DRUID Working Paper No. 13-09

E Pluribus Unum:

Organizational Size and the Efficacy of Learning

By

Hart E. Posen, Dirk Martignoni and Daniel A. Levinthal

Danish Research Unit for Industrial Dynamics

www.druid.dk

E Pluribus Unum:
Organizational Size and the Efficacy of Learning

Hart E. Posen
School of Business
University of Wisconsin-Madison
E-mail: hposen@bus.wisc.edu

Dirk Martignoni
Department of Business Administration
University of Zurich
E-mail: dirk.martignoni@uzh.ch

Daniel A. Levinthal
Wharton School
University of Pennsylvania
E-mail: levinthal@wharton.upenn.edu

January 2013

Abstract:
Learning from experience is a central theme in the management literature. While in general experiential learning is viewed as efficacious, the literature increasingly points to the difficulties inherent in the learning process — many of which stem from a deficit of information about the merits of alternative solutions. It seems plausible that larger organizations, with their capacity to simultaneously pursue a variety of potential solutions to a given challenge, may overcome this deficit. Such a perspective suggests that the efficacy of an organization's learning process should be an increasing function of organizational size. While this logic is intuitively appealing, we find that it does not fully capture the nuances of the organizational learning process. We employ a computational model and find that larger organizations, as characterized by their scale in pursuing parallel initiatives: (a) explore less than smaller organizations, (b) are less likely to discover the very best alternative, and yet (c) on average identify better alternatives. Increasing the number of parallel initiatives guides the search process towards viable alternatives, but it does so at the cost of inhibiting search breadth. Thus, in our model, the characteristics of learning by larger organizations do not result from differences in inertia or incentives that may impede learning and innovation, but rather from the properties of the organizational learning process itself.

ISBN 978-87-7873-349-8

DRUID society

Acknowledgements
The authors would like to thank seminar participants at the Academy of Management Conference 2012, DRUID Conference 2012, LMU Munich, Strategic Management Society Conference 2012, Tilburg Conference on Innovation 2012, and the University of Michigan (Strategy Seminar, Psychology-Decision Consortium Seminar) for helpful comments. We wish to thank Annabelle Grewer, Markus Lang, and Sidney Winter for assistance in
improving the paper.

1. Introduction

Organizations have the capacity in terms of human and financial capital to simultaneously pursue a variety of potential solutions to any given challenge. Organizational units, for example, a firm's production facilities, geographic divisions, or technology development centers, may individually seek solutions to the same or related problems. These units may hold different beliefs about the best path forward and engage in processes of experiential learning to identify possible solutions (March and Olsen 1975, March, Sproull, and Tamuz 1991). It seems natural to assume that because these units are within the boundaries of one organization, they may pool their experience during the learning process, and collectively identify solutions superior to those that an individual unit alone could identify. Indeed, to the extent that organizational size is defined associated with the number of units seeking solutions to the same or related problems, this logic also suggests that the efficacy of an organization's learning process should be an increasing function of organizational size (when size is related to the degree of parallelism).[1] While this logic seems sensible and intuitive, we demonstrate that increasing the number of units within an organization seeking solutions to a given problem may diminish, rather than enhance the efficacy of experiential learning. As a consequence, differences in innovativeness across organizations of varying size may result not from differences in inertia or incentives that may impede learning, but rather from the properties of the collective learning process.

The simultaneous pursuit of multiple potential solutions to a given problem (Nelson 1961) and the pooling of experience is a phenomenon widely observed in organizations, across employees (March 1991, Greve 2003, Puranam, Raveendran, and Knudsen 2012), teams or groups (Kim & Burton 2002, Rivkin 2001, Taylor and Greve 2006), and projects (Eggers 2012). For expositional purposes, we focus on learning by organizational units, examples of which may include: (1) multiple project teams in a large technology firm seeking uses for the same set of basic technologies; and (2) multiple geographically

[1] This conception of size is consistent with work that highlights the internal structure of the organization as a metric of size, rather than the number of employees, sales, or assets (e.g., Burton, Minton, and Obel 1991, Kimberly 1976).

dispersed production facilities within a large firm, each producing the same products and seeking to improve production efficiency or reduce defects. In each of these examples, learning by individual units gives rise to the opportunity to pool experience across units about partial solutions, effective approaches, and dead-ends. This collective learning is largely an organizational phenomenon because knowledge tends to flow more freely within an organization than across organizational boundaries (Kogut and Zander 1992). Thus, while multiple independent firms may seek solutions to the same problem (Afuah and Tucci 2012, Greve 1995, 2005, Rivkin 2000, Csaszar and Siggelkow 2010), and sometimes collaborate in doing so, the potential and ability to pool experience in an inter-organizational context is more limited.

The question at the heart of this study is: why does organizational size alter the efficacy with which firms find attractive solutions to the challenges that they face? For instance, anecdotal evidence often points to the limited innovativeness of larger organizations, and folk wisdom, like "the big can, the small do," is prominent in the popular press.[2] More systematic evidence suggests that while the extent of effort and the variety of potential solutions tends to increase with organization size, it does so at less than a proportional rate (e.g., Scherer 1965, Bound, Cummins, Griliches, Hall, and Jaffe 1982, Katila and Ahuja 2002). A number of existing arguments in the literature speak to the observation that larger organizations may be less innovative than small organizations, these include: (a) a greater degree of bureaucracy and inertia (e.g., Hannan and Freeman 1984), (b) reduced incentives to innovate due to self-cannibalization (e.g., Reinganum 1983), and (c) the challenges of implementing incentive-intensive employment contracts (Zenger and Lazzarini 2004). We offer an alternative, complementary, explanation for this phenomenon — we point to the properties of learning when larger organization size enables organizations to pursue multiple, simultaneous solutions to any given challenge.

On the surface, organizational learning seems an unlikely theoretical tool to explain why large organizations learn and innovate less effectively than small organizations. Indeed, intuitively, one might expect that increasing the number of organizational units pursuing solutions to a given problem should

[2] http://www.economist.com/blogs/freeexchange/2011/06/innovation

enhance the efficacy with which an organization discovers attractive solutions, putting aside issues of inertia and incentives. The theoretical logic underlying this intuition rests on the observation that organizational units engage in learning in task environments in which the true value of an alternative cannot be known with certainty (e.g., Benner and Tripsas 2012, March 1996, Denrell and March 2001, Puranam, Powell, and Singh 2006, Knudsen and Levinthal 2007, Posen and Levinthal 2012). Under uncertainty, learning units face an experience constraint (March, Sproull, and Tamuz 1991) that engenders problems of both limited depth of experience with any particular alternative, and limited breadth of experience across the set of alternatives (Turner, Bettis, and Burton 2002).

The problem of limited depth arises because any particular experience with an alternative may be misleading or unrepresentative to the extent there is variation in possible outcomes, and small samples of experience will exacerbate this effect (March, Sproull, and Tamuz 1991). This leads to an over-emphasis on a chance negative outcome and the possible premature foreclosure of a search process (Denrell and March 2001). The problem of limited breadth arises because other, as yet poorly understood, alternatives may be superior to the currently preferred choice. This need to garner experience with a broader set of alternatives generates the trade-off between exploration and exploitation (March 1991). In an organizational setting, the experience of individual units can be pooled, and one unit may rely upon the experience gained by other units sampling the same and other alternatives. As such, increasing the number of units in an organization may lead to a deeper and broader base of experience, potentially alleviating the problems of limited depth and breadth.

While the logic above is intuitively appealing, we find that it does not fully capture the nuances of the organizational learning process. Our model suggests that the consequences of increasing the size of organizations, when size is associated with the number of units seeking solutions to similar challenges, are three-fold. First, larger organizations explore less broadly than do smaller organizations. Second, larger organizations are less likely to discover the very best alternative (superior extreme performance). Third, despite reduced exploration and maximal performance, larger organizations on average identify better alternatives (superior mean performance). These results do not simply reflect declining returns to increasing organization size. Rather, these results reflect an inherent trade-off between mean and extreme performance that emerges endogenously from the process of experiential learning.

The intuition underlying our results is as follows. The problems of limited depth and breadth derive from a deficit of information about potential alternative solutions. As March et al. (1991: 1) note "history is not generous with experience". Increasing the number of organizational units seeking solutions to a problem tends to alleviate this issue, but in doing so it may cause another issue — an abundance of information about moderately attractive alternative solutions. In particular, increasing organizational size may lead to an over-emphasis on a chance positive outcome from a good, but not necessarily outstanding, alternative achieved by one unit in the organization that may "seduce" other units away from potentially superior alternatives. As a consequence, increasing organizational size alters not only the number of alternatives explored, which we call the quantity of exploration, but also the quality of alternatives explored. The returns to larger size then hinge on how increasing the number of units alters the relative strength of these two effects.

