Exploration, exploitation, and variability: Competition for primacy revisited

Juha Uotila

Warwick Business School

The University of Warwick

Coventry CV4 7AL United Kingdom

juha.uotila@wbs.ac.uk

(forthcoming in Strategic Organization)

Exploration, exploitation, and variability: Competition for primacy revisited

Abstract

This study examines how exploration and exploitation contribute to variability in organizational performance and how this variability influences competitions for primacy, i.e. contexts in which the ability to generate exceptionally high levels of performance is a key success factor. The results of simulation analyses conducted using the NK model, the mutual learning model, and the multiarmed bandit model all show that while exploration introduces more internal variability in organizations, exploitation tends to increase the variability in performance between different organizations. Variability and uniqueness in relation to one's competitors, brought about by a single-minded focus on exploitation, is found to be advantageous in a race to finish first. In contrast, in contexts where avoiding especially low performance is the key consideration, exploration is found to be relatively more beneficial. These results suggest a need for a more nuanced understanding of the relationships between exploration, exploitation, and organizational risk taking.

Keywords: Exploration/Exploitation, Performance, Variability, Competition, Simulation

Introduction

Balancing exploration and exploitation is a key challenge for organizations (Lavie et al., 2010; March, 1991; Gupta et al., 2006). However, while the importance of the two processes is widely accepted, understanding of how competitive conditions influence the relative benefits of the two adaptation types is still limited (Lavie et al., 2010). Furthermore, while exploration and exploitation differ in their associated risk profiles (March, 1991; Levinthal and March, 1993), the specific risk and variability of outcomes associated with the two modes of adaptation have received little explicit research attention.

This study argues that the variability of outcomes from exploration and exploitation and the competitive implications of the two types of adaptation are interlinked and addresses two important questions related to the organizational effects of exploration and exploitation. First, how do exploration and exploitation influence the variability of organizational performance, and second, how does this variability influence competitive outcomes in competitions for primacy? This study investigates these questions using the three canonical simulation modeling approaches to the study of exploration and exploitation: the NK model (Kauffman, 1993), the mutual learning model (March, 1991), and the multiarmed bandit model (Robbins, 1952).

This study finds that although exploration provides more variation in terms of its immediate performance effects within the organization, exploitation is generally associated with an increase in performance variation between different organizations. Because of this increased variability between organizations, an overly exploitative strategy is found to be beneficial in a competition for primacy, especially in the short term when the negative effects of lower long-term performance caused by such a strategy do not yet dominate the positive effects of higher performance variability. This study thus contributes to the organizational adaptation literature by demonstrating the value of the organizational uniqueness stemming from exploitation in settings where finishing near the top is important. The implications of these findings for theory and practice are discussed.

Exploration, exploitation, and variability

Exploration and exploitation are considered the two key forms of organizational adaptation (March, 1991; Gupta et al., 2006; Lavie et al., 2010). Exploitation allows the organization to refine and utilize its existing knowledge, competences and opportunities, whereas exploration allows the organization to find completely new knowledge, competences and opportunities. The importance of properly balancing these two processes is well-established

in the organizational adaptation literature (e.g., Lavie et al., 2010; Gupta et al., 2006; Uotila et al., 2009; He and Wong, 2004).

In his seminal article, March (1991) argued that returns from exploration and exploitation differ in their variability and timing and, because of this, organizations need to consider their risk preferences when allocating resources between the two types of activities. In his analysis of competition for primacy, March (1991: 81-85) further showed that a firm's likelihood of competitive success is a function of not only the mean but also the variance of its expected performance. Based on the assumption that exploration tends to be the riskier of the two alternatives, March argued that exploration is especially important in competitions for primacy: when finishing first among a large number of organizations is what matters, increasing the variability of performance can be more important than increasing the mean performance. The subsequent literature has taken the riskiness of exploration, and thus its relative importance over exploitation in competitions for primacy, largely as granted (e.g., Voss et al., 2008; Levinthal and March, 1993).

However, there are two reasons to expect that the role of exploration and exploitation in competitions for primacy may be less straightforward. First, the inherent riskiness and

performance variability associated with individual exploration actions do not necessarily aggregate to higher performance variability when considered at the level of the organization that may undertake a large number of such exploratory moves, when compared to a firm following a more exploitative strategy. For example, Levinthal (1997) noted that exploitative adaptation can take different organizations to completely different solutions, indicating that not only exploration but also exploitation can be a source of outcome variability at the firm level. However, there has been little comparative analysis of the organization-level performance variability stemming from exploration and exploitation.

The second factor to be considered is the time dimension. Even if exploration was associated with higher performance variability at the firm level, the returns to exploitation are generally considered to be more immediate and thus higher in the short run (March, 1991). Because success in competitions for primacy is a positive function of both the expected mean performance and the variability of this performance (March, 1991; Denrell and March, 2001), the total effects of exploration and exploitation in such competitions are unclear, especially in the short term where increasing exploration relative to exploitation might increase performance variability at the expense of the mean performance. Thus,

whether and how the duration of the competition for primacy influences the optimal exploration–exploitation balance is an open question.

This study addresses these gaps by taking a formal look at the effects of exploration and exploitation on the variability of organizational outcomes and consequently on competition for primacy, over different time horizons, using a formal simulation modeling approach of organizational adaptation. In his analysis of competition for primacy, March (1991) assumed each firm to exhibit constant performance over the course of the competition, with no explicit modeling of learning or adaptation – either exploratory or exploitative in nature – taking place. The present study's approach of allowing the organizations to adapt through exploration and exploitation serves to relax this assumption, thus helping to unpack the variation of organizational performance and providing a more nuanced understanding of exploration, exploitation, and performance variability over time.

Method

Computational modeling is a useful tool for examining undertheorized process phenomena with a simple underlying theoretical basis (Davis et al., 2007), especially when they involve

processes such as organizational search and learning that are difficult to observe empirically (Burton and Obel, 2011). For the modeling of exploration and exploitation, there are three canonical simulation models in the literature: the NK model, the mutual learning model, and the multiarmed bandit model. Typically, simulation studies of exploration and exploitation have selected one of these frameworks and adapted it to suit the particular needs of their research questions (e.g., Posen and Levinthal, 2012; Rivkin and Siggelkow, 2007; Miller et al., 2006). However, because the analysis of competition for relative position operates on the basis of performance distributions resulting from exploratory and exploitative strategies – something that all three models yield – the competitive analysis of the present paper can be applied to any underlying exploration–exploitation framework. Consequently, to investigate the extent to which the observed patterns are generalizable across different formalizations of exploration and exploitation, this study employs all three frameworks to analyze the implications of exploration and exploitation on performance variability and competition.

The NK model

The NK model (Kauffman, 1993) is arguably the most widely used formalization of exploration and exploitation in management science (e.g., Rivkin and Siggelkow, 2007;

Siggelkow and Levinthal, 2003; Levinthal, 1997). In the NK model, organizations adapt in a rugged performance landscape that consists of N choice dimensions, each of which can take the value of 0 or 1 and represent the elementary choices that the organizations can make regarding for example their strategy, structure, technologies, and so on. Organizations are represented by the vector of binary choices they make in this landscape, for example 0101001000 in a landscape with N = 10. Each landscape position, i.e. vector of choices, is associated with a corresponding performance payoff, and organizations attempt to find the landscape position with the best payoff possible.

