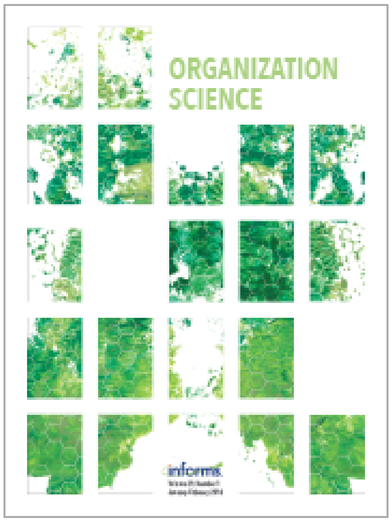


Organization Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Adaptation as Information Restriction: The Hot Stove Effect

Jerker Denrell, James G. March,

Adaptation as Information Restriction: The Hot Stove Effect

Jerker Denrell • James G. March
Institute of International Business, Stockholm School of Economics, Box 6501, 113 83 Stockholm, Sweden
71 Cubberley, Stanford University, Stanford, California 94305-3096
jerker.denrell@hhs.se • march@leland.stanford.edu

Abstract

Individuals and social systems are often portrayed as risk averse and resistant to change. Such propensities are characteristically attributed to individual, organizational, and cultural traits such as risk aversion, uncertainty-avoidance, discounting, and an unwillingness to change. This paper explores an alternative interpretation of such phenomena. We show how the reproduction of successful actions inherent in adaptive processes, such as learning and competitive selection and reproduction, results in a bias against alternatives that initially may appear to be worse than they actually are. In particular, learning and selection are biased against both risky and novel alternatives. Because the biases are products of the tendency to reproduce success that is inherent in the sequential sampling of adaptation, they are reduced whenever the reproduction of success is attenuated. In particular, when adaptation is slowed, made imprecise, or recalled less reliably, the propensity to engage in risky and new activities is increased. These protections against the error of rejecting potentially good alternatives on inadequate experiential evidence are costly, however. They increase the likelihood of persisting with alternatives that are poor in the long run as well as in the short run.

(*Organizational Learning*; *Selection*; *Risk Taking*; *Change*; *Exploration*; *Innovation*)

> We should be careful to get out of an experience only the wisdom that is in it—and stop there; lest we be like the cat that sits down on a hot stove lid. She will never sit down on a hot stove lid again—and that is well; but also she will never sit down on a cold one. (Twain 1897, p. 124).

Individuals and social systems are often portrayed as risk averse and resistant to change. A standard interpretation of such propensities attributes them to individual, organizational, and cultural traits. Within this interpretation, risk aversion and change aversion are fundamental properties of individuals and organizations. The tendencies reflected in these traits may vary among individuals and groups and may be augmented or overcome by incentives, norms, selection, or situational factors (March 1994, pp. 40–55), but the traits themselves are fixed and unexplained. This paper explores an alternative interpretation of risk taking and change in social systems, one that pictures these predispositions as evolving from experience at the individual or population level. In particular, we show how the reproduction of success, inherent in the sequential sampling of adaptive processes, results in a bias against both risky and novel alternatives.

Adaptation as Sequential Sampling

Modern treatments of organizational development over time are primarily variations on two themes of adaptation. The first theme is experiential learning, the idea that organizations and the people in them modify their actions on the basis of an evaluation of their experiences (Cyert and March 1963, Huber 1991, Haleblian and Finkelstein 1999). The second theme is competitive selection and reproduction, the idea that organizations and the people in them are essentially unchanging, but survive and reproduce at different rates depending on their performance (Hannan and Freeman 1977, Nelson and Winter 1982, Aldrich 1999).

Traditionally, both forms of adaptation have been presented as instruments for improving the fit between organizations (or populations of organizations) and their environments. Indeed, presumptions of the efficiency of learning and competitive selection in reaching optimal solutions have been portrayed as justifications for theories of rational choice (Friedman 1953). Such a portrayal is misleading. Although there is no question that both forms of adaptation can lead to major transformations of organizations (Yelle 1979, Haveman 1992, Greve 1996, Usher and Evans 1996, Sutton and Barto 1998), neither learning nor competitive selection and reproduction can guarantee the discovery and adoption of optimum practices. Explicit models of adaptation have demonstrated that adaptive processes are prone to settling into stable suboptima (Levinthal and March 1981, Herriott et al. 1985, Kauffman 1993, Carroll and Harrison 1994, Levinthal 1997). Except in a tautological sense, the fittest do not necessarily survive (Nelson and Winter 1982, Carroll and Harrison 1994, Barnett 1997, Gimeno et. al. 1997), and learning from experience does not necessarily locate global maxima (Busemeyer et al. 1986, Levinthal and March 1993, Miner and Mezias 1996, Levinthal 1997). In this paper we demonstrate that the sequential character of adaptation implies two specific kinds of suboptimality that are important for organizational development.

An adaptive process can be seen as a process of sequential sampling. At each point in a sequential sampling process, one alternative is sampled from a set of alternatives; the probability of sampling any particular alternative depends on the past history of observations (Wald 1947, Feller 1968). An adaptive process is a sequential sampling process in the sense that alternatives that have experienced relatively good outcomes in the past are more likely to be sampled than are alternatives that have experienced relatively poor outcomes (Holland 1975). Several models have been developed to explore optimal stopping rules in this kind of situation (Rothschild 1974, Easley and Kiefer 1988, Goldberg 1989, Kaelbling 1993, Blume and Easley 1995). The focus here is different. We examine the behavioral consequences of the sequential sampling character of adaptation.

In general, sequential sampling processes lead to improvement in performance as alternatives with good outcomes in previous samples come to dominate future samples. However, the differential reproduction of successful alternatives not only affects the returns to action but also the accumulation of information. Because alternatives that do well are reproduced, adaptation will generate further information about their potential. The additional information tends to correct any errors involving alternatives that initially appear better than they actually are. On the other hand, alternatives that do poorly are likely to be avoided. As a result, it is less likely that additional information about their potential will be obtained. Consequently, adaptation will often fail to correct errors involving alternatives that initially appear worse than they actually are. In deference to Mark Twain and his homily about cats with which we began this paper, we call this phenomenon the hot stove effect. The hot stove effect refers to the asymmetry in the capability of adaptive processes to correct early sampling errors.

The hot stove effect leads to a bias against new alternatives that require practice and against alternatives involving risk. Consider new alternatives. There are numerous reasons why new alternatives fare poorly in comparison with the status quo. We address here a particular feature of adaptive processes that operates to sustain the status quo—the so-called "competency trap." Most new alternatives require practice to realize their full potential reliably. As a result, the outcomes of new alternatives tend initially to be lower than their potential and unreliable (high variability). With repeated experience, average performance improves and variability is reduced. However, even a new alternative that has the potential to improve with practice and eventually to surpass existing alternatives is initially likely to perform poorly in comparison with old and established alternatives. The short run disadvantage leads to avoidance, which restricts the practice that would develop the performance. As a result, the error is unlikely to be corrected. On average, and without any particular individual or organizational resistance to change, this leads to a bias against new alternatives in competition with established ones (David 1985, Levitt and March 1988, Arthur 1989).

The same mechanism produces a bias against risky alternatives. When early experiences are more favorable to a risky alternative than its true average value warrants, the probability of sampling the risky alternative will increase and subsequent sampling will correct the error. However, when early experiences are less favorable to a risky alternative than its true average value warrants, the probability of choosing the risky alternative will decrease. As a result, the error is less likely to be corrected. On average, this asymmetry of sequential sampling leads to a bias against risky alternatives. The bias is sometimes labeled "risk aversion" when it is observed, but it is not attributable to any traitlike risk preferences on the part of individuals or organizations. Adaptation itself produces behavior that can be described as risk averse (March 1996).

To demonstrate the effects more precisely, we consider two classic modes of adaptation—experiential learning and competitive selection—each of which exhibits reduced sampling of failures. Each of these domains is rich with models. The intricacies of their fine detail are essential, but we focus on their basic features to illustrate how those features are implicated in information restriction.

Experiential Learning

In an experiential learning process, the propensity to select a particular alternative in period $t + 1$ is based on the outcome of choices in previous periods (Cyert and March 1963, Nelson and Winter 1982, Levitt and March 1988, Huber 1991, Miner and Mezias 1996). Although it is clear that such updating can lead to improvements (Yelle 1979, Argote and Epple 1990), analytical models of experiential learning have demonstrated that the process does not necessarily converge to a global optimum.