We anchor our formal development on the multi-armed bandit model. This model is the canonical representation of exploration and exploitation under conditions of uncertainty (Holland 1975). In the management literature, March formulates much of his discussion of learning in terms of the bandit model (Denrell and March, 2001, March 1996, 2003, 2010). The bandit model takes its name from a slot machine analogy in which a unit seeks to maximize the flow of returns over time. In each period, a unit chooses one alternative from a set of policy alternatives, with the payoff to a choice reflecting a draw from a probability distribution with an unknown, alternative-specific mean. A unit chooses based on its beliefs about the expected returns to each of the alternatives. Thus, a unit is portrayed as possessing a mental model or cognitive map derived from its own prior experience, which encapsulates their understanding of the merits of the available set of choices. We extend this standard, single-unit bandit model to a multi-unit model — an organization consisting of multiple, individual learning units. We assume that units in an organization pool experience such that a unit forms beliefs about the relative merits of alternatives, not only based on its own experience, but also on the experience of others in the organization. Within this structure, we examine the implications of increasing the number of units in an organization.

This paper proceeds as follows. In the next section, we describe the theoretical background and setup of the multi-unit, multi-armed bandit computational model. In section 3, we present the results of simulation experiments in which we examine the properties of the model of experiential learning as a function of organization size (number of units engaged in parallel search). In the final section, we discuss the implications of these results for theory and practice.

2. Model

The bandit model (Gittins 1979, Berry and Fristedt 1985), which is the basis of our analysis, has been the subject of significant study because its underlying structure captures core issues in a variety of realistic economic situations ranging from research settings such as R&D (Hardwick and Stout 1992), to strategic issues, such as product pricing (Bergemann and Välimäki 1996), and consumer choice (Gans, Knox, and Croson 2007). The bandit model has recently begun to emerge in the management literature (e.g., Denrell and March 2001, Posen and Levinthal 2012), enabling a stronger connection (and docking) between results in management and other disciplines (Burton and Obel 1995).

There are two common features underlying economic problems that are modeled in a bandit framework. First, information about the returns to an alternative can only be gathered by sampling it. Second, feedback from samples is subject to uncertainty that gives rise to variation in possible outcomes, and as such, any particular experience may be misleading or unrepresentative.

Formally, the bandit model reflects a sequential choice problem where, at each point in time, t, a unit must choose among N alternatives. The realized outcome of a particular choice is a draw from an unknown, alternative specific, probability distribution. If the process is Bernoulli, then the choice of alternative n results in an outcome of $r = (1, 0)$, reflecting a positive outcome of $r = 1$ with probability p_n and a negative outcome $r = 0$ with probability $1\text{-}p_n$. As such, the state of the environment can be described by the vector of payoffs to the alternatives, $P = [p_1, \ldots, p_N]$.

Consider unit *m's* beliefs, $q_{m,n,t}$, about alternative n at time t, which are encapsulated by the vector $Q_{m,t} = [q_{m,1,t}, \ldots, q_{m,N,t}]$ where $0 \leq q_{m,n,t} \leq 1$. To refine these beliefs and maximize the value of the stream of rewards, the unit engages in learning. In the initial period t=0, unit m has prior beliefs, $Q_m(\text{t=0}) = [q_{m,1}, \ldots, q_{m,N}]$, about the value of the N reward outcome probabilities $P = [p_1, \ldots, p_N]$. In particular, we assume that units have initial beliefs across the alternatives, $Q_m(\text{t=0}) = [q_{m,1}, \ldots, q_{m,50}]$, that are homogeneous and set to 0.5, which is equal to the mean value of the actual distribution of payoffs. In each subsequent period, the unit makes a choice from the set of alternatives. By acting – making a choice of an alternative – the unit m receives feedback from the environment in the form of an outcome signal as a success or failure, $r_{m,n} = (1,0)$.

We assume that beliefs at any given point in time reflect the average reward over a unit's entire history of samples of a given alternative (March 1996, Sutton and Barto 1998). This simple average updating is a special case of the more general fractional adjustment updating methodology (Bush and Mosteller 1955). Thus, for a unit, m, beliefs about the merits of alternative n at time t, reflect the average reward history (yielding rewards $r_1, r_2, \ldots, r_{k_n}$) associated with the k_n samples of that alternative. As such:

$$q_{n,t} = \frac{\sum_{s=1}^{k_n} r_{n,s}}{k_n}, \qquad (1)$$

where k_n denotes the number of times alternative n has been sampled by period t. In any period, t, only the belief about the currently sampled alternative, n, is updated; for all other alternatives, j, beliefs remain unchanged such that $q_{j,t} = q_{j,t-1}$. Under this assumption, the belief updating process employing the average reward history (per Equation 1) is equivalent to Bayesian updating.

In an organization, units may not only rely on their own experiences but also on the experiences of all other units. We denote ω as the extent to which the unit h weights its own experience compared to the aggregated experiences of all other units in the organization. Unit h's estimate of the value of alternative n in period t is given by the linear combination of its individual experience with this alternative (first term in equation (2), weighted by 1- ω), and the aggregated experience of all other units (second term in equation (2), weighted by ω):

$$q_{h,n,t} = (1-\omega)\left(\frac{\sum_{s=1}^{k_{h,n}} r_{h,n,s}}{k_{h,n}}\right) + \omega\left(\frac{\sum_{m=1}^{M}\sum_{s=1}^{k_{m,n}} r_{m,n,s}}{\sum_{m=1}^{M} k_{m,n}}\right), \tag{2}$$

where m indexes individual units in the organization of size M. This belief updating mechanism implies that all units' experiences with a given alternative contribute equally to belief formation.

As a result, when ω=0, all units make choices based on beliefs that arise from their own experience alone. In contrast, when ω=1, all units in an organization make choices based on beliefs that weigh their own and others' experience equally.[3]

In each period, each unit, h, independently chooses alternatives to sample based on their beliefs, $q_{h,n,t}$. While there is a wide variety of plausible exploration strategies, perhaps the simplest and best-known strategy is that of selecting, at each point in time, the alternative with the highest belief, max(q_1,…, q_N), reflecting the highest expected reward (Auer, Cesa-Bianchi, and Fischer 2002, Sutton and Barto 1998). This rule is "greedy" (Sutton and Barto 1998) in the sense that at each point in time it maximizes the expected reward in the next period. If they are indifferent across alternatives, because several alternatives appear equally attractive, units randomly select one of them (Sutton and Barto 1998).[4] In our sensitivity analysis, we relax this assumption of "greedy" search behavior, and allow units to engage in either random exploration that is uninformed by beliefs (as in Csaszar and Siggelkow 2010,

[3] Note that when ω=0, a multi-unit organization is identical to an equal number of atomistic units.

[4] In our sensitivity analysis, we examine the possibility that individual units employ less greedy choice rules. Our results are largely robust to this possibility.

Ethiraj and Levinthal 2009, Levinthal 1997) or exploration where random choice is weighted by beliefs about the relative merits of alternatives (Posen and Levinthal 2012, Sutton and Barto 1998).

We use the multi-unit and multi-armed bandit model described above to analyze the effect of organizational size on learning and exploration. The structure of the simulation requires initializing both the opportunity structure of the task environment and the actors' beliefs about the merits of alternatives.

The opportunity structure of the environment is defined by initializing the payoff probabilities for each alternative. We formulate a 50-arm bandit model such that the vector of alternatives' payoff probabilities is $P = [p_1, \ldots, p_{50}]$. Each alternative is allocated a payoff probability, p_n, that is a draw from a uniform distribution [0,1]. This produces a distribution of payoff probabilities across alternatives that is symmetrical with a mean of 0.50 and standard deviation of 0.29. The choice of a particular alternative n leads to a positive/negative outcome with probabilities (p_n, $1-p_n$), and an associated reward of $r_{i,t}=(1,0)$.

We examine organizations ranging in size from M=1 through M=10 units. Each organization searches on a uniquely specified task environment. In our main analysis, we assume that within an organization units pool their experience and update beliefs according to Equation 2 where ω=1. To average over the stochastic outcome, we examine fifty thousand iterations for each level of size. The simulation runs for 200 periods by which time a steady state is reached. To make sensible comparisons across organizations of different sizes, we normalize by the number of units in the organization.

3. Analysis

In the following analysis, we begin by examining the impact of size (number of units searching in parallel) on performance outcomes, focusing on the mechanisms through which the size of the organization alters the dynamics of exploration and the efficacy of learning. In the first experiment, we make two assumptions about how units learn and choose among alternatives: (a) units are greedy, exploiting the alternative they deem to currently be the best, only exploring if they are indifferent among a set of such alternatives[5]; and (b) units act on the basis of beliefs that result from the aggregation of the experience of all other units (ω=1). In the sensitivity analysis, we relax these assumptions, allowing units to explore more broadly, and to weight their own experience more strongly than the experience of others in the organization (ω<1).

Main Experiment

In Figure 1, we plot two performance metrics (averaged over the first 200 periods). Performance is measured as the average cumulative reward stream for all units in the organization. Discovery of the best alternative is measured as the cumulative probability that at least one unit in the organization has sampled the best alternative.

< Insert Figure 1 about here >

The key observation in the figure above is the paradox of size – larger organizations (those with more units searching in parallel) are less likely to discover the best alternative, but on average generate higher performance.[6] This paradox of size emerges even though we hold effort constant across organizations of different sizes; that is, we divide both the cumulative rewards and the probability of discovering the best alternative by the number of units in the organization.