In the NK model, the payoff contribution of each of the N choice dimensions depends on K other, randomly assigned choice dimensions, reflecting the fact that the effect of an organizational choice typically also depends on a number of other choices that the organization makes. Specifically, for each choice dimension n of landscape position i, where $n = 1, \ldots, N$, a payoff contribution P_{in} is randomly drawn from the uniform distribution U[0, 1] for each combination of binary choices c in both the focal dimension n and the K other dimensions on which that choice depends, i.e. $P_{in} = f_{in}(c_{in}, c_{in1}, \ldots, c_{inK})$, where $n1, \ldots, nK$ are the K interacting choice dimensions and f_{in} is a function yielding a different performance contribution, randomly initialized from U[0, 1], for each possible

combination of interacting binary choice inputs. The overall performance payoff of landscape position *i* is the average of the payoff contributions of its N choice dimensions, i.e. $P_i = (P_{i1} + P_{i2} + \ldots + P_{iN})/\mathrm{N}$.

Organizations, each starting from a randomly assigned landscape position, attempt to improve their positions through either local or distant search. In local search, the organization samples landscape positions that differ from its current position in only one choice dimension. For example, the organization in position 0101001000 could sample the landscape position 0111001000 through local search, experimenting with changing its choice in the third choice dimension. Because the organization builds its adaptation efforts on incrementally refining its current solution to the competitive problem posed by the environment rather than looking for completely new alternatives, this form of search is considered a form of exploitative adaptation. Conversely, in distant search the organization explores landscape positions that are further away from its current position, sampling a random vector of choices anywhere in the landscape. Such exploratory search can help the organization to find completely different solutions to the competitive problem. If the organization finds, through its local or distant search efforts, a landscape position that

offers better performance than its current one, it adopts this new position and continues its search from there in the next time period.

The mutual learning model

In the mutual learning model (e.g., Miller et al., 2006; March, 1991), the organization and its members are modeled as learning from each other over time. Consistent with March's (1991) implementation, the firms are modeled as operating in an external reality with m dimensions, each of which has a value of 1 or –1 with equal probability. In each firm, there are n individuals, each with their own set of beliefs about reality, and an organizational code that reflects the received truth as established within each firm. The beliefs on each dimension, held by the n individuals and the organizational code, can take a value of 1, 0, or –1. At the start of the simulation, all the m dimensions of the organizational code are set as 0, reflecting no belief about the environment, and the m dimensions of the n individuals are each randomly set as either 1, 0, or –1, with equal probability.

The mutual learning process involves two kinds of learning. First, each individual in a firm learns from that firm's organizational code so that, if the code has a non-zero value that differs from the belief of the individual, the individual adopts the value held by the code

with probability p_1. Thus, p_1 reflects the socialization rate within the firm. Second, the organizational code also learns from those individuals whose beliefs correspond to the reality in more dimensions than the code does. Specifically, if the code differs in a particular dimension from the majority view among the individuals with superior knowledge, the code adopts this majority view with the probability $1 - (1 - p_2)^{(k-j)}$, where k is the number of superior-knowledge individuals that have a belief different from the code and j the number of superior-knowledge individuals that have the same belief as the code. The parameter p_2 thus reflects the rate at which the organizational code learns from the organizational members.

In the baseline model of mutual learning, the balance between exploration and exploitation can be adjusted by changing the socialization rate p_1. When the socialization rate is high, the organizational members quickly learn the beliefs reflected by the organizational code, and the received knowledge spreads quickly throughout the organization allowing for rapid exploitation of this knowledge. Conversely, when the socialization rate is low, divergent beliefs persist for longer in the organization, resulting in a longer period of exploration, with more learning opportunities for the organizational code and typically higher equilibrium knowledge (March, 1991).

March (1991) also discusses an open-system extension to the mutual learning model that adds to the model both environmental turbulence and personnel turnover. This model introduces two new parameters: turnover rate p_3 and the rate of environmental turbulence p_4. In each time period, each individual in a firm has a probability p_3 of leaving and being replaced by another individual with a randomly initialized set of beliefs. Further, for each dimension of the external reality, there is a probability p_4 that its value shifts from 1 to –1 or vice versa. In the open-system mutual learning model, the turnover rate p_3 is used to denote a firm's learning strategy: a higher rate of turnover implies a more exploratory strategy as the firm is constantly hiring new personnel to gain new beliefs that it hopes more accurately reflect the changing reality.

The multiarmed bandit model

The third canonical model, the multiarmed bandit model (e.g., Posen and Levinthal, 2012; Robbins, 1952; Stieglitz et al., 2016), conceptualizes the environment as a slot machine with N arms, each of which carries a different probability p_n, of winning a fixed payoff. These probabilities are unknown to the organizations. In the baseline model, p_n is set as x_n^{10}, with x_n taken as a random draw from U[0, 1] and n = 1, … , N. This setup generates a slot machine in which a very small number of arms carry a high probability of winning,

whereas for most arms, the winning probability is less than 0.1, making the need to find a high-performing arm a vital consideration.

Exploitation is modeled as the organization pulling the arm that it considers, based on its prior experience, to have the highest winning probability, i.e. the arm for which the organization's historical payoff to pull ratio is the highest, whereas exploration is modeled as the organization pulling a random arm in order to increase its information about the payoffs. After each pull, the organization updates its beliefs based on whether it receives a payoff or not. If there are several arms with the same perceived probability of winning, such as in the beginning of the simulation, the organization chooses one based on a randomly initialized order of preference.

The following analyses model the competitive efforts of a number of different organizations and examine how the extent to which an organization uses an exploitative versus an exploratory adaptation strategy affects its performance, the variability of this performance, and the competitive implications of this variability, using each of the three frameworks described above. Specifically, for each model, the adaptation efforts of 30 organizations are simulated. Ten of these organizations are Exploiters, following an

exploitative adaptation strategy as defined by each model, ten organizations are Explorers and follow an exploratory strategy, and ten organizations are Hybrids, following a strategy that is a mix between the Exploiter and Explorer strategies.

Results

Variability of returns from exploratory and exploitative search

The first analyses employed the NK framework that allows for an unambiguous measurement of not only overall firm performance but also of the immediate performance implications of individual exploratory and exploitative actions in terms of the improvements in landscape position yielded by such actions. In the simulations presented below, the landscape consisted of N = 24 choice dimensions, with each choice dimension being dependent on K = 3 other choice dimensions.[1]

To examine the risk profiles of exploration and exploitation, it is instructive to first look at the immediate probability of success of the two types of search among the 10 Exploiter firms following a local search strategy and sampling a randomly assigned local landscape position each turn, 10 Explorer firms following a distant search strategy and sampling each

turn a random landscape position over the entire landscape, and 10 Hybrid firms using each turn either an Exploiter strategy or an Explorer strategy, each with the probability 0.5.[2] Thus, there were on average 15 firms attempting both types of search in each turn. All the results below are the averages over 10,000 simulation runs on different landscapes.

Insert Figure 1 about here

Figure 1 shows the success rate per turn of both exploitative search and exploratory search, defined as the probability of improvement in performance over the previous time period, as well as how great the improvements from the two search types were when successful. Each data point was calculated as the average after the previous data point; for example, the data point "100" denotes the success rates and average performance improvements of exploration and exploitation after data point "50", i.e. for turns 51–100. As seen in Figures 1a and 1b, the success rate of exploitation was initially much higher than the success rate of exploration. Further, as Figure 1c shows, in all time frames a successful exploration attempt led, on average, to a greater performance improvement than a successful exploitation attempt. These results are in line with the notion that exploration tends to be more risky and

uncertain than exploitation, with exploitation offering a more certain and immediate payoff whereas the payoff from exploration is less certain but potentially greater (March, 1991). However, while the success rate of exploitation was initially high, the potential for exploitation was also exhausted faster than the potential for exploration, and the success rate of exploitation eventually fell below that of exploration.[3]

The benefits from atomistic exploitative and exploratory search attempts only tell us about the immediate rewards of these two modes of adaptation. However, what is strategically more interesting is the quality of the landscape positions that these adaptation modes allow the firm to reach and thus the ultimate performance payoff that the organization is able to gain. The following experiments therefore examine the average firm-level performance and the average standard deviation of performance among the 10 Exploiter firms, the 10 Explorer firms, and the 10 Hybrid firms, respectively, at different points in time.