As a result of ambiguity, environmental turbulence, interdependency, and path dependency introduced by competence multipliers, experiential learning can produce superstitious learning (Lave and March 1975, Lounamaa and March 1987, Levitt and March 1988) and converge to inferior alternatives (Levinthal and March 1981, Herriott et al. 1985, Levitt and March 1988, Levinthal 1997). Although the idea that experiential learning is prone to mistakes is well known (Huber 1991, Levinthal and March 1993, Lant 1994, Miner and Mezias 1996), it is less commonly observed that the failures tend to be systematic. In this section we show more precisely how experiential learning leads to a bias against risky and new alternatives.

There is a long tradition of computer simulation models of experiential learning within organizational research (Levinthal and March 1981; Herriott et al. 1985; March 1991; Lant and Mezias 1990, 1992; Lant 1994). Although the models differ in their details, most of them assume that organizations repeat actions that appear successful and change actions that appear unsuccessful, and that performance is evaluated relative to an aspiration level that depends on the history of past performances (Cyert and March 1963, Lant 1992, Greve 1998).

A basic model consistent with this tradition is an experiential learning model of an individual or organization choosing between two alternatives. The probability that the first alternative is chosen in period t is denoted P_t.[1] Learning is assumed to change this probability on the basis of a comparison between the realized outcome from the choice of an alternative at a particular time and an aspiration level. The aspiration level summarizes past experience as a mix between the previous aspiration and the previously realized outcome, thus as an exponentially weighted average of past outcomes (Levinthal and March 1981). Formally, the aspiration level at $t + 1$ is given by:

$$L_{t+1} = L_t (1 - b) + O_t b, \qquad (1)$$

where L_t is the aspiration level and O_t is the outcome at t. The mix between the most recent outcome and the most recent aspiration is controlled by b, a non-negative fraction that reflects the rate at which aspirations adjust to experience.

Consistent with stochastic learning models (Bush and Mosteller 1955, Coombs et al. 1970), we assume that positive experience with an alternative increases, and negative experience decreases, the probability that this alternative will be chosen in the future. The change in the probability is assumed to be proportional to the difference between P_t and the learning limit (that is, one in the case of an increase and zero in the case of a decrease). Specifically, if the first alternative is tried and yields an outcome better than the aspiration at time t, or if the second alternative is tried and yields an outcome worse than the aspiration, the probability of choosing the first alternative in the next time period, P_{t+1}, increases in the following way:

$$P_{t+1} = P_t + a(1 - P_t) \qquad (2)$$

Here, a is a positive fraction parameter that defines the speed of learning. The higher the value of a, the more a single experience will affect the subsequent probability of choosing the risky alternative.

On the other hand, if the first alternative is tried and yields an outcome that is worse than the aspiration, or if the second alternative is tried and yields an outcome better than the aspiration, the probability of choosing the first alternative decreases in the following way:

$$P_{t+1} = (1 - a)\, P_t \qquad (3)$$

Regardless of which alternative is chosen, if the realized outcome is identical to the aspiration, the probability of choosing the risky alternative in the next time period remains unchanged.

The Reproduction of Reliability Through Experiential Learning

To demonstrate how the model of experiential learning specified above generates a bias against risky alternatives, we examine a two-alternative setting in which one alternative has a certain outcome, Y, and the other alternative involves risk. The risk of the latter alternative is measured by the variability in the probability distribution over possible outcomes conditional on its choice. In the remainder of the present treatment, we consider a risky alternative characterized by a normally distributed outcome distribution with mean X and standard deviation S, and we consider the effect of varying the value of S.

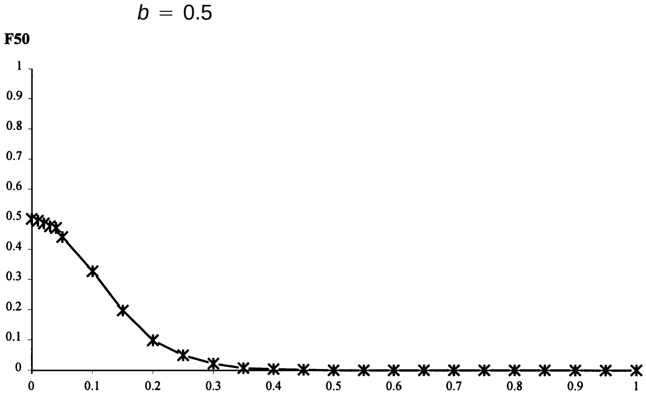
We define the probability that the risky alternative will be chosen at time t as P_t. We assume that initially a learner is equally likely to choose the risky alternative or the certain one ($P_0 = 0.5$) and that the initial aspiration is identical to the expectation at $t = 0$ ($L_0 = [P_0 X + \{1 - P_0\} Y]$), where X is the expectation of the risky alternative, and Y is the expectation of the certain alternative. Thus, the learner has no initial bias with respect to risk and has an initial aspiration that is equal to the expectation. Alternative assumptions about the initial aspiration level were explored. As long as they are not extreme, they do not affect the general results below.

We ask what happens to the likelihood of choosing the risky alternative over time in this learning situation. It seems reasonable to expect that the results might depend on the relation between the expected returns from the two alternatives, with a risky alternative becoming more likely to be chosen over time the greater its expected return. That expectation is correct. However, there is also a systematic bias in favor of reliability. Suppose that the risky alternative and the certain alternative have the same long-run expectation, $X = Y$. To examine the aggregate statistical properties of learning in such a situation, we simulate the learning of 5,000 learners over 50 time periods and record the fraction of the 5,000 who choose the risky alternative on the fiftieth choice, F_{50}. As Figure 1 shows, the results indicate a clear learning bias against risky alternatives. The fraction of learners choosing the risky alternative at time 50 depends on the learning rate, a, but for all levels of a, the fraction choosing the risky alternative is less than one-half.

The effect is due to the natural learning tendency to avoid alternatives that produce poor outcomes and the high likelihood that a risky alternative will do so. The normal variation of a risky alternative produces periods of good returns and periods of poor returns. Whenever the risky alternative is chosen and generates a poor outcome, the probability of choosing it again decreases. This optional stopping feature of learning leads the learner to abandon a risky alternative that, at some point, has a run of bad luck and thus fails to experience any subsequent run of good luck. After some time, most individuals will have a low probability of choosing the risky alternative. In that sense, at least, learning produces behavior that might be classified as risk aversion (March 1996).

The bias against risky alternatives can be overcome if the risky alternative has a sufficient advantage in expected value. Suppose the expected value of the certain alternative (Y) is equal to 10. What expected value for the risky alternative (X^*) is required to make that alternative equally likely to be chosen after learning from the outcome of 50 choices? As might be anticipated, the expected value advantage required for the risky alternative depends on the standard deviation (S) and the learning rate (a). On the basis of simulations similar to those above, we can estimate the expected value of the risky alternative that is required for the risky alternative to be chosen half of the time at trial 50. This estimate is given in Tables 1a and 1b for selected values of S and a. The expected value advantage required to compensate for the bias against risky alternatives can be quite large.

Figure 1 **The Fraction of 5,000 Individuals Who Choose the Risky Alternatives at the End of Period 50 as a Function of *a*. Based on Averages from 25 Sets of 5,000 Simulations Where** $X = 10$, $Y = 10$, $S = 10$, $b = 0.5$



It should be noted that the bias against risky alternatives is sensitive to the aspiration level updating process assumed. If the updating function includes a positive constant (Lant 1992), the aspiration level will be above an exponentially weighted moving average of realized past outcomes; thus it is likely to be above the expected value of the alternatives. As a result, the bias will be attenuated. If the aspiration level is above the expected value of the alternatives, there will be a bias against the certain alternative unless the constant is so large as to make all outcomes failures and learning superstitious (Lave and March 1975). On the other hand, if the updating function includes a negative constant, the bias will be strengthened. The certain alternative is much more likely to be defined as a success than is the risky alternative, unless the negative constant is so large (in absolute terms) as to essentially make all outcomes successes and learning superstitious.

The Reproduction of the Status Quo Through Experiential Learning

New alternatives frequently require competence to realize their full potential. If competence depends on experience, experiential learning may converge to stable suboptima when the outcome of alternatives depends on their frequency of use (Levinthal and March 1981, Herriott et al. 1985). In this section we show that the same learning model that produces a bias against risky alternatives also produces a bias against new alternatives that require practice.