Our objective in the subsequent analysis is to examine how increasing organizational size gives rise to this paradox. Our model setup rules out many of the usual explanations for this paradox, including: (a) a greater degree of bureaucracy and inertia (e.g., Hannan and Freeman 1984), (b) reduced incentives to innovate due to self-cannibalization (e.g., Reinganum 1983), and (c) the challenges of implementing incentive-intensive employment contracts (Zenger and Lazzarini 2004). We focus on the process of

[5] In the bandit model, units are assumed to make choices based on their beliefs about the relative merits of alternatives. For simplicity, in this analysis we focus on units pursuing a greedy strategy, always choosing the alternative on which they have the highest belief, and choosing randomly between alternatives that are believed to be equally good. Relaxing the greedy assumption by allowing for additional exploration (e.g., ε probability of choosing a different arm at random, a strategy called ε-greedy) increases the level of exploration, but does not change the basic intuition of the results developed here.

[6] The exception is organizations of size two, which we will return to later in the analysis.

organizational learning under uncertainty, and analyze two key factors that may alter the efficacy of search and learning: the quantity of exploration and the quality of alternatives explored.

Quantity of Exploration

We begin by examining differences in the extent to which organizations of different sizes explore the set of alternatives, which we term the quantity of exploration. Because units are modeled as greedy (i.e., choosing the alternative that appears most promising), they are likely to under-explore the set of alternatives. As such, we consider the possibility that pooling experience in an organization increases the quantity (extent) of exploration, moving the units closer to the optimal balance between exploration and exploitation.

< Insert Figure 2 about here >

In Figure 2, we examine two metrics of the quantity of exploration. In Panel A, we plot the cumulative number of exploration events. Specifically, this is the cumulative count of events in which a unit chooses a different alternative in period t+1 than it chose in period t. To normalize results across organizations of different sizes, we measure the quantity of exploration of ten units organized in different arrangements: one ten-unit, five two-unit, two five-unit, and one ten-unit organization. Exploration decreases from 29.8 change events for ten one-unit organizations to 17.6 for one ten-unit organization. The key result of this figure is that the size-normalized exploration of larger organizations is forty percent less than that of smaller organizations.

To explain why units in larger organizations are less likely to change alternatives over time, we also track the extent to which, conditional on changing alternatives, the alternative to which they change is new to the organization. For organizations of size one, exploration usually entails sampling alternatives not previously sampled (the black portion of the bar is 5 times out of nearly 30 change events). In contrast, when units in a large organization explore by choosing an alternative different from its current choice, this new alternative is rarely new to the organization. In a large organization, the positive experience of one unit with an alternative leads to an increased probability that others in the organization will sample that alternative. Likewise, when one unit has a negative experience with an alternative, others in the organization are less likely to pursue it. This result is consistent with Levinthal and March's (1993) observation that the best strategy for any unit is often to exploit the successful explorations of other units. Pooling experience in organizations tends to generate directed exploration, which leads to the early identification of a relatively good alternative within the restricted consideration set. Yet, at the same time, it also leads to a rapid decrease in the level of exploration.

In Figure 2 Panel B, we examine the implications of directed exploration for the breadth and depth of exploration by focusing on the number of unique alternatives explored, rather than the number of exploration events. The two polar cases of organizing ten units generate substantially different patterns of exploration, again normalized such that exploration reflects one ten-unit or ten one-unit organizations. A population of ten, one-unit organizations explores more than 28 different alternatives (of the 50 available), while a ten-unit organization explores less than ten alternatives, a 66% reduction in exploration. Moreover, larger organizations devote a disproportionate share of their exploration effort to testing an alternative only once. If we consider only alternatives sampled at least twice, organizations of ten units explore 78% fewer alternatives than those explored by ten one-unit organizations.

Quality of Exploration and Average Performance

The quantity of exploration is insufficient to explain why larger organizations exhibit superior average performance. As such, we examine the conjecture that larger organizations explore better quality alternatives.

In Figure 3, we plot the distribution of alternatives sampled, conditional on changing alternatives. The horizontal axis indicates the quality of the alternative sampled (ranked from best, rank=1, to worst, rank=50). The vertical axis indicates the probability that an alternative of a given rank is selected. For one-unit organizations, the choice of alternatives is uniform across the distribution of alternative payoffs. In contrast, larger organizations tend to choose better alternatives when they explore. In this sense, larger organizations exhibit a higher quality of exploration.

< Insert Figure 3 about here >

The quality of exploration increases with size for two reasons. First, a unit in an organization responds to rewards received by other units in the organization. When an alternative is sampled, a positive reward suggests that the alternative is unlikely to be extremely bad, while a negative reward suggests that it is unlikely to be extremely good (Kim and Miner 2007). Pooling such experience guides units in an organization toward good alternatives and away from bad ones. This effect increases with the size of the organization.

Second, by concentrating exploration on a narrow set of alternatives, directed exploration from pooled experience generates more accurate beliefs about the relative merits of alternatives. Because the challenge faced by organizations is one of evaluative uncertainty, multiple samples of an alternative will provide more accurate estimates (beliefs) of its true value. Larger organizations explore a smaller quantity of alternatives, but on the alternatives they explore they engage in repeated samples (per Figure 2 Panel B). This behavior on the part of larger organizations generates more accurate beliefs and gives rise to a virtuous cycle as search effort is further refined and focused on alternatives that appear to be of high value.[7] Moreover, directed exploration and repeated samples tend to mitigate the "hot stove effect," in which alternatives that generate a single negative reward are unlikely to be sampled again by a unitary unit (Denrell and March 2001, March, Sproull, and Tamuz 1991).

Combined Effect of the Quantity and Quality of Exploration

The pooling of experience in organizations tends to reduce the quantity of exploration, but increases its quality. Independently, these two effects function in opposite directions. Decreasing the quantity of exploration alone reduces performance, while increasing the quality of exploration increases performance. Over a broad range of organizational sizes, from two through ten units, the net of this trade-off is an increase in average performance. However, recall from Figure 1 that increasing size from one to two units

[7] This dynamic further reinforces the process that limits the quantity of exploration in larger organizations. Not only do larger organizations develop more accurate beliefs early in the search process, they also focus on superior alternatives. A more accurate belief about a good alternative is higher in magnitude than an accurate belief about a mediocre alternative.

decreases average performance, suggesting that the marginal effect of a change in quantity or quality of exploration may not be constant.

< Insert Figure 4 about here >

To examine this possibility, in Figure 4 we graph the individual contributions of the quantity and quality of exploration, and their combined effect. In the top panel, the quantity of exploration (right scale) is measured by the number of times a unit chooses a different alternative in period t than was chosen in period t-1. This difference is decreasing with organizational size, from 2.98 for ten one-unit organizations to 1.76 for a ten-unit organization. The quality of exploration (left scale) is measured by the average difference between the actual value of the new alternative explored and the alternative abandoned, conditional on exploring. For a one-unit organization, an exploration event leads to the selection of an alternative that is, on average, 0.11 better than the prior alternative, while for a ten-unit organization the average improvement is 0.19.

The combined returns to increasing size reflect the product of the quantity and quality of exploration, which we plot in the lower panel of Figure 4. Consistent with our hypothesis that the impact of size is the result of both the quality and quantity of exploration effects, the resulting graph reconstructs the qualitative pattern in Figure 1. We observe that an increase in size from one-unit organizations to two-unit organizations decreases average performance because the benefits of increased quality of exploration are more than offset by the decrease in the quantity of exploration. Further increases in size have, on net, positive average performance implications because quality effects dominate.

In sum, increasing organizational size has two competing effects on experiential learning: it decreases the quantity of exploration and increases the quality of exploration. When the latter effect dominates the former, increasing size can lead to superior average performance while also decreasing the level of exploration.

Quality of Exploration and the Discovery of the Best Alternative

Up to this point, we have examined the mechanisms that give rise to the observation that larger organizations, on average, find superior solutions. We have yet to examine why larger organizations are less likely to discover the very best solution.

The pooling of experience in organizations is a double-edged sword with respect to the efficacy of experiential learning. On one side of the blade, pooling experience directs exploration to a deep search of a relatively small portion of the set of alternatives. This is beneficial when the problem is one of uncertainty in evaluation and shallow sampling of an alternative may generate beliefs that do not accurately reflect the true value of an alternative.

On the other side of the blade, pooling experience reduces the quantity of exploration and decreases its breadth. This decline in exploration quantity and breadth accounts for the decreased discovery probability with increasing organizational size. Moreover, the observed reduction in discovery probability understates the cost associated with pooling experience in larger organizations. Increasing size not only reduces discovery, but also increases the probability that, conditional on discovery, the best alternative will be abandoned.

In Figure 5, we report the probability that a unit explores a different alternative in period t+1, conditional on having chosen the best alternative in period t (averaged over the first 10 periods). While one-unit organizations rarely abandon the best alternative (less than one percent), the probability of abandoning the best alternative increases with size, reaching nearly nine percent for ten-unit organizations.

< Insert Figure 5 about here >

Why do organizations discover but then sometimes abandon the best alternative? Consider a one-unit organization. A positive reward on an individual unit's choice always reinforces current behavior, reducing the probability of exploring (abandoning the current preferred alternative). Likewise, a negative reward on an individual unit's current choice always increases the likelihood of exploring. Both of these mechanisms are independent of size.