Insert Figure 2 about here

Figure 2a shows the average performance of the three organization types and depicts dynamics that are in line with those found by Rivkin and Siggelkow (2007): Exploiters, limited to improving one choice at a time, suffered from the poorest initial performance but overtook the Hybrids and Explorers in the medium term because of their more consistent rate of improvement, before falling back to the last place again as they became stuck on local peaks.[4] Figure 2b shows the average standard deviation of this performance as a function of time[5] and offers an interesting departure from the literature that associates increased exploration with increased risk and variability of outcomes (e.g., Lavie et al., 2010; March, 1991): in all time frames, the standard deviation of performance is seen to be highest for Exploiters and lowest for Explorers, with the standard deviation of Hybrid performance somewhere between the two. Thus exploitation, rather than exploration, seems to be associated with increased variability in performance between different organizations.

Competition for primacy revisited

As March (1991) noted, when a large number of organizations compete for a single prize, such that only the best-performing organization receives any payoff, it is typically more important to increase the variance in the organization's performance than it is to increase the mean. March (1991: 83-85) further suggested that because of this, in such situations

exploration becomes more important than exploitation, a view that has been subsequently echoed by other scholars, such as Levinthal and March (1993: 107), who noted that "Exploration is, on average, unfruitful, but it is the only way to finish first". However, as we have seen above, the standard deviation of performance seems to be higher for organizations following an exploitation strategy than for those following an exploration strategy. What are the implications of this for competition for primacy?

To examine this question, the following analysis reports the results of a number of tournaments among the 30 organizations – 10 Exploiters, 10 Explorers and 10 Hybrids – set to last for a different number of turns. For each tournament, the winner was recorded, i.e. the organization with the highest performance of the 30 at the end of the tournament. This represents a winner-take-all situation in which there is only one prize to be gained by the firm with the best performance at a given time. Such winner-take-all situations represent especially intensive competitive conditions, such as those that may arise in network industries where increasing returns can generate lock-in to the leading firm's technology, forcing the losers out of the market (Arthur, 1989; McIntyre and Subramaniam, 2009). To explore which strategy yields the highest probability of winning such tournaments, a total of 160,000 tournaments were conducted, ranging from 1 to 10,000 turns in length, with

10,000 tournaments of each length, and the number of wins for different organization types for each tournament length was recorded. If more than one organization had the same performance at the end of the tournament, one of these was randomly assigned the winner with equal probability. To explore the effects of exploration and exploitation on the ability to avoid particularly bad competitive outcomes, the type of organization – Exploiter, Explorer or Hybrid – that placed last, i.e. had the worst performance of the 30, was also recorded.

Figure 2c shows the probability of Exploiters, Explorers, and Hybrids winning a competition for primacy as a function of the length of the tournament in the NK model. As seen in Figure 2c, Explorers rarely won such tournaments due to their slow rate of performance improvement. In tournaments lasting for a short to medium time, a pure Exploiter strategy was clearly the best bet when competing for primacy, with about 70 percent of wins in the 50 turn tournaments going to Exploiters. This was not only due to the faster short-term performance improvement for Exploiters, but was in large part driven by the increase in performance variability. For example, in 200 turn tournaments the average performance of Hybrids was significantly (i.e., more than two standard deviations) higher than the average performance of Exploiters, but Exploiters still had a greater chance of

finishing first due to the increased variance in their performance. Only in the long term did Hybrid firms become superior when their mean performance became sufficiently higher than that of Exploiters such that also the winner tended to be a Hybrid. Furthermore, as can be seen by comparing Figures 1 and 2c, Hybrids only outperformed Exploiters in competitions for primacy when the tournament was sufficiently long such that the potential for exploitative improvement was almost completely exhausted by the end of the tournament. All in all, the results in Figure 2c show that the increased variability in performance from a single-mindedly exploitative strategy, shown in Figure 2b, also translates to an increased success rate in short to medium length competitions for primacy.[6]

A more detailed look into the dynamics leading to the higher performance variability of Exploiters, depicted in Figure 2b, reveals that there were two processes in play. Initially, the performance improvements of Exploiters were potentially rapid but constrained by their local landscape neighborhood, which led to instances of both exceptionally rapid and exceptionally slow accumulation of performance improvements in the short term and the highest performance variability among the three organization types. This initial variance appeared to be driven by the locally constrained nature of exploitative search rather than landscape ruggedness; robustness tests showed that even on perfectly smooth landscapes (K

= 0), Exploiters had the highest standard deviation of performance in the short term. However, in the analysis depicted in Figure 2, a different dynamic started to set in at around time period 100: an increasing number of Exploiters, having exhausted their potential for incremental improvement, were stuck on local peaks with no further potential for improvement. The performance differences between these local peaks led to the constant equilibrium variance of Exploiter performance, whereas Hybrids and Explorers continued their journey towards the global peak and the standard deviation of their performance approached zero. It is noteworthy that this equilibrium variance of Exploiters resides in the left-hand tail of the performance distribution, and because the probability of finishing first depends on the right-hand tail of the performance distribution (March, 1991: 83), does not contribute positively to competitions for primacy.

As shown in Figure 2d, Explorers had the highest probability of finishing last in the medium term, which can be attributed to their low average performance due to the slow improvement rate. However, what is interesting in Figure 2d is that Exploiters had a higher probability than Hybrids of finishing last, regardless of the length of the tournament. Thus while the average performance of Exploiters and Hybrids was quite similar, as shown in Figure 2a, a pure Exploiter strategy tended to produce the best as well as the worst

performers among these two organization types. Thus, while the risks associated with exploration were found to be internal to the organization, reflected in the less consistent but potentially higher returns than those from exploitation, the risks associated with exploitation appear to be more external, as the path that the organization strives to exploit may be a suboptimal one in relation to the paths chosen by its competitors.

In the above baseline analysis, all firms started from different randomly assigned positions in the landscape. While this is a standard assumption in the NK literature and arguably more realistic than having all organizations start at the same position, it gives Exploiters a degree of initial variance similar to the other firm types. To examine the extent to which the results are driven by this initial variance, Figure 3 shows a replication of the above analysis with all 30 firms starting from the same landscape position in each simulation run. As can be seen in Figure 3, Exploiters, all starting with the same set of choices and only able to change one choice at a time, exhibited the lowest initial performance variability. However, this standard deviation quickly ramped up to become the highest when each Exploiter commenced its local search in a different direction – often towards different peaks, as suggested by the equilibrium standard deviation (0.025) that is only slightly smaller than that of the baseline model with random starting positions (0.030) – after which the model

dynamics were relatively similar to the baseline model. The standard deviation of Exploiter performance exceeded the standard deviation of Hybrid performance for about periods 15–30, and during this time Exploiters had both a higher chance of finishing first and a higher chance of finishing last than Hybrids. Given that the analysis presented in Figure 3 is likely to understate the variability of Exploiters more than that of Hybrids or Explorers, it supports the notion that the process of exploitative, local search contributes to the Exploiters' higher variability, rather than this variability only resulting from the random initial positioning in the baseline model.

