Cognitive and experiential search. *Admin. Sci. Quart.* **45** 113–137.

Greve, H. 2003. *Organizational Learning from Performance Feedback*. Cambridge University Press, Cambridge, UK.

Holland, J. H. 1975. *Adaptation in Natural and Artificial Systems*. University of Michigan Press, Ann Arbor, MI.

Kauffman, S. A. 1993. *The Origins of Order: Self-Organization and Selection in Evolution.* Oxford University Press, New York.

Lant, T. 1992. Aspiration level adaptation: An empirical exploration. *Management Sci.* **38**(5) 623–644.

Lant, T. K., S. J. Mezias. 1990. Managing discontinuous change: A simulation study of organizational learning and entrepreneurship. *Strategic Management J.* **11** 147–179.

Levinthal, D. A. 1997. Adaptation on rugged landscapes. *Management Sci.* **43** 934–950.

Levinthal, D. A., J. G. March. 1981. A model of adaptive organizational search. *J. Econom. Behav. Organ.* **2** 307–333.

Levitt, Barbara, James G. March. 1988. Organizational learning. *Ann. Rev. Sociol.* **14** 319–340.

Lin, Z., K. M. Carley. 1997. Organizational response: The cost performance tradeoff. *Management Sci.* **43** 217–234.

March, J. G. 1991. Exploration and exploitation in organizational learning. *Organ. Sci.* **2** 71–87.

March, J. G., Herbert A. Simon. 1958. *Organizations*. John Wiley & Sons, New York.

Marschak, J., R. Radner. 1972. *Economic Theory of Teams*. Yale University Press, New Haven, CT, London, UK.

Nelson, P. 1970. Information and consumer behavior. *J. Political Econom.* **78** 311–329.

Nelson, R. 1982. The role of knowledge in R&D efficiency. *Quart. J. Econom.* **97** 453–470.

Nelson, R., S. Winter. 1982. *An Evolutionary Theory of Economic Change*. Harvard University Press, Cambridge, MA.

Pfeffer, J. 1981. *Power in Organizations*. Ballinger Publishing, Cambridge, MA.

Rivkin, J. W., 2000. Invitation of complex strategies. *Management Sci.* **46** 824–844.

Rivkin, J. W., N. Siggelkow. 2003. Balancing search and stability: Interdependencies among elements of organizational design. *Management Sci.* **49** 290–311.

Rosenkopf, L., A. Nerkar. 2001. Beyond local search: Boundary-spanning, exploration, and impact in the optical disk industry. *Strategic Management J.* **22** 287–306.

Sah, R., J. Stiglitz. 1986. The architecture of economic systems: Hierarchies and polyarchies. *Amer. Econom. Rev.* **76** 716–727.

Seshadri, S., Z. Shapira. 2003. The flow of ideas and timing of evaluation as determinants of knowledge creation. *Indust. Corporate Change* **12** 1099–1124.

Siggelkow, N., J. W. Rivkin. 2005. Speed and search: Designing organizations for turbulence and complexity. *Organ. Sci.* **16**(2) 101–122.

Simon, H. A. 1955. A behavioral model of rational choice. *Amer. Econom. Rev.* **69** 99–118.

Simon, H. A. 1962. The architecture of complexity. *Proc. Amer. Philos. Soc.* **106** 467–482.

Sorenson, O. 2002. Interorganizational complexity and computation. J. A. C. Baum, ed. *Companion to Organizations*. Blackwell, Oxford, UK, 664–685.

Stigler, G. 1961. The economics of information. *J. Political Econom.* **69** 213–225.

Stuart, T., J. Podolny. 1996. Local search and the evolution of technological capabilities. *Strategic Management J.* **17** 21–39.

Weber, M. 1942. A. Henderson, T. Parsons, trans. *The Theory of Social and Economic Organization.* Free Press, New York.

Wright, S. 1931. Evolution in Mendelian populations. *Genetics* **16** 97–159.