A second dynamic happens in multi-unit organizations, where the rewards received by other units also affect subsequent choices by the focal unit. Consider the situation in which a single unit in a multi-unit organization gets a positive reward after choosing the best alternative, but by chance another unit in the same organization gets a positive reward after choosing a mediocre alternative. Holding initial beliefs constant, the unit that had previously chosen the best alternative is now less likely to do so again in the subsequent period, because it may be seduced by the mediocre alternative that appears to be at least equally attractive. Thus, while others' failures are known to have positive or negative performance implications (Miner, Ji-Yub, Holzinger, and Haunschild 1999), so too can others' successes.

Consider this result in the context of Greve's (1995) observation, in a study of radio broadcasting, that strategy abandonment is contagious across firms when there is significant uncertainty about the merits of the alternative strategies. Our results point to the mirror observation, that a unit currently choosing the best alternative may abandon it if another unit receives positive feedback on an inferior alternative.

To examine this mechanism, we disaggregate the abandonment phenomenon by splitting it into abandonment of the best alternative that occurs after the unit has a negative reward from the best arm (shaded black), and those abandonments that occur after it has a positive reward (shaded grey) in Figure 5. For a one-unit organization, abandonment can only occur after a chance negative reward on the best alternative. The impact of abandonment based on a negative signal from the best alternative is reduced with size (declining height of the black bars) because directed exploration generates deeper sampling of the explored alternatives, and thus more accurate beliefs. In contrast, for larger organizations, almost all abandonment events occur subsequent to the unit getting a positive reward on the best alternative. This phenomenon increases with size.

Underlying this phenomenon is the evolution of beliefs in an organization. In Figure 6, we graphically illustrate how beliefs and choices change over time. To do so, we track the temporal pattern of beliefs of one ten-unit organization. The graphs, in Figure 6, provide two examples that illuminate the pattern of belief formation and the choices of units. The vertical axis reflects time periods, while the horizontal axis reflects the alternative payoffs ranked from best (rank=1) to worst (rank=50). The color reflects the units' beliefs about the alternatives.[8] Darker grey reflects alternatives that are believed to be better (have a higher payoff), while lighter grey reflects those alternatives believed to be worse (lower payoff).

< Insert Figure 6 about here >

In Panel A, we examine a case in which pooling of experience leads to the exploration of higher-quality alternatives. At the start (*t*=0), the ten units in the organization have uninformative priors, and as such, they are indifferent across alternatives. They randomly try alternatives numbered 3, 9, 13, 21, 26, 30, 42, 46 (twice), and 48.[9] Recall that the alternatives are ranked from best to worst (left to right on the x-axis). The units receive a positive reward on alternatives 3, 13, 21, and 26 (on alternative 46, the organization receives one positive and one negative reward). As a result, all units in the organization update their beliefs on these alternatives upwardly as indicated by the slightly darker shading. The units receive negative rewards on the rest of the alternatives, as indicated by the slightly lighter shading. Because the units explore greedily, in *t*=2 they select only from the set of alternatives 3, 13, 21, and 26, for which they hold superior beliefs. In this next sample, alternatives 3, 13, and 21 each generate two positive rewards, while alternative 26 generates zero net reward (two positive and two negative). Thus, in the third period, the organization only selects from alternatives 3, 13, and 21. Alternative 3 generates three positive rewards, alternative 13 a zero net reward (one positive and one negative), and alternative 21 one net positive reward (three positive and two negative). Now, the units are no longer indifferent among

[8] Since there is a complete pooling, all units in the organization hold identical beliefs.

[9] Alternative 46 is sampled by two units, generating one positive and one negative reward and, as a result, beliefs remain unchanged for their initial level. This alternative takes the same color as the background grey.

several alternatives: alternative 3 has consistently generated only positive rewards. As a result, alternative 3 has the highest estimated value and all ten units in the organization converge to this alternative, which continues to generate net positive rewards in subsequent periods. In this way, pooled experience in organizations guides units towards better alternatives.

In Panel B, we examine a case in which pooling of experience leads to both exploration of higher quality alternatives and, sometimes, to myopic behavior in which the best alternative is tried and then rejected. In the initial period, all units are indifferent across all alternatives and choose randomly. By chance in this run, one unit in the organization tries the optimal choice, alternative 1, and garners a positive reward. Other units initially choose alternatives 12, 14, 18, 24, 29 and 33, also garnering positive rewards. The remaining units try alternatives 25, 42, and 47, which generate negative rewards. In the next period, all ten units select among the alternatives that generated positive rewards in the prior period, because they have identical high beliefs on these alternatives. They randomly select alternatives 12, 14, 18, 24, 29, and 33. In doing so, they abandon alternative 1 even though it is the best alternative, and despite the fact no unit received a negative reward from this alternative. They do so because other units in the organization generated positive rewards for other (inferior) alternatives. The ten units get one positive reward on alternatives 12 and 14, net zero positive rewards on 18, and two negative rewards on 24, net one positive reward on alternative 29, and one negative reward on alternative 33. At this point, both alternatives 12 and 14 are believed to be superior. The process continues, and the units in the organization converge on alternative 12. While this alternative is relatively good,[10] it is not the best alternative, which was tried and abandoned.

In sum, pooling experience in organizations guides search towards better alternatives. Not only does this reduce the rate of exploration, it also increases average performance. However, these gains are not without costs, as the same process tends to deflect larger organizations away from the very best

[10] With 50 arms drawn from a uniform distribution [0.0,1.0], the alternative ranked first has an expected payoff of 0.98 and the alternative ranked last has an expected payoff of 0.02. Thus, rank=12 alternative translates to a long-run expected performance of 0.98-11*0.02=0.76.

alternatives.

Pooling of Experience

A key assumption in our model is that organizational units act on the basis of beliefs that result from the aggregation of the experience of all units in the organization. In the earlier results, we set ω=1 in equation 2, implying that units update their beliefs by weighting their own and others' experience equally. Of course, this need not be the case. We consider the implications of relaxing this assumption, by examining ω=1.0 and ω=0.50.

< Insert Figure 7 about here >

In the bottom panel of Figure 7, we examine the quantity of exploration, measured as changing alternatives. Because with ω<1, units weigh their own experience more highly than that of other units in the organization, units act more independently. As such, the quantity of exploration increases, but only very slightly. This small increase in exploration occurs because, in decreasing ω, a unit is slightly less directed in is exploration, more willing to sample an alternative that has not been sampled by others in the organization.

A decrease in weighting of others' experience leads to a substantial increase in average performance, but only a small increase in the probability of discovering the best alternative (top panel in Figure 7). Average performance increases because lower weighting on other units' experience reduces the likelihood that a unit will prematurely abandon a relatively good alternative.

More generally, in decreasing the weight each unit places on other units' experience, the pattern of mean versus extreme performance (discovery of the best alternative), the paradox of size, is still evident. Indeed, the trade-off, in terms of the area between the curves indicating the average performance and the probability of discovering the maximum, grows stronger as the weighting of other's experience declines.

Sensitivity to Alternative Choice Behavior

In the above analysis, we imposed an important assumption on the individual behavior of units in organizations. Organizational units are postulated to be greedy, always exploiting the alternative currently deemed best, and only exploring if they are indifferent among a set of such alternatives. In this sensitivity analysis, we examine the implications of relaxing this assumption. We confine our discussion here to the qualitative results, but supply more detailed quantitative results in the appendix.

Earlier, we found that larger organizations explore less than smaller organizations. Would increasing the level of exploration in the form of a less greedy strategy change this result? To examine this possibility, we employ two mechanisms to induce additional exploration.

First, we consider a setting in which units in the organization pursue an overt exploration strategy whereby they engage in random exploration (Csaszar and Siggelkow 2010, Ethiraj and Levinthal 2009, Levinthal 1997) independent of their beliefs with a probability of five percent. This exploration strategy is typically referred to as ε-greedy, where in this case we are setting $\varepsilon=0.05$. The earlier experiments with a greedy strategy are equivalent to setting $\varepsilon=0$ (Sutton and Barto 1998). In terms of the average level of exploration across size, increasing ε from 0 to 0.05 generates a ten-fold increase in the quantity of exploration. Nonetheless, as with a greedy strategy, exploration decreases in organizational size. As before, average performance increases with size, while the probability of discovering the best alternative decreases with size. Thus, the paradox of size persists, even with a search strategy that ensures some baseline level of exploration.

We examine a second setting in which units in the organization pursue an overt exploration strategy, but in this second setting there is a level of intelligence, or non-randomness, in the exploration behavior. In particular, we examine organizations pursuing an exploration strategy that weights the alternative chosen for exploration by the current beliefs regarding the merits of the alternative. We employ the softmax choice rule attributable to Luce (1959) and employed widely (Camerer and Ho 1999, Gans et al. 2007, Sutton and Barto 1998, Vermorel and Mohri 2005, Weber et al. 2004, Posen and Levinthal 2012).[11] The softmax choice rule is tunable (like ε-greedy), ranging from purely exploitative when $\tau=0$ (equivalent

[11] This strategy formulation takes the form of random choice based on a Gibbs (Boltzmann) distribution. The probability of selecting alternative i, m_i, is defined as $m_i = e^{(q_i/\tau)} / \sum_{i=1}^{N} e^{(q_i/\tau)}$.

to $\varepsilon=0$) to increasingly exploratory as τ increases ($\tau>>0$). Using the softmax choice rule, units do not blindly explore, as they do with an ε-greedy choice rule, rather increasing τ increases the propensity to explore conditional on the current set of beliefs about the merits of alternatives. The results for a modestly exploratory strategy of $\tau=0.05$ suggest that employing this more intelligent exploratory strategy increases average performance and the probability of discovering the best alternative, relative to the results from Experiment 1. Nevertheless, the paradox of size previously identified continues to persist in this setting.