Insert Figure 3 about here

Exploration as knowledge diversity: the mutual learning model

The NK model conceptualized the exploration–exploitation dilemma as that between distant and local search. However, the dilemma has also been considered to take many other forms, a fundamental one being the tradeoff between knowledge diversity and rapid knowledge utilization (March, 1991; Miller and Martignoni, 2016). The mutual learning model is used to examine whether these results hold in the latter context. In the simulations

below, Exploiter firms were defined as having a rapid socialization rate at $p_1 = 1$, Explorers as having a slow socialization rate at $p_1 = 0.1$, and Hybrids as having the socialization rate set at the intermediate value of $p_1 = 0.5$. The learning rate by the organizational code was held constant at $p_2 = 0.5$, the number of knowledge dimensions was $m = 24$ and the number of individuals was $n = 100$.[7]

Insert Figure 4 about here

Figure 4 shows the results of tournaments of varying length with 10 Exploiters, 10 Explorers, and 10 Hybrids, as defined above, in the mutual learning model. An exploitative learning strategy can be seen to have led to the highest performance, defined as the fraction of correct beliefs among the organization's members, in the short term, as well as the highest variability in performance.[8] The rapid socialization of Exploiters led to the early organizational beliefs – whether true or false – to be rapidly entrenched throughout the organization, frequently leading to both the best and the worst outcomes in the short term. In contrast, Explorers, with their slow socialization, did not exhibit such a degree of idiosyncrasy in their beliefs but rather slowly refined their knowledge bases to more closely

correspond with the reality, leading to a lower standard deviation and a higher equilibrium performance. The optimal learning strategy can be seen to depend on the length of the tournament, with increasingly long tournaments calling for an increasingly exploration-oriented strategy. Further, in some cases the variability-increasing effects of faster socialization dominated the effects of average performance; for example, in 20-period tournaments, the average performance of Explorers (0.739) was much higher than that of the faster-socializing Hybrids (0.677), but because of the larger variability in performance, nearly two-thirds (63.5%) of 20-period tournament winners were Hybrids.

Figure 5 reports the results from the open-system mutual learning model. Again, the simulations modeled the competition between 10 Exploiters, 10 Explorers, and 10 Hybrids, with the turnover rates p_3 set to zero for Exploiters, one for Explorers, and 0.5 for Hybrids. The rate of turbulence p_4 was held constant at $p_4 = 0.02$ and the mutual learning rates p_1 and p_2 both held constant at 0.5. Again, a purely exploitative strategy can be seen to exhibit the best short-term performance as well as the highest performance variability across all time periods. Because Exploiters had no turnover, they had no way to revise their established beliefs in the face of a changing environment, and by period 100 their average performance had regressed to 0.5, i.e. no better than chance, consistent with the results of March (1991:

80). In contrast, Explorers only had one period to learn from new recruits before they were again replaced by new individuals, and consequently the highest long-term performance in these models was exhibited by Hybrids that offered the best balance between exploration and exploitation among the three firm types. However, interestingly, because of the high standard deviation of performance among Exploiters, their probability of winning a competition for primacy remained clearly non-zero at about 17%, unlike that of the – on average, higher-performing – Explorers; because each Exploiter maintained its idiosyncratic set of beliefs, occasionally the belief set of one of the Exploiters became relevant again when the environment changed back and forth due to turbulence.

Insert Figure 5 about here

Exploration as experimentation: the multiarmed bandit model

The multiarmed bandit model provides yet another conceptualization of the exploration–exploitation dilemma, this time as the trade-off between experimenting to gain more knowledge of the available options versus selecting the option that the current knowledge suggests is the best. Figure 6 reports the adaptation efforts of 30 firms to a 100-armed

bandit: 10 Exploiters that always pulled the arm for which they had the initial preference, 10 Explorers that always pulled a randomly chosen arm, and 10 Hybrids that, with equal probability, pulled either a random arm or the arm that they considered to have the highest payoff probability. Each organization's performance was defined as the fraction of pulls that yielded a payoff, and the organization with the highest number of winning pulls was designated the winner. If more than one organization had the same number of winning pulls, one was again randomly assigned the winner.

As can be expected, in terms of average performance the Hybrid strategy was clearly the best of the three, being the only strategy where the organization actually exploited the information it gained from exploration to increase its expected payoff, whereas the long-term performance of both a pure Exploiter strategy and a pure Explorer strategy was, on average, no better than random. The multiarmed bandit analysis further confirms the result from the NK and mutual learning models that a purely exploitative strategy tends to lead to the largest standard deviations between the performances of the different organizations, as shown in Figure 6b.

Insert Figure 6 about here

The multiarmed bandit results regarding competition for primacy, shown in Figure 6c, also confirm the benefits of single-minded exploitation in tournaments lasting for a short period of time. The best strategy of winning a rapid race for primacy seems to be a purely exploitative one, picking one arm and sticking with it rather than wasting time exploring the options. The flipside of this is that, as seen in Figure 6d, such stubborn exploitation was also the recipe for finishing last if you happened to start with a low-payoff arm. The increased variance from a purely exploitative strategy again led to both the very best and the very worst outcomes. However, in the long term, the higher average performance from the Hybrid strategy was again found to dominate the effects of the higher variability of Exploiters, and in tournaments lasting for a sufficiently long period of time, the winner tended to be a Hybrid.[9]

Discussion

While some of the specific results are unique to each simulation model, three general patterns are relatively consistent across the different models and parameter values. First,

exploitation is generally more beneficial in the short term – both in terms of average performance effects and in terms of winning a competition for primacy – whereas longer time horizons in most cases call for an increased amount of exploration. This finding is consistent with existing literature on the temporal effects of exploration and exploitation (e.g., March, 1991; Levinthal and March, 1993). However, as a novel contribution, increasing exploitation relative to exploration was found to generally lead to an increase in performance variability in most of the simulated contexts and over most time periods. This second finding partly challenges the established view of exploration as the main source of variability in performance (e.g., He and Wong, 2004; March, 1991), suggesting that exploitative processes may be more central to increasing variability between organizations. Third, the increase in performance variability stemming from a highly exploitative strategy was found to be beneficial in competitions for primacy, and in many contexts exploitative firms were found to have a higher probability of finishing first – or last – than their average performance would suggest.

While the results confirm the notion that exploration is, in its immediate impact, more risky than exploitation – as reflected by the lower initial success rate but higher potential payoff of exploratory search attempts in the NK model – they also provide a finding that seems

counterintuitive in light of the established literature: the variance in performance between different organizations, both in the short term and in the long term, is typically larger for organizations following an exploitative strategy than it is for organizations following an exploratory strategy. The key to understanding this apparent contradiction is in the different nature of the two types of variability. The variability brought about by exploration is internal to the organization (March, 1991; McGrath, 2001): exploration gives the organization a variety of options and new knowledge, and therefore facilitates the organization to radically change its landscape position or knowledge base. In contrast, the variability brought about by exploitation occurs between different organizations: with each organization following its own idiosyncratic developmental path, these organizations tend to differ from each other more over time than a group of organizations all exploring over the full range of available knowledge or options. In a competition for primacy, the type of variability that provides competitive advantage is that between different organizations – the type promoted by single-minded exploitation.

It is noteworthy that variability is only beneficial in competitions for primacy when it occurs in the right-hand tail of the performance distribution, potentially providing exceptionally high levels of performance (March, 1991: 83). For example, the equilibrium

variability of Exploiters in NK landscapes is caused by some of these firms being stuck on local peaks, thus occurring in the left-hand side of the distribution and not providing long-term benefits for Exploiters in competitions for primacy. In contrast, the short-term variability stemming from the Exploiters' adaptation strategy that is highly sensitive to the early direction that these firms' adaptation efforts take also influences the right-hand side of the distribution and can be beneficial in competitions for primacy.