Suppose there is an existing alternative with a certain return of Y and a new alternative with a payoff subject to risk. We assume that the risky alternative generates an outcome with an average potential of X ($X > Y$), subject not only to the random fluctuations of risk but also the systematic fluctuations due to variations in competence. The realized outcome from the new alternative at t is a draw from a normal distribution with a mean equal to $c_t X$, where c_t is the competence at t ($0 < c_t \leqq 1$). The

Table 1a **The Approximate Value of X Required for 50% of the Choices at Time 50 to Be Choices of the Risky Alternative, for Various Values of S. Each Entry is Based on 5000 Simulations Where $Y = 10$, $a = 0.2$, and $b = 0.5$**

S	5	10	15	20	25
X^*	13.7	17.2	21.7	25.1	28.8

Table 1b **The Approximate Value of X Required for 50% of the Choices at Time 50 to Be Choices of the Risky Alternative, for Various Values of a. Each Entry is Based on 5000 Simulations Where $Y = 10$, $S = 10$, and $b = 0.5$**

a	0.1	0.2	0.3	0.4	0.4
X^*	13.4	17.2	20.2	21.7	22.9

standard deviation of the outcome distribution at t is $(S/c_t)^k$, where $0 < k \leqq 1$. Competence at the new alternative increases each time that alternative is chosen, thus increasing the expected performance and reducing the variation (although the standard deviation always remains positive). Over repeated choices of the new alternative, the outcome from using that alternative approaches an expected return of $X > Y$ and a standard deviation of S^k, but in the short run, when c_t is relatively small, outcomes from the new alternative are, on average, inferior to outcomes from the existing (certain) alternative.

Competence in using the new alternative increases with each utilization in a way analogous to standard learning curves (Yelle 1979, Argote and Epple 1990). Thus,

$$c_{t+1} = c_t \text{ if the existing alternative is chosen.}$$
$$c_{c_{t+1}} = c_t + d(1 - c_t) \text{ if the new alternative is chosen. (4)}$$

The learning parameter, d, is a non-negative fraction and reflects the competence learning rate.

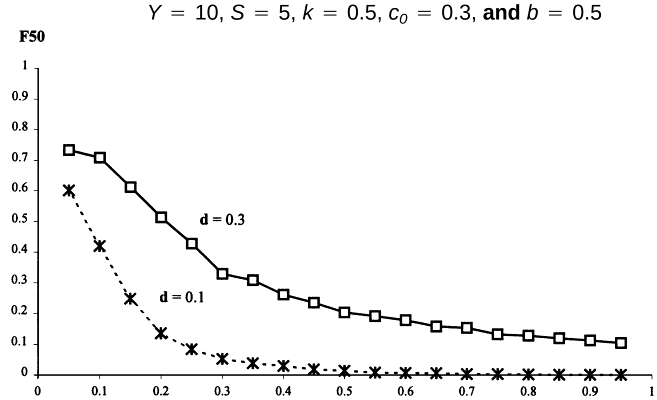
In this modification of the original experiential learning model, the adapter simultaneously learns which alternative to choose and gains competence on the chosen alternative. As intuition might suggest, the probability of choosing the new alternative frequently falls to a very low level. As the probability of choosing the new alternative becomes very low, it becomes unlikely that a learner will experiment any further with it. As a result, competence remains low and the probability of choosing the new alternative remains low. Only those individuals who persist with the new alternative gain enough competence and eventually increase the probability of choosing it.

The proportion of individuals who choose the new alternative depends on the competence learning parameter, d, by which competence at the new alternative increases with experience. The proportion of individuals who choose the new alternative also depends on the choice learning parameter. The smaller the value of a, the larger the proportion of individuals who will choose the new alternative at the end of 50 periods. Figure 2 shows that the fraction of all individuals choosing the new alternative as a varies from 0.05 to 0.95 for several values of d, when $X = 15$ and $Y = 10$. As Figure 2 suggests, one way of avoiding a bias against alternatives that require practice is rapid learning of competence and slow learning of choices—longer experimentation with alternatives that seem poor. The latter would clearly be a good idea save for the obvious complication that in a world in which seemingly poor alternatives most often are poor, extended periods of experimentation with apparently poor alternatives is usually a very costly strategy.

Robustness

The above simulations show that a behavioral model of organizational learning produces a bias against risky and new alternatives. The results are not limited to a narrow range of parameter values or to the specific models outlined above. For example, a bias against risky alternatives is also a property of learning models that select alternatives on the basis of the average outcome they have generated in the past (Sutton and Barto 1998), and also of models in which the magnitudes of learning effects depend on the magnitudes of success and failure. Among the models that we have considered, only those that insert an explicit preference for alternatives with a high variance (i.e., Kaelbling's Interval Estimation Method, Kaelbling 1993) avoid the bias against risky alternatives.

Figure 2 **The Fraction of 5,000 Individuals Who Choose the New Alternative at the End of Period 50 as a Function of a and d. Each Line Is Based on Averages from 19 Sets of 5,000 Simulations Where $X = 15$, $Y = 10$, $S = 5$, $k = 0.5$, $c_0 = 0.3$, and $b = 0.5$**



All these models are strictly models of autonomous experiential learning. Individuals and organizations also routinely learn from each other (Bandura 1977, Levitt and March 1988, Miner and Haunschild 1995, Miner and Raghavan 1999). Whether such imitation will attenuate or aggravate the above biases depends on the mode and sequence of imitation (Haunschild and Miner 1997, Miner and Raghavan 1999). If imitation involves observing both the actions of others and their associated outcomes, it is possible that the bias against risky alternatives will be weakened. If knowledge about a risky alternative is shared among several organizations, the false impressions gained from the limited sample of one organization can be corrected, provided the shared knowledge itself does not affect the individual actions. If poor outcomes in any organization lead to avoidance by all organizations, however, the bias against risky alternatives is likely to be strengthened, rather than weakened, by sharing. Finally, if imitation is based on the frequency of particular actions, but not on any observation of their consequences (DiMaggio and Powell 1983, Tolbert and Zucker 1983), a bias against risky or new alternatives may be aggravated through imitation.

Competitive Selection and Reproduction

The actions in a population of organizations can be seen as the results of adaptive processes in which the characteristics of a population of actions in period $t + 1$ are a function of the relative performance of organizations exhibiting those actions in period t (Hannan and Freeman 1977, Nelson and Winter 1982, Barnett 1997, Aldrich 1999). The population changes by virtue of selection—the differential survival of higher performers—and by reproduction (the differential adoption of the attributes of higher performers by new members of the population).

Although selection and reproduction can lead to major transformations of populations (Usher and Evans 1996, Baldwin 1998), they, like learning processes, do not necessarily lead to global optima. Competitive selection processes are history and frequency dependent and are prone to settling into stable suboptima (Kauffman 1993, Carroll and Harrison 1994, Levinthal 1997). However, deviations from optima are not random. Processes of competitive selection and reproduction favor some alternatives over others. Here, we focus on the tendency for such processes to be biased against risky and new alternatives.

Most empirical studies of competitive selection examine selection and reproduction as a function of characteristics of the population, such as density (Hannan and Carroll 1992) or the level of concentration (Baldwin 1998) and do not include explicit measures of individual or organizational performance. There are several simulation models, however, in which selection and reproduction are treated as a function of individual performance. These models range from stylized reduced form models, in which the probability of failure is assumed to be a decreasing function of relative or absolute performance (Lant and Mezias 1990, 1992; Levinthal 1991, 1997; Mezias and Eisner 1997), to detailed models of the competitive process and of capital markets (Nelson and Winter 1982).

To illustrate the bias against risky alternatives, we consider a reduced form model in which the probability of failure is assumed to be a decreasing function of relative performance. Specifically, we consider a population of unchanging organizations drawn from two distinct types. The types differ in their propensities to choose one or the other of two alternatives. As a result of their differing propensities to choose the two alternatives, the two types realize different outcomes. We assume that organizations with low performance relative to others fail to survive and are replaced. Because the two types survive and reproduce at different rates, the process generates a changing population of invariant types. There is, however, no mutation of types. The model is strictly a model of competitive selection and reproduction among invariant types.

We assume a constant population size of $N_{1,t} + N_{2,t}$ organizations, where $N_{j,t}$ is the number of organizations of type j. In the beginning each of the two types is equally represented in the population. Their representation in the population changes over time in response to differential survival and reproduction. Survival can be modeled in many different ways (Lant and Mezias 1990, 1992; Levinthal 1991; Mezias and Lant 1994; Mezias and Eisner 1997). To capture the importance of relative performance without relying upon an explicit model of the competitive process, we assume the probability of failure depends on a ranking of organizations according to performance. Specifically, we assume that lower ranked organizations fail and that some fraction, w, of the population is eliminated in each period.

Each eliminated organization is replaced by a new organization. Denote the probability that the replacement will be of type j by $r_{j,t}$. Several different assumptions about $r_{j,t}$ are possible and have been used in the literature. First, reproduction may be random among the types. That is, the type of a new organization may be randomly selected from the set of surviving types without regard to the number of survivors of the several types.

Second, reproduction may be sensitive to the number of survivors (Lant and Mezias 1990, 1992; Hannan and

Carroll 1992; Carroll and Harrison 1994). To reflect this we define:

$$r_{j,t} = N_{j,t}^{\,h}/(N_{j,t}^{\,h} + N_{i,t}^{\,h}). \tag{5}$$

The hypothetical process is one in which relative density confers legitimacy or in which new replacements are produced by survivors. Lower values of h reduce the reproductive advantage of numbers. Higher values of h increase the reproductive advantage of numbers.