In sum, we find that pooling of experiences within organizations affects both the quantity of exploration and the quality of exploration. This generates the paradox of size – an explicit trade-off in which size confers both an advantage to average performance and a disadvantage to the probability of discovering the best alternative.

4. Discussion

Organizations often operate in task environments in which the true value of an alternative cannot be known with certainty (e.g., March 1996, Denrell and March 2001, Knudsen and Levinthal 2007, Posen and Levinthal 2012). In this context, a boundedly rational organization attempts to learn from experience (March and Olsen 1975, March, Sproull, and Tamuz 1991, March 2010). While experiential learning is in general efficacious, it suffers because "history is not generous with experience" (March, Sproull, and Tamuz 1991: p.1). This creates problems of both limited search depth and limited search breadth (Levinthal and March 1993). The challenge, epitomized in the inherent need to balance exploration and exploitation (March 1991), is finding a means to reap the benefits of experiential learning while overcoming its limitations.

Organizations have the capacity in terms of human and financial capital to simultaneously pursue a variety of potential solutions to any given challenge. Individual organizational units, for example, production facilities, geographic divisions, or R&D projects, each holding potentially different beliefs about the best path forward, may seek solutions to the same, or related problems. Thus, while learning from experience can be conceptualized as an individual activity, more often it reflects organized collections of units seeking solutions to the same, or related, problems. It seems natural to assume that, because units within an organization can pool their experience in the learning process, they can together achieve outcomes superior to those that they could achieve alone. This logic implies that the efficacy of learning should be an increasing function of the number of units over which experience is pooled.

Our results demonstrate that this logic is only partially correct. Efforts to relax the experience constraint, by increasing the number of units over which experience is pooled, are a double-edged sword. Pooling experience alters not only the number of alternatives explored, which we call the quantity of exploration, but also the quality of alternatives explored. The returns to larger size then hinge on how increasing the number of units alters the relative strength of these two effects.

While our model is general in the sense that collective learning with experience pooling occurs in a broad variety of settings, including individuals in teams or groups, units in an organization, and firms in an industry, we wish to focus on the implications in the context of organizations. Consider the three basic properties of our model. Larger organizations (a) explore less than smaller organizations, (b) are less likely to discover the very best alternative, and yet (c) despite reduced exploration and maximal performance, on average identify better alternatives. These properties, which we refer to as the paradox of size for experiential learning, offer important insight regarding the common intuition about how exploration and innovation scale, or do not scale, with organizational size.

Our results point to the possibility that larger organizations may exhibit difficulties in learning and innovating, not because of a deficit of information about potential alternative solutions, but rather due to an abundance of information. A large organization is likely to have identified a promising set of alternative solutions, with this likelihood increasing with the size of the organization. In contrast, a small organization has a more modest experience base, not only in terms of a possible modest history with any particular alternative (limited depth), but also, in a cross-sectional sense, a modest number of units carrying out experiential samples across the set of alternatives (limited breadth). This observation has two consequences for larger organizations. First, failure, or negative feedback with respect to a given alternative, is likely to lead to a search of the inventory of possibly more promising alternatives already present within the organization, rather than the broader search that a smaller organization may pursue because it harbors fewer promising alternatives to which to turn in the face of negative feedback. Second, a chance positive outcome achieved by one unit in the organization, sampling an alternative that is only modestly attractive, may "seduce" other units away from potentially superior alternatives. Indeed, even positive direct experience with the best alternative may be swamped by the positive experiences of others. Thus, increasing organizational size alters the dynamics and efficacy of learning as a result of the unintended consequences of possessing an abundance of information about alternatives.

While we have examined *how* organization size matters for the efficacy of experiential learning, we wish to be careful about the conditions under which our conclusions hold. In this study, we examine an organization learning under conditions of uncertainty. Given uncertainty, we find that the returns to larger organizational size hinge on the relative strength of a quantity and quality of exploration effects. Yet the challenge of uncertainty is only one of two distinct challenges to the efficacy of learning. The other challenge is that of variety.

The challenge of variety arises when experiential learning necessitates search among a vast number of alternative choices (e.g., Bruderer and Singh 1996, Ethiraj, Levinthal, and Roy 2008, Levinthal 1997, March 1991, Rivkin 2000, 2001, Lenox, Rockart, and Lewin 2006). A large variety of solutions exist when there are many, potentially interdependent, policy choice dimensions on which the firm needs to make decisions and, as a result, variety grows in a combinatoric fashion. While variety is present in our model, because there are N possible arms from which to choose, the extent of variety is modest relative to the extent of uncertainty. Given a large variety of alternatives, good alternatives may be distant from the currently preferred alternative and potentially overlooked by a myopic search process, and no individual unit can possibly search the entire space of alternatives. Thus, search resembles the task of finding a needle in a haystack. In models of search that focus on the problem of variety (Fang, Lee, and Schilling 2010, March 1991, Miller, Zhao, and Calantone 2006), the efficacy of experiential learning is a function of the diversity of alternatives explored. As a corollary of this property, to the extent that increasing the number of units in the organization serves to increase diversity (Afuah and Tucci 2012), increasing organizational size would enhance the efficacy of experiential learning. Relatedly, Taylor and Greve (2006 p.724) examine "whether structures that lead to variance-enhancing behaviors differ from those that lead to higher mean performance." They focus on the recombination of policy choice dimensions, arguing that increasing the number of units increases the discovery of very good alternatives. Thus, while increasing organizational size engenders a trade-off between mean and extreme performance when the challenge of learning entails both uncertainty and, to a lesser extent some degree of variety, a challenge of combinatory variety will tend to respond more positively to increased organizational size. Thus, the implications of organization size for search processes are importantly contingent on the nature of the task environment.

Our results are also related to work on bandwagons because, like our study, this literature is interested in how the choices of one actor are affected by the observable choices of other actors. In this line of work, bandwagons may be "rational" in that observing others' choices may provide information about the relative merits of alternatives (Banerjee 1992, Terlaak and King 2007), or social in that the perceived value of an alternative is assumed to be a direct function of the number of adopters (Abrahamson and Rosenkopf 1993). Our model differs from these in that units observe not only the choices of others, but also the outcome of these choices. Indeed, this is what we mean by experience pooling. Thus, neither social returns nor observable choice behavior alone drive the results of our model. Rather, units' choices and outcomes affect the nature of information about the set of alternatives. As such, our model does not generate an aggregate result that is good or bad (e.g., diffusion of inherently good or bad innovations). Instead, our model generates an aggregate result that is good or bad depending on an organization's objective function — average versus extreme performance.

Finally, our work also relates to the problem of allocating resources in organizations and economies. While a large organization may have sufficient resources to pursue multiple simultaneous projects that engender unique solutions to a problem, the likelihood of finding a very good (relative to an fairly good) solution is predicated on maintaining a separation between these projects and thereby limiting experience pooling. Separation engenders in an organization “multiple minds” in the sense of a variety of opinions about the path to a good solution. Yet the temptation to pool experience within an organization, to allow each of the projects to be informed of the others successes and failures, is hard to forestall. Illustrative of this challenge, an M.I.T. research scientist commenting on innovation in large pharmaceutical firms argued that “large size…can end up being an impediment…Very often when you are going for real innovation, …you have to go against prevailing wisdom, and it’s hard to go against prevailing wisdom when … you have some vice president who says, ‘No, that doesn’t make sense.’”[12] In our model, this vice president generates hierarchical pressure to pool experience. Such pooling tends to alter the learning paths of these projects, and the quality of outcomes achieved, not necessarily for the better. Thus, the challenge of the organization is that of maintaining multiple minds for some period of time, to reap the benefits of the parallel search efforts, while mitigating the cost of doing so. Alternatively, to the extent that such separation (limited pooling) is indeed difficult to maintain within organizations, then perhaps finding the very best solution to a problem is the province of markets, where pooling of experience is likely to be more limited (Kogut and Zander 1992).

In sum, our study attempts to begin to understand the implications of collective learning and experience pooling in organizations. Our starting point is that organizations have the capacity in terms of human and financial capital to simultaneously pursue a variety of potential solutions to any given challenge. We find that the impact of organizational size, when characterized as the number of units pursuing solutions to the same or related problems, is not unambiguously positive.