Furthermore, Exploiters' success in competitions for primacy stemming from their high performance variability is contingent on their mean performance not falling too far below that of the more long-term oriented Hybrids and Explorers, and consequently in most cases Exploiters' success is limited to competitions lasting for a short to medium period of time. However, there are reasons to believe that the short-term effects may be more relevant in the real world than the long-term equilibrium conditions. First, many competitive situations require relatively rapid results (D'Aveni, 1994; Eisenhardt, 1989). Second, business environments can change, either radically and thus in effect resetting the search process, or incrementally through constant turbulence, which may also increase the benefits of exploitation relative to exploration (Posen and Levinthal, 2012; Marino et al., 2015).[10]

Theoretical and practical implications

This study answers recent calls to examine how different competitive conditions influence the benefits of exploration and exploration (Lavie et al., 2010). The results corroborate the argument that "In competition to achieve relatively high positions, variability has a positive effect. In competition to avoid relatively low positions, variability has a negative effect" (March, 1991: 83). However, it is not the internal type of variability brought about by individual exploration actions but variability between organizations, often resulting from a generally exploitative strategy, that has a positive effect in a competition for finishing at the top and a negative effect in a race to avoid finishing last. Especially when faced with a choice between exploratory and exploitative courses of action that have similar expected mean performance implications, it is useful to understand how the two types of strategies influence performance variability.

Uniqueness from competitors is key to winning competitions for primacy and can be achieved through a focused exploitation effort. Thus, the results of this study suggest that fast-paced competitions for primacy may not require broad exploration but rather a combination of luck and a single-minded focus on the firm's distinctive competences (Barney, 1991; Prahalad and Hamel, 1990). When finishing first among a number of

competitors is important, an optimal strategy may be to go with the solution that the firm is uniquely positioned to exploit and begin refining it without wasting resources for further exploration, hoping that such a solution will turn out to be the winner. Even though, on average, introducing more exploration could lead to finding solutions with better potential, in the face of intense competition such exploration may be a luxury that the organization cannot afford. The more intense the competition is and the more important it is to be among the very best, the more important is the kind of variability that focused exploitation can bring.[11] These results therefore support the finding by Jansen, Van Den Bosch and Volberda (2006) that competitive intensity may increase the relative benefits of exploitation.

In contrast, when it is vital to avoid finishing last, exploration becomes important. The results of the present study highlight that, in addition to avoiding competence traps and obsolescence, exploration has a further potential benefit: the internal variability generated by exploration increases the organization's portfolio of options, allowing it to hedge its bets in complex and uncertain competitive contexts when particularly low-performing choices need to be avoided. In such contexts, it is advisable to engage in at least some exploration –

traditionally considered the “risky” alternative (March, 1991; Voss et al., 2008) – to reduce the risk of ending up with a particularly bad solution.

These results can also contribute to understanding why large, successful firms tend to emphasize exploitation more than would be optimal (Uotila et al., 2009; Levinthal and March, 1993). If overexploitation leads to the very best and the very worst outcomes, it may be not only the result but also, in part, the cause of the observed success of such firms. Due to selection bias, overly exploitative firms may be overrepresented among the largest firms: whereas less fortunate exploiters get weeded out, the luckiest ones rise to the top of their industries. However, when their luck runs out, the main legacy of their success may be an unbalanced search strategy going into the future.

Limitations and future research

The results of this study suggest that the competitive effects of exploration and exploitation warrant further research attention. While the models used in the present paper could be extended to cover different competitive conditions, the abstract nature of formal simulation modeling also brings about the need to evaluate the applicability of these results empirically. Even though obtaining similar results from three different types of simulation

models increases the confidence that these findings are theoretically relevant, simulation analysis cannot answer whether and in what situations the variability and uniqueness brought about by exploitation, or the variance-reducing hedging brought about by exploration, are strategically significant. Empirical work using qualitative or quantitative data on real-world organizations is needed to assess the practical applicability of these results. Examining the kinds of dynamics discussed in the present study in the context of managerial risk-taking (e.g., March and Shapira, 1987; Kahneman and Lovallo, 1993) could also provide fruitful avenues for further research.

The present paper has assumed that the organizations operate as singular units. However, similar logic could be extended to different levels of analysis. For example, Levinthal and Marino (2015) conceptualized the organization as a set of evolving practices, each modeled as a separate NK landscape, on which intra-organizational selection pressures operate. The results of the present paper suggest that, in such cases, variability and consequently the efficiency of selection processes may increase when the practices are allowed to evolve exploitatively through rapid, local adaptation. Future research could increasingly examine such dynamics of exploration, exploitation, and performance variability across different

levels of analysis, for example through a more nuanced modeling of internal organizational structures (e.g., Levinthal and Marino, 2015; Davis et al., 2009).

There is a particularly salient limitation of the present study that should be kept in mind when interpreting the results, stemming from how novelty is modeled in the existing computational models of exploration and exploration: all competing firms are assumed to conduct their exploration and exploitation efforts in a predetermined and limited search space. All three canonical exploration–exploitation models – the NK model, the multiarmed bandit model, and the mutual learning model – share the same limitation in that exploration only entails gaining better knowledge or a better position with a pre-existing set of alternatives. What the established approaches do not readily facilitate is the modeling of particularly extreme and novel results from exploration, such as game-changing technological advances, that March (1991) may have had in mind when arguing for the higher variability of the returns to exploration. While modeling broader search spaces – larger landscapes, knowledge structures with a larger number of beliefs, or bandits with a larger number of arms – would ensure that there always remain undiscovered areas to facilitate exploration, these novel areas are, in essence, qualitatively similar to the ones already discovered, limiting the upside potential from exploration.[12] When the returns to

exploration are bounded by the same performance function as the returns to exploitation, the existing models are likely to underestimate the effects of exploration on performance variability in dynamic competitive contexts.[13] As Felin et al. (2014: 278) argued, "The focus on factors such as bounded rationality, calculation, and search on phase spaces and landscapes is inappropriate for understanding the emergence of entrepreneurial and economic novelty."

This suggests a fundamental limitation in the existing computational models of exploration and exploitation: the treatment of novelty. Future theoretical work on organizational learning and adaptation should therefore strive to develop models of exploration as a more proactive process of novelty generation, rather than only a process of evaluating and choosing from a given set of alternatives. Reconsidering the modeling of novelty in our theoretical treatment of exploration and exploitation could be instrumental in advancing our understanding of the effects of exploration and exploitation on variability and competition.

Conclusion

The existing exploration–exploitation literature has equated exploration with variability in organizational performance. Using formal simulation of exploration and exploitation, this study has shown that exploitation, rather than exploration, is the key process that generates variability in performance between different organizations. This type of variability brought about by exploitation is vital in competitions for primacy, where it is important for organizations to generate their own, unique advantage in an effort to stand out from their competitors. Conversely, exploration can be valuable as a risk-reducing strategy when it is vital to avoid particularly low-performance outcomes.

Notes

[1] Robustness tests using different values for landscape size N (from 8 to 50), complexity K (from 1 to 12), and the number of organizations (from 1 to 100 of each type) produced results qualitatively similar to the baseline model.

[2] While it is easy to find search strategies that outperform all three used in the present paper, these three simple strategies are used as representative examples of the relationships between exploration, exploitation, performance variability, and competition for primacy. The search for optimal search strategies in different competitive contexts is outside the scope of the present paper.

[3] The average time it took for a purely exploitative search to reach a local peak with no further improvement potential was 154 time periods, with a standard deviation of 89.

[4] For a more detailed explanation of these patterns, see Rivkin and Siggelkow (2007: 1081).

[5] Because the variable of interest is the variability of performance within a particular population, standard deviations were calculated among each set of firms adapting on the same landscape and then averaged over the different landscapes in order to remove the effect of different landscapes on the standard deviations of outcomes. Standard deviations calculated over all trials, i.e. across all the different landscapes, were larger but exhibited a qualitatively similar pattern.