A third feature of the surviving population that might affect $r_{j,t}$ is the aggregate performance of the survivors of each type (March and Shapira 1992). To reflect this, we define:

$$r_{j,t} = T_{j,t}^{\,h}/(T_{j,t}^{\,h} + T_{i,t}^{\,h}), \tag{6}$$

where $T_{j,t}$ is the aggregate performances of all survivors of the type j. The hypothetical process is one in which reproduction is driven by the total prior performance of a particular type. Prior performance of survivors provides the resources for reproducing their types. The exponent h affects the sensitivity of reproduction to the aggregate performance of a type.

Finally, a fourth feature of the surviving population that might affect $r_{j,t}$ is the average performance of each type. To reflect this, we define:

$$r_{j,t} = A_{j,t}^{\,h}/(A_{j,t}^{\,h} + A_{i,t}^{\,h}), \tag{7}$$

where $A_{j,t}$ is the average performance of survivors of organizations of type j. The hypothetical process is one of imitation and diffusion. New organizations mimic existing types that have high average performance (Argote et al. 1990). The exponent h affects the sensitivity of reproduction to the average performance of a type.

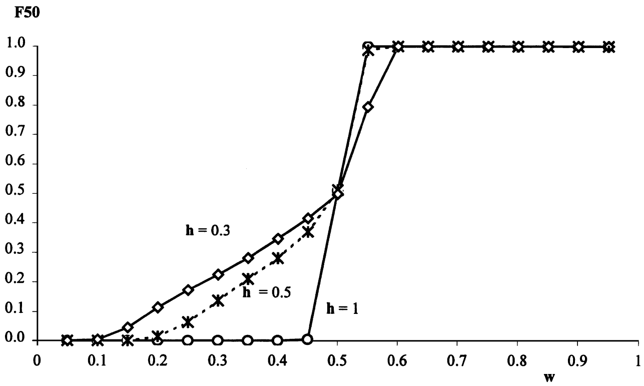
The Reproduction of Reliability Through Competitive Selection and Reproduction

To examine the reproduction of reliability we consider a situation in which the first type of organization consistently chooses a certain alternative, while the second type consistently chooses a risky alternative. The certain alternative is characterized by a reliable return, Y. The risky alternative is characterized by a normally distributed outcome distribution with mean X and standard deviation S. Each organization realizes an outcome for each time period, resulting in a ranking of organizations according to their most recent performance. Higher ranked performers survive; poorer performers are eliminated. Each period, a fraction, w, of the population is eliminated.

As long as w is less than one-half, the variability in the performances of the risky alternative assures that eliminated organizations will be drawn disproportionately from the type that chooses the risky alternative. If reproduction is random, selection thus leads to a steady decrease in the proportion of organizations of the risky type when w is less than one-half and a steady increase in the proportion of organizations of the risky type when w is greater than one-half. If reproduction is responsive to the number of survivors, or is randomly drawn from surviving organizations with positive performance (Lant and Mezias 1990, 1992), the bias is aggravated. The adaptive process (in both its "death" and "birth" processes) steadily increases the fraction of the population that avoids the risky alternative if w is less than one-half, and steadily decreases that fraction if w is more than one-half. The results after 50 trials, when reproduction is responsive to the number of survivors, are illustrated in Figure 3. The case shown is for $X = Y = 10$ and $S = 10$, and for various values of w and h. As is clear from Figure 3, unless selection pressure eliminates a majority of the population each time period (i.e., $w > 0.5$), the selection process reproduces reliability. This result holds even if the alternatives vary less profoundly in their risk and (under some circumstances) even if the expected value for the risky alternative is greater than that of the certain alternative.

Suppose, next, that reproduction depends on the total aggregate performance of survivors of a particular type, i.e., follows Equation (6). In general, the results are similar to the case involving numbers alone. The process usually reproduces reliability as long as selection eliminates less than one-half of the population each time period (i.e., $w < 0.5$). However, dependence on total aggregate performance of survivors makes that result weaker. If reproduction depends on the aggregate performance of all actors who have existed in a particular time period (including those who are eliminated in that period), reliability is reproduced as long as $w < 0.5$. However, when reproduction depends only on survivors, the fact that the average performance of survivors of the risk-taking types will, in general, be greater than that of the risk-averse types assures that the reproduction of reliability will be moderated. This also holds if reproduction depends on average performance of the two types, i.e., follows Equation (7). When reproduction responds to the average performance of the types, the risk-taking type continues to be eliminated in greater numbers than the risk-averse type; but it is favored in reproduction. As a result, it is possible (though not guaranteed) that the risk-taking type will maintain a substantial presence in the population even if $w < 0.5$. However, this can only occur if reproduction is based on the average performance of survivors, rather than the average performance of all organizations of each type.

Figure 3 **The Fraction of the Population Who Choose the Risky Alternative at the End of Period 50 as a Function of w and h. Each Line Is Based on Averages from 19 Sets of 1,000 Simulations Where $X = 10$, $Y = 10$, and $S = 10$. The Constant Population Size is 100.**



The Reproduction of the Status Quo Through Competitive Selection and Reproduction

Competitive selection and reproduction processes are biased not only against unreliable alternatives but also against alternatives that require practice. The fundamental intuition is that since alternatives that require practice will initially be inferior to other alternatives, it is possible that they will be eliminated before their potential can be revealed (Nelson and Winter 1982, Elster 1984, Levitt and March 1988, Carroll and Harrison 1994). Whether this will happen depends on the rate of competence gain relative to the speed of selection and on the possibilities for reproductive transfer of competence (Argote et al. 1990).

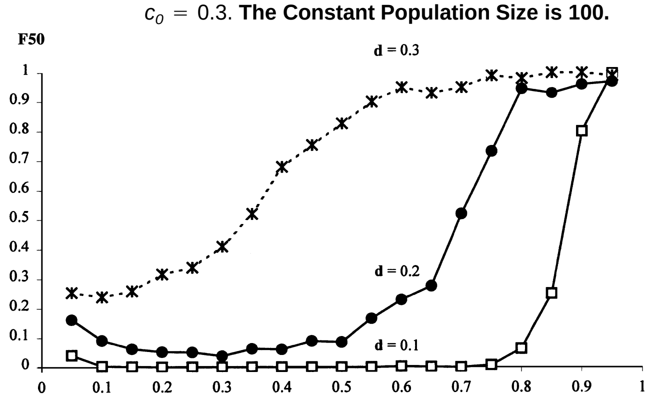
To examine this, we assume that replacements are drawn from two types according to the number of survivors, as in Equation (5). As in the parallel case within the learning model, we assume that the established alternative generates a certain outcome of Y. The new alternative has a potential expected value of $X > Y$ but generates a realized outcome at t that is a draw from a normal distribution with a mean equal to c_tX, where c_t is the competence at t ($0 < c_t \leqq 1$). The standard deviation of the outcome distribution at t is $(S/c_t)^k$, where $0 < k \leqq 1$. Competence at the new alternative increases each time that the alternative is chosen according to (4). Although the new alternative may initially only be represented by a few organizations, we examine the case where both types are equally represented in the initial population.

Consider, first, the case in which there is no within-type intergenerational reproductive transfer of competence. Learning affects reproduction rates within a generation by affecting competence, and thereby relative performance, but competence is not reproduced. In terms of the above model, this implies that whenever an organization is eliminated and replaced with another organization, the initial competence of the new organization is set at c_0.

Given these assumptions, it is clear that any organization choosing an alternative that requires practice will initially exhibit inferior performance. The longer it survives, the better its performance. If its potential is greater than that of the established alternative and it survives long enough, its performance will become reliably better than the established alternative. However, even though the potential of the new alternative is greater than the performance of the certain alternative, i.e., $X > Y$, and thus could ultimately come to dominate, the new alternative can fail to survive in the population long enough to achieve the level of competence that leads to survival and reproduction. The likelihood of this outcome depends on w and d, as illustrated in Figure 4. Only if competence is gained relatively quickly (i.e., d is relatively large), or if strong selection pressure favors risky alternatives (i.e. w is relatively large), will a new alternative be able to survive and dominate. The effect of w is sensitive to the standard deviation of the alternative that requires practice. If the standard deviation (i.e., S) is low (for example $S = 1$), organizations choosing the alternative that requires practice are reliably worse than the established alternative and are thus unlikely to survive for long, unless the selection pressure (i.e. w) is very low.