Our point is not to suggest that larger organizations are somehow inferior to smaller organizations (or vice versa). Rather, we have sought to build a basic understanding of the dynamic implications of increasing organization size with respect to processes of organizational learning. Much as Michel (1915)

[12] Seligson, H. New York Times. November 24. 2012. “Hatching Ideas, and Companies, by the Dozens at M.I.T.”

argued for the “iron law of oligarchy,” expressing his pessimism regarding the possibility of true democratic expression in a polity, we argue that the dynamics of collective learning processes temper the latent diversity of experience and opinion within organizations. Organization size will tend to tilt the inherent balance in an organization away from exploration and towards the exploitation end of the continuum, even in the absence of features of bureaucracy and incentive conflict that are typically invoked to explain the apparent conservatism of larger organizations. While we would not claim that the dynamics of collective learning are the singular or even necessarily primary basis of explanation, we would argue that this process represents an important contribution to our understanding of this phenomenon.

Bibliography

Abrahamson, E., & L. Rosenkopf. 1993. Institutional and competitive bandwagons: Using mathematical modeling as a tool to explore innovation diffusion. *Academy of Management Review*, 18, 487-517.

Afuah, A., & C. Tucci. 2012. Crowdsourcing as a solution to distant search. *Academy of Management Review*, 37(3), 355-375.

Auer, P., N. Cesa-Bianchi, P. Fischer. 2002. Finite-time analysis of the multi-armed bandit problem. *Machine Learning*. **47**(2–3) 235–256.

Banerjee, A. V. 1992. A simple model of herd behavior. *Quarterly Journal of Economics*, 107(3), 797-817.

Benner, M. J., & M. Tripsas. 2012. The influence of prior industry affiliation on framing in nascent industries: The evolution of digital cameras. *Strategic Management Journal*, 33(3), 277-302.

Berry, D. A., B. Fristedt. 1985. *Bandit Problems: Sequential Allocation of Experiments*. Chapman and Hall, London.

Bergemann, D., J. Välimäki. 2001. Stationary multi-choice bandit problems. *Journal of Economic Dynamics and Control*. **25**(10) 1585–94.

Bound, J., C. Cummins, Z. Griliches, B. Hall A. B. Jaffe. 1982. Who does R&D and who patents? In Z. Griliches (Ed.), *R&D, patents, and productivity*. National Bureau of Economic Research Cambridge, Mass. 21-54)

Bruderer, E., J. S. Singh. 1996. Organizational evolution, learning, and selection: A genetic-algorithm-based model. *Academy of Management Journal.* **39**(5) 1322–1349.

Burton, R. M., J. W. Minton, & B. Obel. 1991. Organizational size and efficiency: An information processing view. Scandinavian Journal of Management, 7(2), 79-93.

Burton, R., & B. Obel. 1995. The validity of computational models in organization science: From model realism to purpose of the model. *Computational & Mathematical Organization Theory*, 1(1), 57-71.

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

Camerer, C., T. H. Ho. 1999. Experience-weighted attraction learning in normal form games. *Econometrica* **67**(4) 827–874.

Csaszar, F., N. Siggelkow. 2010. How much to copy? Determinants of effective imitation breadth. *Organization Science.* **21**(3) 661-676.

Denrell, J., J.G. March. 2001. Adaptation as information restriction: The hot stove effect. *Organization Science.* **12**(5) 523-538.

Eggers, J.P. 2012. All Experience is Not Created Equal: Learning, Adapting and Focusing in Product Portfolio Management. *Strategic Management Journal*. **33**(3): 315-335.

Ethiraj, S.K., D.A. Levinthal, R. Roy. 2008. The dual role of modularity: Innovation and imitation, *Management Science.* **54**(5) 939-955.

Ethiraj, S.K., D. A. Levinthal. 2009. Hoping for A to Z while rewarding only A: Complex organizations and multiple goals. *Organization Science.* **20**(1) 4-21.

Fang, C., J. Lee, M. A. Schilling. 2010. Balancing exploration and exploitation through structural design: The isolation of subgroups and organizational learning. *Organization Science.* **21**(3) 625–642.

Gans, N., G. Knox, R. Croson. 2007. Simple models of discrete choice and their performance in bandit experiments. *Manufacturing Service Operations Management.* **9**(4) 383–408.

Gittins, J.C. 1979. Bandit processes and dynamic allocation indices. *Journal of the Royal Statistical Society. Series B (Methodological).* **41**(2) 148–177.

Greve, H. 1995. Jumping ship - the diffusion of strategy abandonment. *Administrative Science Quarterly*, **40**(3), 444-473.

Greve, H. R. 2003. *Organizational Learning from Performance Feedback: A Behavioral Perspective on Innovation and Change*. Cambridge University Press, Cambridge, UK.

Greve, H. R. 2005. Inter-organizational learning and heterogeneous social structure. O*rganization Studies*. 26 (7): 1025-1047

Hannan, M., and J. Freeman. 1984. Structural inertia and organizational Change. *American Sociological Review*, 49(2): 149-164.

Hardwick, J.P., Q.F. Stout. 1991. Bandit strategies for ethical sequential allocation. *Computing Science and Statistics*. **23** 421–424.

Holland, J. 1975. *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control & Artificial Intelligence*. University of Michigan Press, Ann Arbor.

Katila, R., & G. Ahuja. 2002. Something old, something new: A longitudinal study of search behavior and new product introduction. *Academy of Management Journal*, 45(6), 1183-1194.

Kim, J., & R. Burton. (2002). The effect of task uncertainty and decentralization on project team performance. *Computational & Mathematical Organization Theory*, 8(4), 365-384.

Kim, J. Y., & Miner, A. (2007). Vicarious learning from the failures and near-failures of others: Evidence from the U.S. Commercial banking industry. *Academy of Management Journal*, 50(3), 687-714.

Kimberly, J. R. 1976. Organizational size and the structuralist perspective: A review, critique, and proposal. Administrative Science Quarterly, 21(4), 571-597.

Knudsen, T., D.A. Levinthal. 2007. Two faces of search: Alternative generation and alternative evaluation. *Organizational Science*. **18**(1) 39-54.

Kogut, B., U. Zander, U. 1992. Knowledge of the firm, combinative capabilities, and the replication of technology. *Organization Science* **3**(3) 383-397.

Lenox, M., S. Rockart, A.Y. Lewin. 2006. Interdependency, competition, and the distribution of firm and industry profits. *Management Science*. **52**(5) 757-772.

Levinthal, D.A. 1997. Adaptation on rugged landscapes. *Management Science*. **43**(7) 934–950.

Levinthal, D. A., J. G. March. 1993. The myopia of learning. *Strategic Management J*. **14** 95–112.

Levinthal, D. A., & Wu, B. 2010. Opportunity costs and non-scale free capabilities: Profit maximization, corporate scope, and profit margins. *Strategic Management Journal*, 31(7): 780-801

Levitt, B., & J. March, J. 1988. Organizational learning*. Annual Review of Sociology*, 14, 319-340.

Luce, R. 1959. *Individual Choice Behavior: A Theoretical Analysis*. Wiley, New York.

March, J., L. Sproull, & M. Tamuz. 1991. Learning from samples of one or fewer. *Organization Science*, 2(1), 58-71.

March, J.G. 1991. Exploration and exploitation in organizational learning. *Organization Science*. **2**(1) 71–87.

March, J.G. 1996. Learning to be risk averse. *Psychological Review*. **103**(2) 309–319.

March, J.G. 2010. *The Ambiguities of Experience*. Cornell University Press, Ithaca, NY.

March, J.G., L.S. Sproull, M. Tamuz. 1991. Learning from samples of one or fewer. *Organization Science*. **2**(1) 1-13.

Michels, Robert. 1915. Political Parties: *A Sociological Study of the Oligarchical Tendencies of Modern Democracy*. Translated into English by Eden Paul and Cedar Paul. New York: The Free Press.

Miller, K.D., M. Zhao, R.J. Calantone. 2006. Adding interpersonal learning and tacit knowledge to March's exploration-exploitation model. *Academy of Management Journal.* **49**(4) 709-722.

Miner, A. S., K. Ji-Yub, I. W. Holzinger, and P. Haunschild. 1999. Fruits of failure: Organizational failure and population-level learning. *Advances in Strategic Management*, 16, 187-220.

Nelson, R. 1961. Uncertainty, learning, and the economics of parallel research and development efforts. *Review of Economics and Statistics*, 43(4), 351-364.

Puranam, P., B.C. Powell, H. Singh. 2006. Due diligence failure as a signal detection problem. *Strategic Organization* 4: 319

Puranam, P., M. Raveendran, T. Knudsen. 2012. Organization design: The epistemic interdependence perspective. Academy of Management Review, 37(3), 419-440.

Posen, H.E, D.A. Levinthal. 2012. Chasing a moving target: Exploitation and exploration in dynamic environments. *Management Science*. **58**(3):587-601

Posen, H.E., Lee, J., Yi, S. 2013. The power of imperfect imitation. *Strategic Management Journal*, 34(2), 149-164.

Scherer, F. M. 1965. Firm size, market structure, opportunity, and the output of patented inventions. *American Economic Review*, 55(5), 1097-1125.

Reinganum, J. 1983. Uncertain innovation and the persistence of monopoly. *American Economic Review*, 73(4): 741-748.

Rivkin, J.W. 2000. Imitation of Complex Strategies. *Management Science*. **46**(6) 824-844.

Rivkin, J.W. 2001. Reproducing knowledge: Replication without imitation at moderate complexity. *Organization Science*. 12(3) 274–293.