[6] An alternative model in which poorly performing firms could be selected out and replaced by another firm of a similar type, as in Levinthal (1997: 938-939), yielded similar short-term results with Exploiters showing the highest performance variability and success in competitions for primacy. In the long run, as Exploiters stuck on local peaks got selected out, firms of all types ended up on the global peak and performance differences disappeared. Another alternative model in which off-line evaluation of landscape positions was not possible but a firm evaluating an alternative (local or distant) landscape position first adopted the new position for one time period before observing its performance implications on-line, returning to its previous landscape position if its performance decreased due to the change, generated the highest standard deviation of performance for

Explorers. However, this higher standard deviation occurred solely in the left-hand tail of the performance distribution as the Explorers and, to a lesser extent, Hybrids constantly implemented low-performing solutions in their search. Thus, while this negative variability points to a further risk of exploration that the baseline model did not address, it did not have an influence on competitions for primacy, and the patterns depicted in Figure 2c were qualitatively similar in the model with on-line evaluation of alternatives: Exploiters had the highest chance of winning competitions for primacy in the short term while Hybrids prevailed in the long run.

[7] Robustness tests using different amounts of knowledge dimensions (from 2 to 100) and individuals per firm (from 2 to 1000) produced qualitatively similar patterns.

[8] This result is robust to different parameter values (p_2, m and n) in the short term; with some combinations of parameter values, the initially highest performance variability of Exploiters was exceeded by that of Hybrids or Explorers in the long term.

[9] Robustness tests showed that these results are relatively insensitive to the number of arms. A smaller number of firms, a payoff distribution with a smaller number of high-performing arms, or a more optimal search strategy for Hybrids (e.g., only pulling a random arm 10% of the time) were found to benefit Hybrids in competitions for primacy and allow them to overtake Exploiters earlier. However, in all experiments, Exploiters had the highest chance of winning competitions for primacy lasting for a very short period of time.

[10] Experiments with the NK model supported this result that Posen and Levinthal (2012) showed in the context of the multiarmed bandit model, and in the NK model with turbulence, Exploiters had the highest probability of winning competitions for primacy in the long term as well.

[11] Experiments with varying numbers of firms supported this notion; in general, the higher the number of competing firms, the more beneficial exploitation was in competitions for primacy.

[12] Robustness tests conducted with search spaces of varying sizes supported this contention.

[13] A similar issue pertains to the modeling of environmental turbulence: as discussed in the context of Figure 5, occasionally the environment reverts to a similar state as it had been in an earlier time period, and a previously obsolete knowledge base becomes optimal again. In many real-world contexts, this is clearly unrealistic, and computational models are likely to overstate the benefits of an exploitative strategy in turbulent environments.

References

Arthur WB (1989) Competing technologies, increasing returns, and lock-in by historical events. *Economic Journal* 99(394): 116-131.

Barney J (1991) Firm resources and sustained competitive advantage. *Journal of Management* 17(1): 99-120.

Burton RM and Obel B (2011) Computational modeling for what-is, what-might-be, and what-should-be studies—and triangulation. *Organization Science* 22(5): 1195-1202.

D'Aveni RA (1994) *Hypercompetition: Managing the dynamics of strategic maneuvering*. New York: Free Press.

Davis JP, Eisenhardt KM and Bingham CB (2007) Developing theory through simulation methods. *Academy of Management Review* 32(2): 480-499.

Davis JP, Eisenhardt KM and Bingham CB (2009) Optimal structure, market dynamism, and the strategy of simple rules. *Administrative Science Quarterly* 54(3): 413-452.

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

Eisenhardt KM (1989) Making fast strategic decisions in high-velocity environments. *Academy of Management Journal* 32(3): 543-576.

Felin T, Kauffman S, Koppl R, et al. (2014) Economic opportunity and evolution: Beyond landscapes and bounded rationality. *Strategic Entrepreneurship Journal* 8(4): 269-282.

Gupta AK, Smith KG and Shalley CE (2006) The interplay between exploration and exploitation. *Academy of Management Journal* 49(4): 693-706.

He Z-L and Wong P-K (2004) Exploration vs. Exploitation: An empirical test of the ambidexterity hypothesis. *Organization Science* 15(4): 481-494.

Jansen JJP, Van Den Bosch FAJ and Volberda HW (2006) Exploratory innovation, exploitative innovation, and performance: Effects of organizational antecedents and environmental moderators. *Management Science* 52(11): 1661-1674.

Kahneman D and Lovallo D (1993) Timid choices and bold forecasts: A cognitive perspective on risk taking. *Management Science* 39(1): 17-31.

Kauffman SA (1993) *The origins of order: Self-organization and selection in evolution*. New York: Oxford University Press.

Lavie D, Stettner U and Tushman ML (2010) Exploration and exploitation within and across organizations. *Academy of Management Annals* 4(1): 109-155.

Levinthal DA (1997) Adaptation on rugged landscapes. *Management Science* 43(7): 934-950.

Levinthal DA and March JG (1993) The myopia of learning. *Strategic Management Journal* 14(Winter Special Issue): 95-112.

Levinthal DA and Marino A (2015) Three facets of organizational adaptation: Selection, variety, and plasticity. *Organization Science* 26(3): 743-755.

March JG (1991) Exploration and exploitation in organizational learning. *Organization Science* 2(1): 71-87.

March JG and Shapira Z (1987) Managerial perspectives on risk and risk taking. *Management Science* 33(11): 1404-1418.

Marino A, Aversa P, Mesquita L, et al. (2015) Driving performance via exploration in changing environments: Evidence from formula one racing. *Organization Science* 26(4): 1079-1100.

McGrath RG (2001) Exploratory learning, innovative capacity and managerial oversight. *Academy of Management Journal* 44(1): 118-131.

McIntyre DP and Subramaniam M (2009) Strategy in network industries: A review and research agenda. *Journal of Management* 35(6): 1494-1517.

Miller KD and Martignoni D (2016) Organizational learning with forgetting: Reconsidering the exploration–exploitation tradeoff. *Strategic Organization* 14(1): 53-72.

Miller KD, Zhao M and Calantone RJ (2006) Adding interpersonal learning and tacit knowledge to march's exploration-exploitation model. *Academy of Management Journal* 49(4): 709-722.

Posen HE and Levinthal DA (2012) Chasing a moving target: Exploitation and exploration in dynamic environments. *Management Science* 58(3): 587-601.

Prahalad CK and Hamel G (1990) The core competence of the corporation. *Harvard Business Review* 68(3): 79-91.

Rivkin JW and Siggelkow N (2007) Patterned interactions in complex systems: Implications for exploration. *Management Science* 53(7): 1068-1085.

Robbins H (1952) Some aspects of the sequential design of experiments. *Bulletin of the American Mathematical Society* 58: 527-535.

Siggelkow N and Levinthal DA (2003) Temporarily divide to conquer: Centralized, decentralized, and reintegrated organizational approaches to exploration and adaptation. *Organization Science* 14(6): 650-669.

Stieglitz N, Knudsen T and Becker MC (2016) Adaptation and inertia in dynamic environments. *Strategic Management Journal* 37(9): 1854-1864.

Uotila J, Maula M, Keil T, et al. (2009) Exploration, exploitation, and financial performance: Analysis of s&p 500 corporations. *Strategic Management Journal* 30(2): 221-231.

Voss GB, Sirdeshmukh D and Voss ZG (2008) The effects of slack resources and environmental threat on product exploration and exploitation. *Academy of Management Journal* 51(1): 147-164.