If S is somewhat greater, however, there is always some chance that an organization choosing the alternative that

Figure 4 **The Fraction of the Population Who Choose the New Alternative at the End of Period 50 as a Function of w and d when $h = 0.5$. Each Line Is Based on Averages from 19 Sets of 100 Simulations Where $X = 15$, $Y = 10$, $S = 5$, $k = 0.5$, and $c_0 = 0.3$. The Constant Population Size is 100.**



requires practice will survive and be reproduced. If a sufficient number of organizations of that type survive, reproduction will increase the supply of organizations choosing the alternative that requires practice, which, if they survive, will also become reliably better, and so on until organizations choosing the alternative that requires practice will come to dominate (pending the arrival of a new competitor). This implies that the alternative that requires practice is quite likely to dominate, ultimately, if the total population is large enough. "Ultimately" is usually, however, an extremely long time in the future.

Suppose, on the other hand, that competence is transferred in reproduction. Specifically, suppose we assume that the initial competence of a new organization arriving in period $t + 1$ is equal to the average competence of all surviving organizations of the same type at the end of period *t*. As in the case without transfer of competence, the average performance of surviving organizations choosing the new alternative will surpass the performance of organizations choosing the certain alternative at some point, but they are at risk of becoming extinct before that occurs. Intergenerational transfer of competence reduces that risk, but the risk remains. The fraction of the population choosing the new alternative can go to zero even in the presence of intergenerational transfer of competence.

In general, although there are conditions under which new alternatives can overcome old ones through competitive selection and reproduction, the status quo has a fairly clear advantage, even in competition with alternatives that are, on average, considerably better. It may be worth noting that the effects are sensitive to the size of the population. If the population is large, it is less likely that all individuals choosing an initially inferior, but potentially better, alternative will fail to be reproduced. As a result, it is more likely that some individuals will survive and eventually attain a level of reliable competence sufficient to avoid being eliminated.

Robustness

Somewhat different assumptions about selection and reproduction could be made. Alternative assumptions about reproduction have been discussed above. It was demonstrated that the present results hold for a wide variety of assumptions regarding reproduction. With respect to selection, it should be clear that the results hold for a wide variety of models in which poor performers are unlikely to survive, while average performers and good performers are likely to survive. For example, suppose that organizations fail whenever their cumulative resources fall below some absolute level of performance (Nelson and Winter 1982; Lant and Mezias 1990, 1992; Levinthal 1991; Mezias and Eisner 1997), and this absolute level of performance is below the expected performance of the two types. In this case, because organizations of the risk-taking type will be more likely than organizations of the risk-averse type to generate a performance below the level required for failure, they will be less likely to survive. Similarly, organizations making use of a new alternative will be less likely to survive. These results would be weakened, however, if it were assumed that the expected performance of organizations of a given type was a decreasing function of the number of organizations of this type (Hannan and Carroll 1992). In this case, a high density of a particular type would decrease its expected performance, which eventually would reduce its density.

As noted above, the results are also sensitive to alternative assumptions about imitation and the diffusion of knowledge. If knowledge about the results achieved by others is diffused, it may eliminate the biases, depending on the speed with which it affects the abandonment of apparently inferior alternatives. If competencies are diffused, competence gains from experience are shared through a population, and there is less chance that a potentially good alternative will become extinct before its potential is discovered. A more refined model of diffusion and selection would have to take account of the network of linkages among organizations in a population (MacArthur and Wilson 1967, Rogers 1995). Some of the more relevant elaborations involve elements of "partially permeable membranes" that are found in networks of associations. In a typical network, different parts of the system evolve partly (but not completely) independently. By virtue of being partly separated, they are likely to evolve in different directions. By virtue of being partly connected, the favorable outcomes in one can spread ultimately, but not too rapidly, to another. In this case, there is some chance that risky or new alternatives that are discarded by most will be adopted by some—from whom others can ultimately adopt the alternative (Wright 1931, Weick 1976).

We have not explored elaborations that involve these considerations of the diffusion process, including variations that would attend to issues of legitimacy (DiMaggio and Powell 1983, Tolbert and Zucker 1983). Nor have we explored the effect imitation of a particular organizational behavior can have on the average performance of organizations exhibiting this behavior. In a behavioral model of risk taking and change, such changes in performance could also have systematic effects on risk taking and the probability of change (March 1988, March and Shapira 1992, Shapira 1995, Mezias and Eisner 1997, Greve 1998).

Implications for Risk Taking and Change in Organizations

The simulations demonstrate that behaviors that are often labeled as conservatism or risk aversion may reflect biases in adaptive processes. Because the biases result from the tendency of learning and selection to reproduce successful actions, the biases will be weakened whenever failures are more likely to be reproduced. Thus, they will be reduced in situations where, for whatever reason, adaptation is slowed, made imprecise, or recalled less reliably. In this sense, the survival of risky and new alternatives will be associated with ineffective adaptation, and several forms of organizational ineffectiveness can increase the propensity of an organization to engage in risky and new activities.

Speed of Adaptation

The results suggest that slow adaptation will increase the frequency of risky and novel activities. By reproducing successful alternatives, fast learners quickly eliminate poor-performing alternatives in favor of better ones. The very quickness that makes such a process particularly effective, however, also makes it biased against risky and new alternatives. By continuing to reproduce failures, slow learners are less likely than their quicker cousins to suffer from the hot stove effect.

In the case of the model of stochastic learning, slow adaptive response is represented by the learning parameter, a. Slow learning, where a is relatively small, leads to more gradual adaptation of the probability of choosing different alternatives and thus less profound bias against risky and new alternatives. In the case of the model of competitive selection and reproduction, slow adaptive response is represented by the selection pressure, w, and the reproductive responsiveness parameter, h. Slow reproductive response, where h is relatively small, leads to more gradual adaptation and thus less profound bias against risky and, in particular, new alternatives. The impact of variations in selection pressure, w, on the other hand, is mixed. Smaller selection pressures lead to less bias against new alternatives with low risk than do more substantial selection pressures, but they lead to greater bias against risky alternatives. The latter result stems from the way the outcomes from risky alternatives are arrayed at the extremes of the outcome space and thus are especially vulnerable if only a few actors are eliminated.

We ask whether there are features of organizational life that affect the rate of learning or selection. Of particular importance in this regard are features of organizational inertia (Hannan and Freeman 1984) that delay reactions to feedback about new alternatives. Adaptation is slowed by any rule, practice, norm, heuristic, or information processing bias that allows an alternative to survive for some number of observations regardless of what those observations are. Such inertia is likely to attach to new, untried alternatives in situations in which the organizational enthusiasm required to adopt a practice assures that early unfavorable results will be overlooked or ignored. This inclination to persist is likely to be augmented by individual and organizational biases toward confirming initial expectations (Rabin and Schrag 1999), and by the costs of political renegotiations required for changing organizational action. As a result, alternatives with poor initial performance may survive sufficiently long for their benefits (if any) to be discovered.

Similarly, in situations where a selection process, regardless of its intensity, operates on a type only after a period of time, types with low performance can survive for a considerable time. For example, poorly performing organizations endure into adolescence by drawing on their initial endowments of resources and expectations (Fichman and Levinthal 1991, Levinthal 1991), by locating themselves within the institutional protections of legal and political systems (Meyer and Zucker 1989), and by exploiting the informational limitations of the relevant financial, labor, and product markets. The longer the honeymoon (nonadaptive) period, the slower the adaptive process and the weaker the bias against risky and new alternatives.

Adaptation is also slowed by beliefs associated with such things as social legitimacy, ideology, and imagination. Strong beliefs with respect to a course of action systematically buffer individuals and organizations from learning, particularly from adverse evaluations of a course of action. Consider, for example, the role of management fashions. Communities of belief surrounding organizations allow types to survive (at least in the short run) if their actions fulfill expectations of appropriate behavior. Organizational practices that have become fashionable or socially legitimate within the culture of management persist in the face of adverse evidence about their effectiveness (Meyer and Rowan 1977, Meyer and Zucker 1989, Hannan and Carroll 1992, Abrahamson and Fairchild 1999). In the case of a potentially valuable practice, social legitimization allows survival long enough for the value to be developed or exhibited.

Ideology and imagination are similar belief-based sources of delay that can lead to avoiding the premature abandonment of an alternative that would ultimately prove valuable (Hirschman 1967, Kolakowski 1968). Because they are slow to learn from adverse experience, ideologues are likely both to do poorly on average and to be the first to discover valuable novel and risky innovations. Imagination plays a comparable role. The ability and will to deviate from established ways involve being able to conjure a different world, to see things with different eyes. This talent for deviance is, however, of little use by itself because a person or organization that easily rejects ideas will flit from one to another, abandoning each new idea long before it can become or be identified as a good one. The second component of imagination—the inclination to stick with an idea despite evidence of its inadequacy—is equally important. Without such tenacity, new ideas wither. Thus, it is necessary to combine a willingness to deviate from old ideas with an irrational pigheadedness in support of the new (slow learning). Imagination does this by substituting a fantastical world for an observed world, a substitution that usually leads to disaster but occasionally leads to discovery (March 1995).