Sutton, R. S., A. G. Barto. 1998. *Reinforcement Learning: An Introduction*. MIT Press, Cambridge, MA.

Taylor, A., & H. Greve. 2006. Superman or the fantastic four? Knowledge combination and experience in innovative teams. *Academy of Management Journal*, 49(4), 723-740.

Terlaak, A., & King, A. A. (2007). Follow the small? Information-revealing adoption bandwagons when observers expect larger firms to benefit more from adoption. *Strategic Management Journal*, 28, 1167-1185.

Turner, S., R. Bettis, & R. Burton, R. 2002. Exploring depth versus breadth in knowledge management strategies. *Computational & Mathematical Organization Theory*, 8(1), 49-73.

Vermorel, J., M. Mohri. 2005. Multi-armed bandit algorithms and empirical evaluation. J. Gama, R. Camacho, P. Brazdil, A. Jorge, L. Torgo, eds. *Machine Learning: ECML 2005*, LNAI 3720. Springer- Verlag, Berlin, 437–448.

Weber, E., S. Shafir, A. R. Blais. 2004. Predicting risk sensitivity in humans and lower animals: Risk as variance or coefficient of variation. *Psychological Review* **111**(2) 430–445.

Zenger, T. R., & S. G. Lazzarini. 2004. Compensating for innovation: Do small firms offer high-powered incentives that lure talent and motivate effort? *Managerial and Decision Economics*, 25(67), 329-345.

FIGURES

Figure 1: Paradox of Size — Trade Off between Average and Extreme Performance

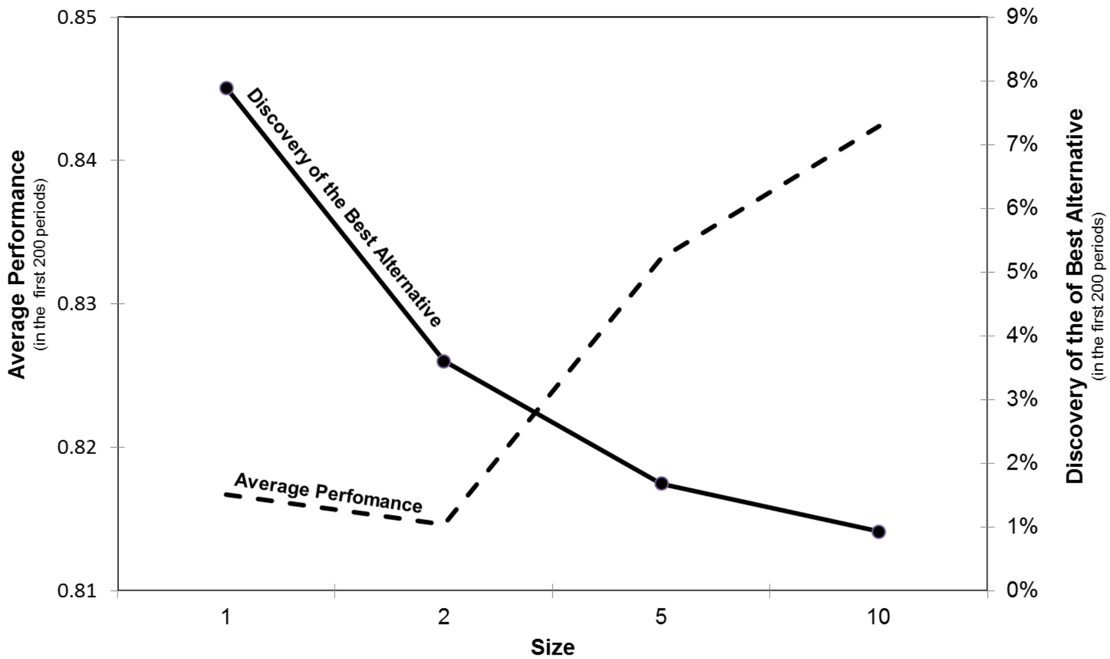
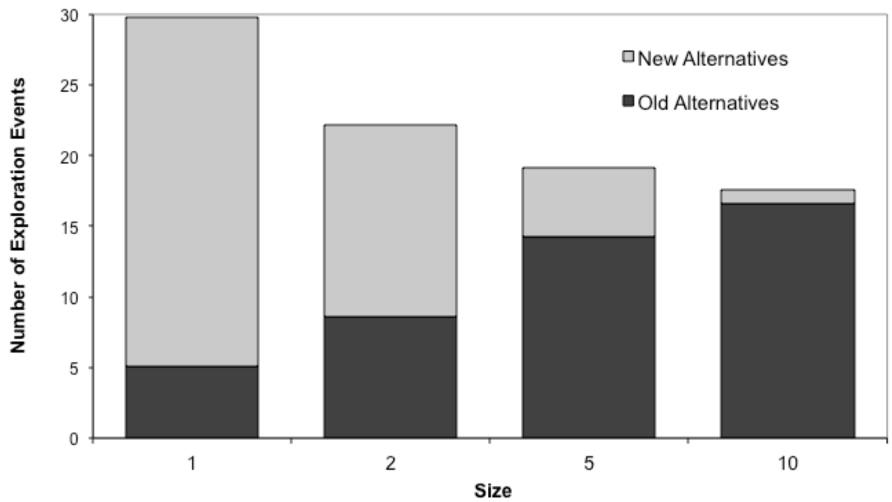