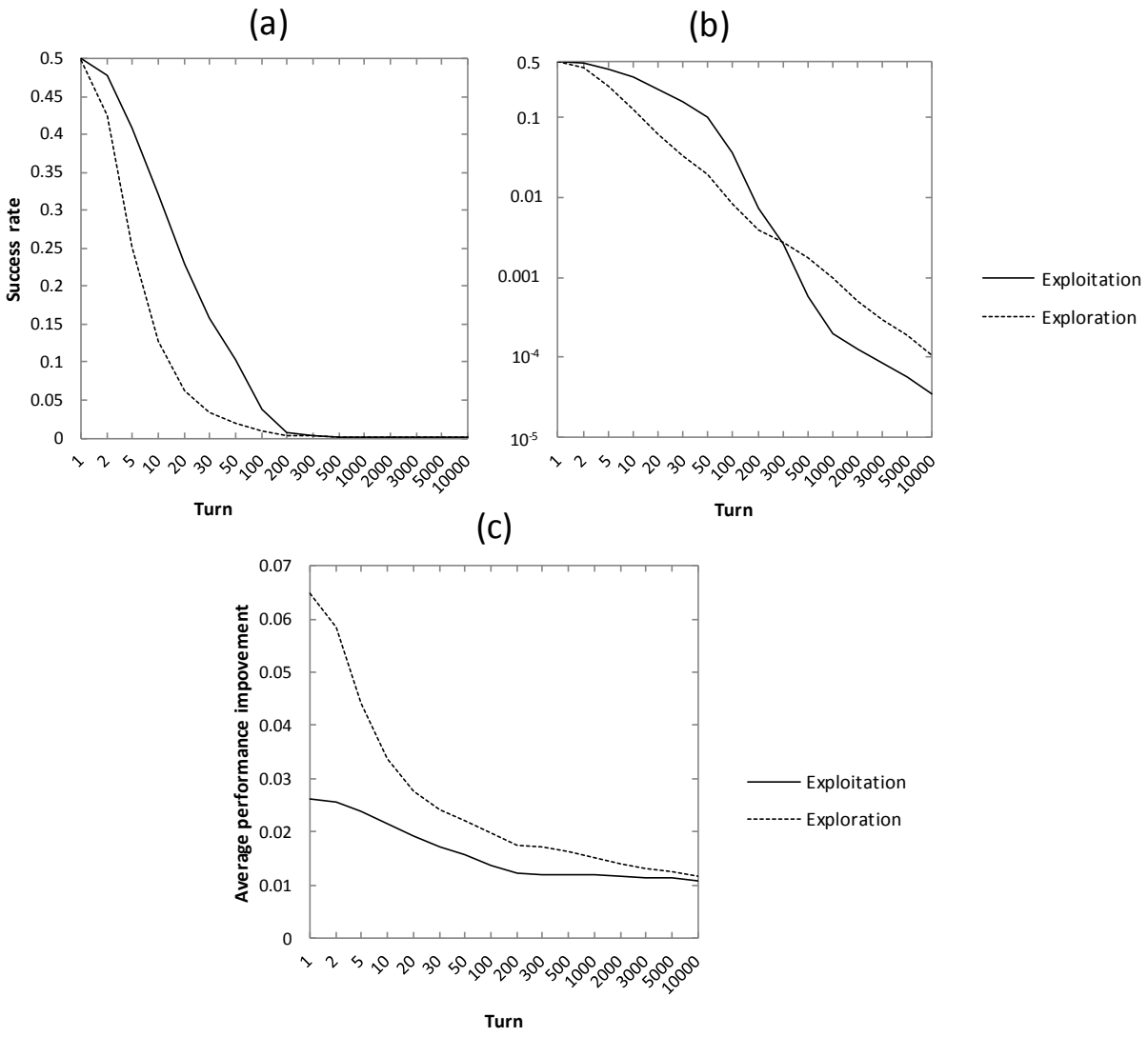

Figure 1. Success rates of exploitation and exploration per turn, in (a) linear and (b) logarithmic scale, and (c) the average performance improvement from a successful exploitation and exploration attempt.

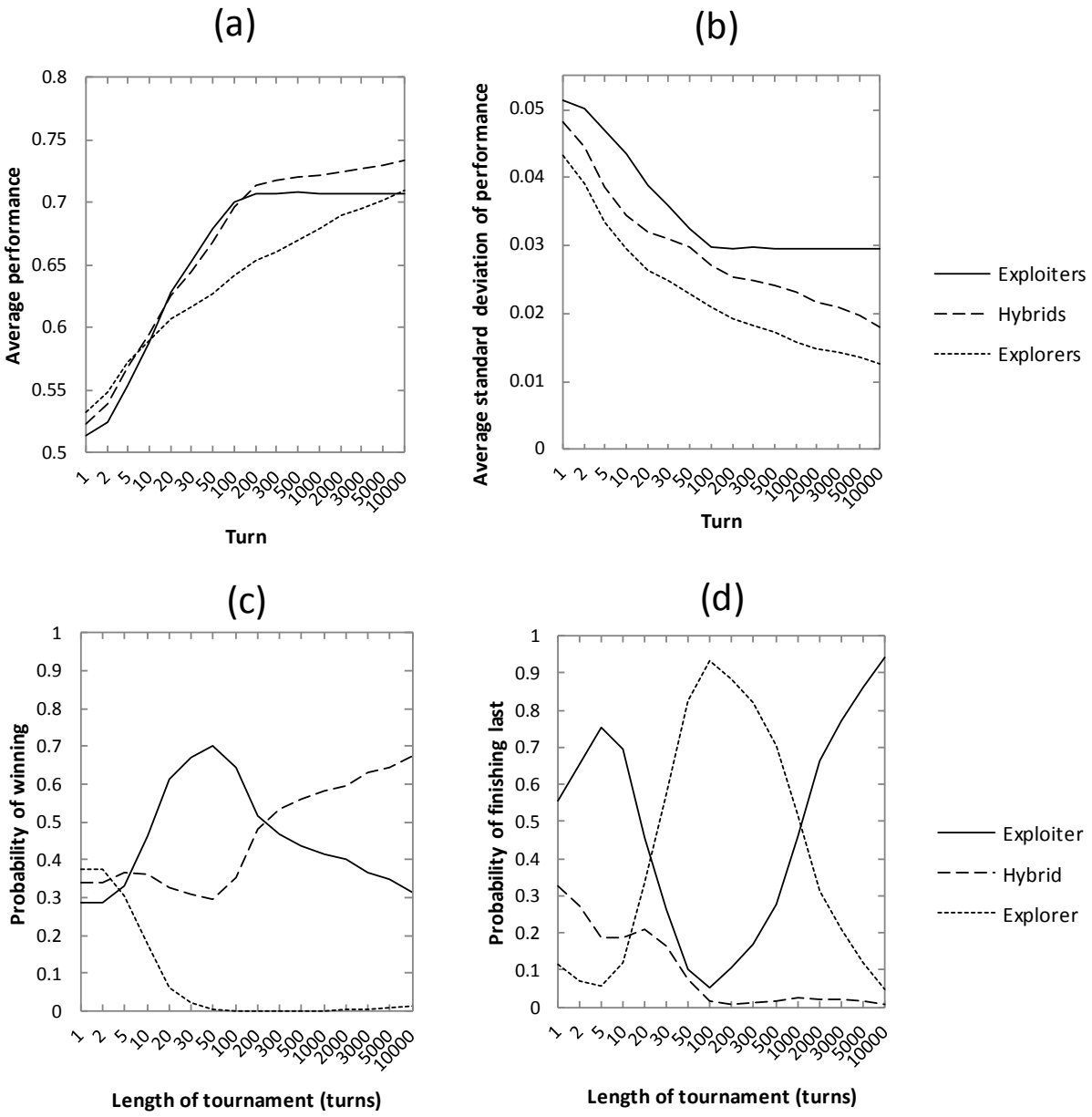

Figure 2. Average (a) performance, (b) standard deviation of performance, (c) probability of winning, and (d) probability of finishing last of different organization types in the NK model.