Precision of Adaptation

The results suggest that imprecise adaptation will increase the probability that an adaptive system will engage in risky and new activities. Imprecision in adaptation increases the probability of reproducing failures. The reproduction of actions that have led to failure results in all of the familiar adverse consequences (Levinthal and March 1981, Lant and Mezias 1990, Lant 1994), but it also results in reducing the bias against new and risky alternatives.

Suppose that information on the outcome of a learning trial is subject to random error. That is, suppose in the case of the stochastic learning model that the realized outcome is as indicated above, plus a draw from a normal distribution with mean zero and standard deviation, E. Then, the discriminations of learning are confused by noise in outcomes. Figure 5 compares the proportion of learners choosing the risky alternative over 50 periods for various values of E. Increasing the noise increases the fraction choosing the risky alternative in a learning situation. The sensitivity to E is more gradual than that displayed in Figure 5 if learning depends on the magnitude of success or failure.

Suppose, similarly, that the performance that determines the elimination of less successful organizations in competitive selection and reproduction is subject to random error. As in the case of learning, the discriminations of selection are confused by noise in outcomes. Figure 6 compares the proportion of the population choosing the risky alternative over 50 periods for various values of E, where reproduction is according to the numbers of each type surviving as shown in Equation (5) and $h = 0.5$. Increasing the noise (reducing the precision of selection) increases the fraction choosing the risky alternative in a competitive selection situation.

Figure 5 **The Effect of *E* on the Reproduction of Reliability Through Experiential Learning. Based on Averages from 11 Sets of 5,000 Simulations Where** $X = 10$, $Y = 10$, $S = 10$, $a = 0.2$, **and** $b = 0.5$.

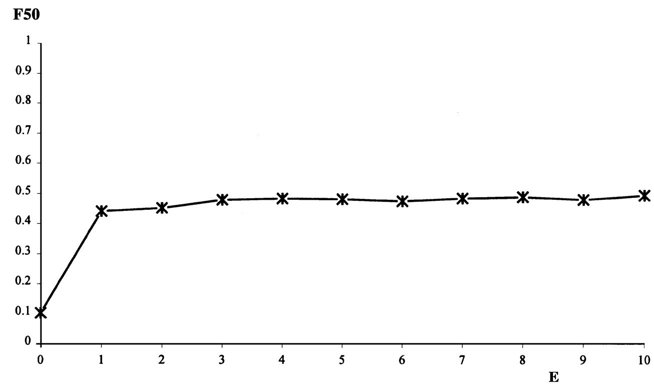
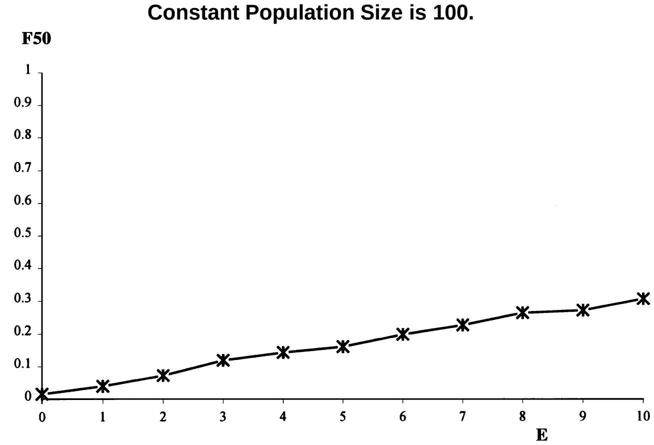


Figure 6 **The Effect of *E* on the Reproduction of Reliability Through Competitive Selection. Based on Averages from 11 Sets of 100 Simulations Where** $X = 10$, $Y = 10$, $S = 10$, $w = 0.2$, $h = 0.5$. **The Constant Population Size is 100.**



The results exhibited in Figures 5 and 6 are not surprising, given what we have seen previously. Because the biases against risky and new alternatives are produced by learning and selection, making the feedback on which those processes depend less precise reduces the biases. If reducing the biases is important to long-run effectiveness, it is also important to reduce the short-run precision of learning and selection (Hedberg et al. 1976, Mezias and Glynn 1993). Noise generates errors in the feedback on which adaptation is based and produces failures (eliminations) and successes (survivals) that are arbitrary relative to the true potentials at the time. These arbitrary "mistakes" in adaptation can be destructive in the short run, but they reduce the biases of adaptation against risky and new alternatives, thus providing advantages in the long run under some conditions.

Organizational contexts of adaptation provide a variety of examples of common phenomena that result in such errors (Aldrich 1999). First, rather than being tied to past history, action can be remarkably arbitrary with respect to any history of successes or failures with previous actions. For example, action is disconnected from experience by a context of rules. Actors function within social structures that develop and impose rules that are independent of the learning or the selection history of the individual actor who executes them. As a result, alternatives may survive even though they have been unsuccessful in the past.

Second, even if action is responsive to historical experience, the interpretation of experience may be indeterminate. Histories are confounded by causal ambiguity, ignorance, and bias. If, as in the case of risky and new alternatives, accurate interpretation of past experience would indicate that the action has not been successful, causal ambiguity, ignorance, and bias may result in more exploration. Alternatives that might reasonably be rejected on the basis of clearer feedback can survive. Causal indeterminacy, complexity, and bias are also important factors in competitive selection. Individual attributes with negative performance implications can survive by "hitchhiking" in a package of attributes with good performance (Gould and Lewontin 1984, Barnett 1997).

Third, even if the interpretation of experience is straightforward, coding of outcomes as successes or failures is made indeterminate by the existence of multiple, incommensurable objectives and criteria. This ambiguity of objectives makes it possible for alternatives with poor performance on any one criterion to survive by exhibiting good performance on other criteria. An innovation survives by attaching itself to one set of criteria until it has developed capabilities with respect to a different set (Amabile 1988).

Finally, the coding of outcomes is made indeterminate by the fact that success and failure are often defined in relative terms. If success is measured relative to an aspiration level, higher aspirations will lead to more failures than lower aspirations and thereby to more and briefer experimentation. Similarly, a selection regime in which only a small fraction of the population survives each time period, and most individuals are eliminated, favors risky alternatives. The result is a variation on a familiar proposition that in a competition for primacy, the importance of variability relative to expected value increases (Lippman and Rumelt 1982, March 1991, Mezias and Glynn 1993, Frank and Cook 1995).

Memory of Adaptation

The results also suggest that poor memory will increase the probability that an adaptive system will engage in risky and new activities. The hot stove effect stems from the stabilization of knowledge concerning alternatives that have resulted in early failures. Because such alternatives are not chosen subsequently, the early unfavorable impressions are not corrected. However, if memory is not perfect, or if the reputations of alternatives are not segregated from each other, the effect will be attenuated. Adaptation stores memories of experience in rules and reputations of alternative rules (Levitt and March 1988, Huber 1991, Walsh and Ungson 1991, March et al. 2000). The memory of failures tends to make their reproduction unlikely, and the lack of reproduction preserves the memory. On the other hand, if failures are remembered poorly, or if the reputation of a particular alternative is affected by experiences with other alternatives, an alternative that is initially unsuccessful can overcome its early bad reputation (March 1973, Nystrom and Starbuck 1984).

Two phenomena associated with adaptive memory are likely to increase the later reproduction of early failures. The first is the decay of memory of past experience. Decays in memory are a characteristic feature of human recall, but perhaps even more so in organizations (Huber 1991). Failures of the past are likely to be forgotten or remembered poorly. As a result, traces of early failure are likely to fade in recall. Organizations with imperfect memories are likely to repeat the failures of the past, with all the attendant negative consequences (Neustadt and May 1986), but they are also less likely to reproduce reliability and the status quo (Hedberg et al. 1976, Brunsson and Olsen 1993). Variations in turnover of personnel or authorities (March 1991, Carley 1992), in documentation of the past (Huber 1991), in departmentalization and fragmentation (Huber 1991), and in estimates of the rate of change in the world are all features of an organizational setting that are likely to lead to variations in the reliability of memory.

The second feature of adaptation that is likely to increase the later reproduction of early failures is the generalization of experience. Experiences (good or bad) with one alternative can spread to other alternatives. When this happens, differences among alternatives are muted, with consequent improvement in the relative position of alternatives that have themselves experienced failures. The most obvious mechanism is the generation of self-confidence (Miller 1993). Success with an alternative is presumably some mixture of the potential of the alternative, the capabilities of the individual or organization in executing that alternative, and good fortune. Individuals and organizations are likely to overestimate the contribution of their capabilities to the outcome and underestimate the contribution of the alternative and luck. Because capabilities can be imagined to transfer to other alternatives, success on one alternative tends to inflate expectations of success on other alternatives. The result is that the relative likelihood of choosing an alternative rejected early increases somewhat over time for actors who are successful in other domains.