Figure 2

Panel A: Quantity of Exploration — Period-to-Period Alternative Changes



Panel B: Quantity of Exploration — Number of Distinct Alternatives Explored

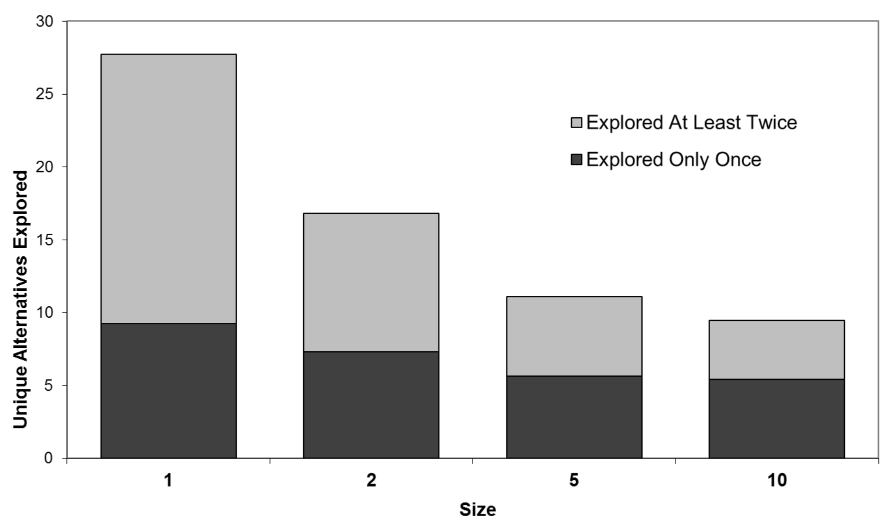


Figure 3: Quality of Exploration — Rank of Alternatives Explored

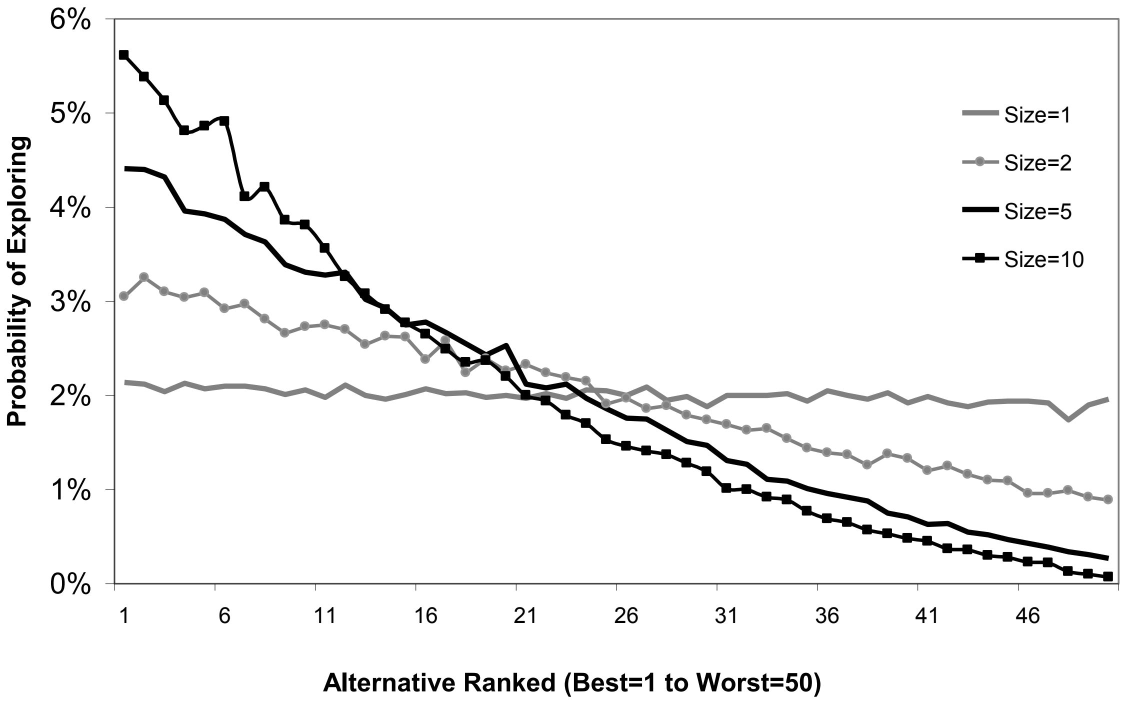


Figure 4: Trade-Off between the Quality and Quantity of Exploration

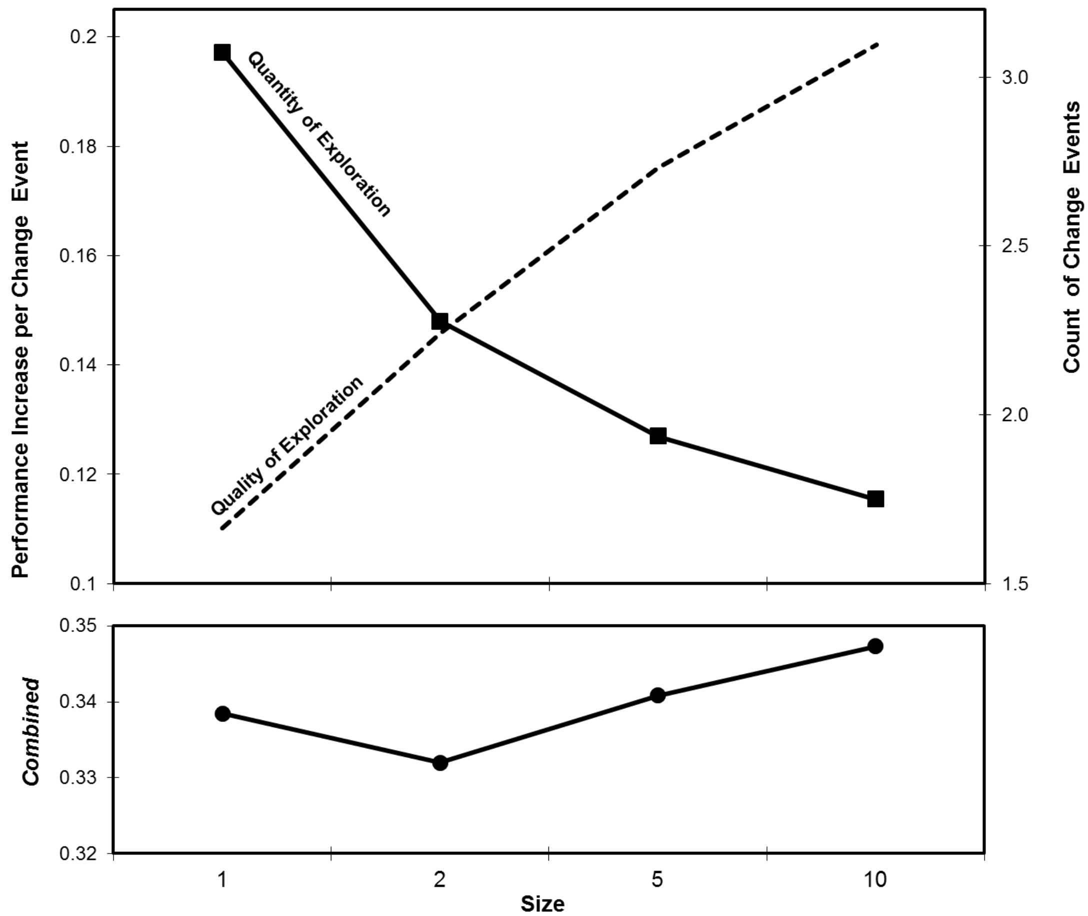


Figure 5: Rejecting the Best Alternative — Probability and Trigger (First 20 Periods)

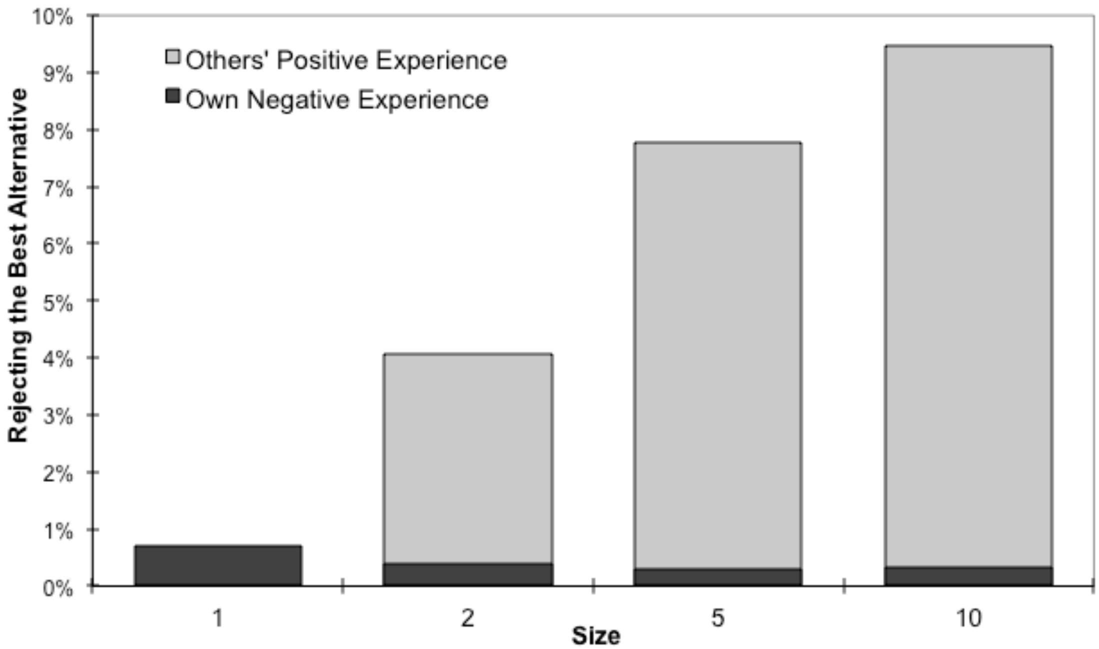


Figure 6: Illustration of Belief Updating in Organizations of Ten Units

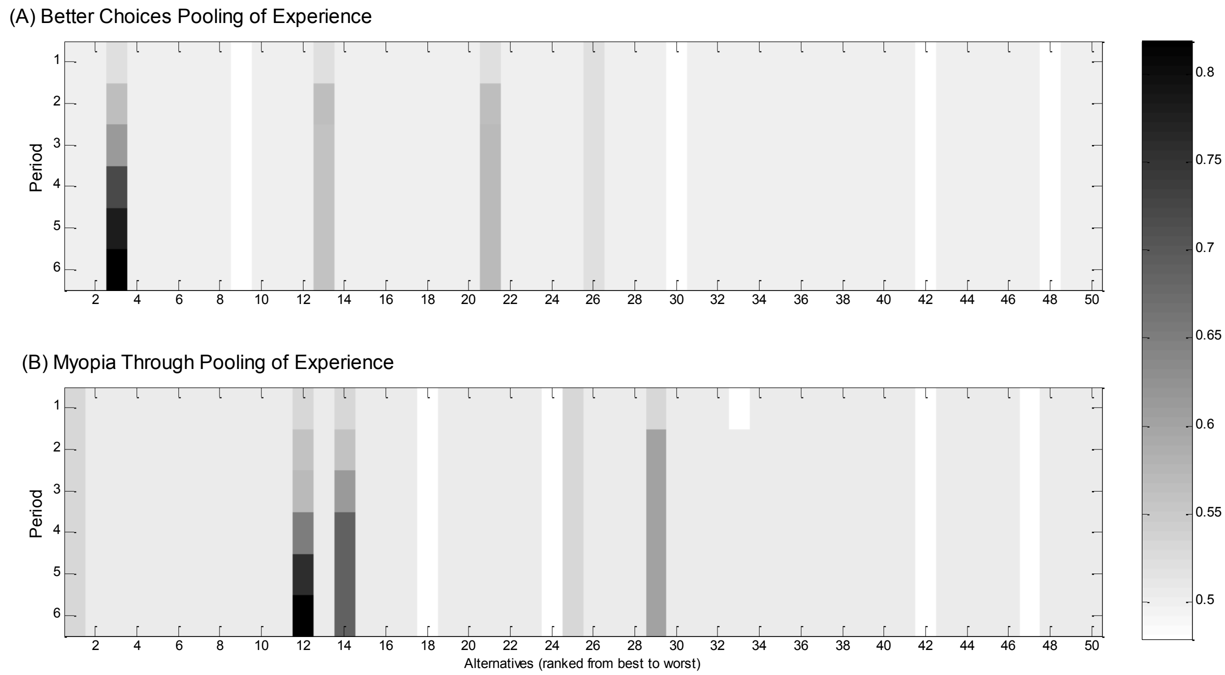


Figure 7: Altering the Weighting of Experience (ω=0.50 and ω=1.00)