Managing Reactions to Risk and Change

Most proposals for managing reactions to risk and change start from the premise that individual and organizational propensities to take or avoid risk and to embrace or resist change are buried in their utility functions and are susceptible primarily to countervailing incentives. The fact that models of experiential learning and competitive selection can produce behavior that looks like risk aversion and resistance to change does not negate the role of individual, organizational, and cultural predispositions in producing observed risk taking and change behavior. However, it does suggest that the search for levers by which to control the taking of risks and resistance to change should include attention to adaptive processes and particularly to persistence with alternatives having poor early outcomes.

Because the bias against risky and new activities is a product of the tendency to reproduce actions that have been successful, the bias will be reduced whenever failures are more likely to be reproduced. Thus, the bias will be reduced in situations where, for whatever reason, adaptation is slowed, made imprecise, or recalled less reliably. By examining the sources of slow, imprecise adaptation and unreliable recall, we have implicitly suggested how several features of adaptive ineffectiveness can increase the propensity of organizations to engage in risky and new activities.

For example, to the extent that inertia and ideology lead to slow adaptation, inertia and ideology will be associated with increased propensities to engage in risky activities and persist in new ones. Similarly, to the extent that causal complexity and ignorance lead to imprecise adaptation, causal complexity and ignorance will be associated with increased propensities to engage in risky activities and persist in new ones. More generally, any organizational characteristic that leads to persistence with alternatives that exhibit poor early results will increase the likelihood of risk taking and change in that organization. Because adaptation is precisely a process for reproducing successes, these characteristics are, for the most part, characteristics of short-run inefficiency in adaptation.

Conclusion

We have shown that biases against new and risky alternatives may be less properties of individuals, organizations, or cultures than of learning and of competitive selection and reproduction themselves. The basic adaptive mechanism that produces biases in favor of reliable alternatives and the status quo is the differential reproduction of actions that have led to success. Because alternatives with a record of good performance are reproduced, alternatives with high potential that initially do poorly are likely to be avoided in the future. Early misrepresentation of the relative values of alternatives can thus have a lasting effect.

Because they exhibit variability in results, risky alternatives are likely to be misestimated on the basis of small samples of experience. Because they are likely to require practice to realize their full potential, new alternatives are likely to be underestimated on the basis of small samples of experience. Thus, an adaptive system will exhibit behavior that looks like risk aversion and resistance to change even in situations in which the long-run expected returns from new alternatives and risky alternatives are notably higher than those from existing alternatives and reliable alternatives. Because the biases are a product of the tendency to reproduce successes, the biases will be smaller whenever failures are more likely to be reproduced. Thus, according to this argument, the survival of risky and new alternatives will be associated with slow and imprecise adaptation and unreliable recall, all classical characteristics of short-run inefficiency in adaptation.

These features of adaptation pose a difficult dilemma for organizations, one that is more likely to be confronted by an ecology of micro-organizational phenomena than by deliberate strategic calculation, particularly because the latter involves knotty intertemporal and interpersonal comparisons. Without resolving how that dilemma might be faced in detail, we can observe that the biases we have identified as stemming from rapid, precise adaptation, are relatively easy to justify in a world in which risky and new alternatives are systematically and persistently poor ones. In such a world, the benefits of increasing risk taking and reducing resistance to change are likely to be smaller than the costs. On the other hand, in a world in which risky and new alternatives have a fair chance of being good ones, any unconditional enthusiasm for fast, precise learning is likely to place an organization at a longer-term competitive disadvantage.

Acknowledgments

The research has been supported by grants from the Spencer Foundation. The authors are grateful for the comments of Sidney Winter, participants at seminars at the Wharton School, the Prince Bertil Symposium, the Conference on Knowledge and Innovation in Helsinki, the editor, and two anonymous reviewers.

Appendix. List of Symbols

- $A_{i,t}$ The average performance (outcomes) for all survivors of type j in period t
- E Noise in realized outcomes; the standard deviation of the normally distributed error
- F_t The fraction of the population having a particular property at period t
- L_t The aspiration level at period t
- $N_{j,t}$ The number of organizations of type j in period t
- O_t The performance outcome) at period t
- P_t The probability of choosing the risky alternative (or the alternative that requires practice) at period t
- S The base line standard deviation of performance
- $T_{i,t}$ The aggregate performance (outcomes) for all survivors of type j in period t
- X The expected performance (outcome) of the risky alternative or the alternative that requires practice
- Y The expected performance (outcome) of the certain alternative or the established alternative
- a The rate of adjustment in the probability of choosing the risky alternative (or the alternative that requires practice)
- b The rate of adjustment in the aspiration level
- c_t The competence in period t
- d The competence learning rate
- h Parameter regulating the sensitivity of reproduction to the average performance (outcome) of a type
- k Parameter regulating the rate of reduction in the standard deviation
- $r_{j,t}$ The probability in period t that a replacement will be of type j
- w The proportion of the population eliminated in each period

Endnote

[1]A complete list of short definitions of the symbols used in the models is provided in the Appendix.

References

Abrahamson, E., G. Fairchild. 1999. Management fashion: Lifecycles, triggers, and collective learning processes. *Admin. Sci. Quart.* **44**(4) 708–214.

Aldrich, H. 1999. *Organizations Evolving*. Sage Publications Inc., Thousands Oaks, CA.

Amabile, T. M. 1988. A model of creativity and innovation in organizations. B. Staw, L. L. Cummings, eds. *Res. Organ. Behavior.* **10** 123–167.

Argote, L., S. L. Beckman, D. Epple. 1990. The persistence and transfer of learning in industrial settings. *Management Sci.* **36**(2) 140–154.

——, D. Epple. 1990. Learning curves in manufacturing. *Sci.* **247** (Feb.) 920–924.

Arthur, B. W. 1989. Competing technologies, increasing returns, and lock-in by historical events. *Econom. J.* **99**(394) 116–131.

Baldwin, J. R. 1998. *The Dynamics of Industrial Competition: A North American Perspective*. Cambridge University Press, Cambridge, MA.

Bandura, A. 1977. *Social Learning Theory*. Prentice Hall, Englewood Cliffs, NJ.

Barnett, W. P. 1997. The dynamics of competitive intensity. *Admin. Sci. Quart.* **42**(1) 128–160.

Blume, L. E., D. Easley. 1995. What has the rational learning literature taught us? A. Kirman, M. Salmon, eds. *Learning and Rationality in Economics*. Blackwell, Oxford, U.K.

Brunsson, N., J. P. Olsen. 1993. *The Reforming Organization*. Routledge, London, U.K.

Busemeyer, J. R., K. N. Swenson, A. Lazarte. 1986. An adaptive approach to resource allocation. *Organizational Behavior and Human Decision Processes* **38**(3) 318–341.

Bush, R. R., F. Mosteller. 1955. *Stochastic Models for Learning*. John Wiley & Sons, New York.

Carley, K. 1992. Organizational learning and personnel turnover. *Organ. Sci.* **3**(1) 20–46.

Carroll, G. R., R. J. Harrison. 1994. Historical efficiency of competition between organizational populations. *Amer. J. Sociology* **100**(3) 720–749.

Coombs, C. H., R. M. Dawes, A. Tversky. 1970. *Mathematical Psychology: An Elementary Introduction*. Prentice-Hall, Englewood Cliffs, NJ.

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

David, P. A. 1985. Clio and the economics of QWERTY. *Amer. Econom. Rev.* **75**(2) 332–337.

DiMaggio, P. J., W. W. Powell. 1983. The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *Amer. Soc. Rev.* **48**(2) 147–160.

Easley, D., N. M. Kiefer. 1988. Controlling a stochastic process with unknown parameters. *Econometrica* **56**(5) 1045–1064.

Elster, J. 1984. *Ulysses and the Sirens: Studies in Rationality and Irrationality*, 2nd ed. Cambridge University Press, Cambridge, U.K.

Feller, W. 1968. *An Introduction to Probability Theory and Its Applications*, vol. 1, 3rd ed. John Wiley & Sons, New York.

Fichman, M., D. A. Levinthal. 1991. Honeymoons and the liability of adolescence: A new perspective on duration dependence in social and organizational relationships. *Acad. Management Rev.* **16**(2) 442–468.

Frank, R. H., P. J. Cook. 1995. *The Winner-Take-It-All Society: Why Few at the Top Get So Much More Than the Rest of Us*. The Free Press, New York.

Friedman, M. 1953. The Methodology of Positive Economics. *Essays in Positive Economics*. Chicago University Press, Chicago, IL.

Gimeno, J., T. B. Folta, A. C. Cooper, C. Y. Woo. 1997. Survival of the fittest? Entrepreneurial human capital and the persistence of underperforming firms. *Admin. Sci. Quart.* **42**(4) 750–783.

Goldberg, D. E. 1989. *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley, Reading, MA.

Gould, S. J., R. C. Lewontin. 1984. The spandrels of San Marco and the Panglossian paradigm: A critique of the adaptationist program. E. Sober, ed. *Conceptual Issues in Evolutionary Biology*. MIT Press, Cambridge, MA. 252–270.

Greve, H. R. 1996. Patterns of competition: The diffusion of a market position in radio broadcasting. *Admin. Sci. Quart.* **41**(1) 29–60.

——. 1998. Performance, aspirations, and risky organizational change. *Admin. Sci. Quart.* **43**(1) 58–86.

Haleblian, J., S. Finkelstein. 1999. The influence of organizational acquisition experience on acquisition performance: A behavioral learning perspective. *Admin. Sci. Quart.* **44**(1) 29–56.

Hannan, M. T., G. R. Carroll. 1992. *Dynamics of Organizational Populations*. Oxford University Press, Oxford, U.K.

——, J. Freeman. 1977. The population ecology of organizations. *Amer. J. Soc.* **82**(5) 929–964.

——, ——. 1984. Structural inertia and organizational change. *Amer. Soc. Rev.* **49**(2) 149–164.

Haunschild, P. R., A. S. Miner. 1997. Modes of interorganizational imitation: The effects of outcome salience and uncertainty. *Admin. Sci. Quart.* **42**(3) 472–500.

Haveman, H. A. 1992. Between a rock and a hard place: Organizational change and performance under conditions of fundamental environmental transformation. *Admin. Sci. Quart.* **37**(1) 48–75.

Hedberg, B. L., P. C. Nystrom, W. H. Starbuck. 1976. Camping on seesaws: Prescriptions for a self-organizing organization. *Admin. Sci. Quart.* **21**(1) 41–65.

Herriott, S. R., D. A. Levinthal, J. G. March. 1985. Learning from experience in organizations. *Amer. Econom. Rev. Proc.* **75**(2) 298–302.

Hirschman, A. O. 1967. *Development Projects Observed*. The Brookings Institution, Washington, D.C.

Holland, J. H. 1975. *Adaptation in Natural and Artificial Systems*. The MIT Press, Cambridge, MA.

Huber, G. P. 1991. Organizational learning: The contributing processes and the literatures. *Organ. Sci.* **2**(1) 88–115.

Kaelbling, L. P. 1993. *Learning in Embedded Systems*. The MIT Press, Cambridge, MA.

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

Kolakowski, L. 1968. *Toward a Marxist Humanism: Essays on the Left Today*. Grove Press, Inc, New York.

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

——. 1994. Computer simulations of organizations as experiential learning systems: Implications for organizational theory. K. Carley and M. Prietula, eds. *Computation Organizational Theory*. Erlbaum Associates, Hillsdale, NJ. 195–216.

——, S. J. Mezias. 1990. Managing discontinuous change: A simulation study of organizational learning and entrepreneurship. *Strategic Management J.* **11** (summer) 147–179.

——, ——. 1992. An organizational learning model of convergence and reorientation. *Organ. Sci.* **3**(1) 47–71.

Lave, C. A., J. G. March. 1975. *An Introduction to Models in the Social Sciences*. Harper & Row, New York.

Levinthal, D. A. 1991. Random walks and organizational mortality. *Admin. Sci. Quart.* **36**(3) 397–420.

——. 1997. Adaptation on rugged landscapes. *Management Sci.* **43**(7) 377–415.

——, J. G. March. 1981. A model of adaptive organizational search. *J. Econom. Behavior Organ.* **2**(4) 307–333.

——, ——. 1993. The myopia of learning. *Strategic Management J.* **14**(2) 95–112.

Levitt, B., J. G. March. 1988. Organizational learning. *Ann. Rev. Soc.* **14** 319–340.

Lippman, S. A., R. Rumelt. 1982. Uncertain imitability: An analysis of interfirm differences in efficiency under competition. *Bell J. Econom.* **13** (Autumn) 418–438.

Lounamaa, P., J. G. March. 1987. Adaptive coordination of a learning team. *Management Sci.* **33**(1) 107–123.

MacArthur, R. H., E. O. Wilson. 1967. *The Theory of Island Biogeography*. Princeton University Press, Princeton, NJ.

March, J. G. 1973. Model bias in social action. *Rev. Ed. Res.* **42**(4) 413–29.

——. 1988. Variable risk preferences and adaptive aspirations. *J. Econom. Behavior and Organ.* **9**(1) 5–24.

——. 1991. Exploration and exploitation in organizational learning. *Organ. Sci.* **2**(1) 71–87.

——. 1994. *A Primer on Decision Making*. The Free Press, New York.

——. 1995. The future, disposable organizations, and the rigidities of imagination. *Organ.* **2** 427–440.

——. 1996. Learning to be risk averse. *Psych. Rev.* **103**(2) 309–19.

——, Z. Shapira. 1992. Variable risk preferences and the focus of attention. *Psych. Rev.* **99**(1) 172–183.

——, M. Schulz, X. Zhou. 2000. *The Dynamics of Rules: Change in Written Organizational Codes*. Stanford University Press, Stanford, CA.

Meyer, J. W., B. Rowan. 1977. Institutionalized organizations: Formal structure as myth and ceremony. *Amer. J. Sociology* **83**(2) 340–363.

Meyer, M. W., L. G. Zucker. 1989. *Permanently Failing Organizations*. Sage, Newbury Park, CA.

Mezias, S. J., A. B. Eisner. 1997. Competition, imitation, and innovation: An organizational learning approach. *Adv. Strategic Management* **14** 261–294.

——, M. A. Glynn. 1993. Three faces of corporate renewal: Institution, revolution, and evolution. *Strategic Management J.* **14**(2) 77–101.

——, T. K. Lant. 1994. Mimetic learning and the evolution of organizational populations. J. A. C. Baum, J. V. Singh, eds. *Evolutionary Dynamics of Organizations*. Oxford University Press, New York 179–198.

Miller, D. 1993. The architecture of simplicity. *Acad. Management Rev.* **18**(1) 116–138.

Miner, A. S., P. R. Haunschild. 1995. Population level learning. L. L. Cummings, M. Barry, eds. *Research in Organizational Behavior* **17** 115–166.

——, S. J. Mezias. 1996. Ugly duckling no more: Pasts and futures of organizational learning research. *Organ. Sci.* **7**(1) 88–99.

——, S. V. Raghavan. 1999. Interorganizational imitation: A hidden engine of selection. B. McKelvey, J. A. C. Baum, eds. *Variations in Organization Science: In Honor of Donald T. Campbell*. Sage Publications, London, U.K. 35–62.

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

Neustadt, R. E., E. R. May. 1986. *Thinking in Time: The Uses of History for Decision-Makers*. Free Press. New York.

Nystrom, P. C., W. H. Starbuck. 1984. To avoid organizational crises, unlearn. *Organ. Dynam.* **12** (Spring) 53–65.

Rabin. M., J. L. Schrag. 1999. First impressions matter: A model of confirmatory bias. *Quart. J. Econom.* **114**(1) 37–82.

Rogers, E. M. 1995. *Diffusion of Innovations*, 4th Ed. The Free Press, New York.

Rothschild, M. 1974. A two-armed bandit theory of market pricing. *J. Econom. Theory* **9**(2) 185–202.
Shapira, Zur. 1995. *Risk Taking: A Managerial Perspective.* Russell Sage, New York.
Sutton, R. S., A. G. Barto. 1998. *Reinforcement Learning: An Introduction*. The MIT Press, Cambridge, MA.
Tolbert, P. S., L. G. Zucker. 1983. Institutional forces of change in the formal structure of organizations: The diffusion of civil services reform, 1880–1935. *Admin. Sci. Quart.* **28**(1) 22–39.
Twain, M. 1897. *Following the Equator: A Journey Around the World.* American Publishing Co, Hartford, CT.
Usher, J. M., M. G. Evans. 1996. Life and death along gasoline alley: Darwinian and Lamarckian processes in a differentiating population. *Acad. Management J.* **39**(5) 1428–1466.
Wald, A. 1947. *Sequential Analysis*. John Wiley & Sons, New York.
Walsh, J. P., G. R. Ungson. 1991. Organizational Memory. *Acad. Management Rev.* **16**(1) 57–91.
Weick, K. E. 1976. Educational organizations as loosely coupled systems. *Admin. Sci. Quart.* **21**(1) 1–19.
Wright, S. 1931. Evolution in Mendelian populations. *Genetics* **16**(1) 97–159.
Yelle, L. E. 1979. The learning curve: Historical review and comprehensive survey. *Decision Sci.* **10**(2) 302–328.

Accepted by Theresa Lant.