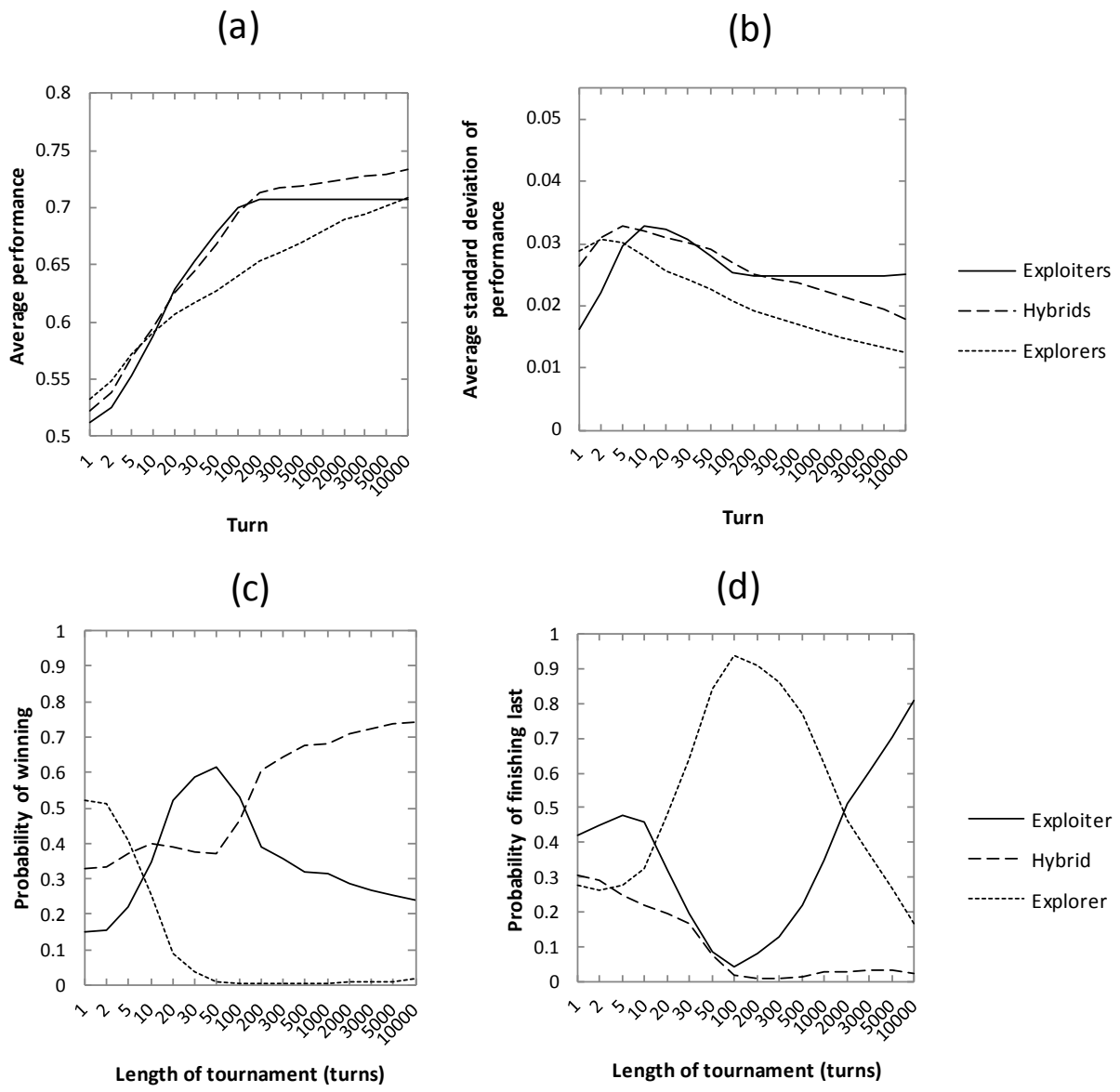

Figure 3. Average (a) performance, (b) standard deviation of performance, (c) probability of winning, and (d) probability of finishing last of different organization types in the NK model with all organizations starting at the same landscape position.

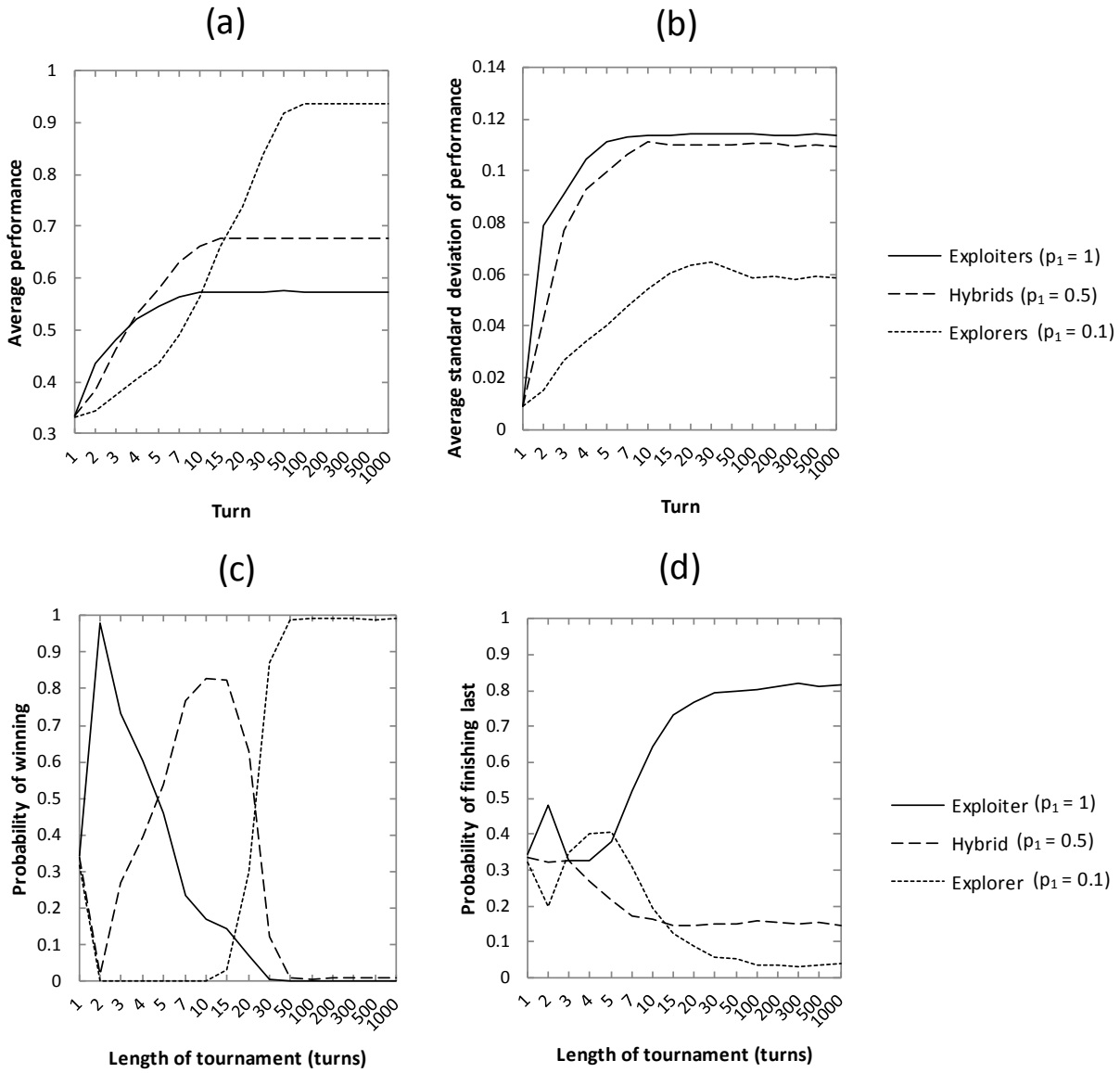

Figure 4. Average (a) performance, (b) standard deviation of performance, (c) probability of winning, and (d) probability of finishing last of different organization types in the baseline mutual learning model (p_2 = 0.5) with different socialization rates (p_1).

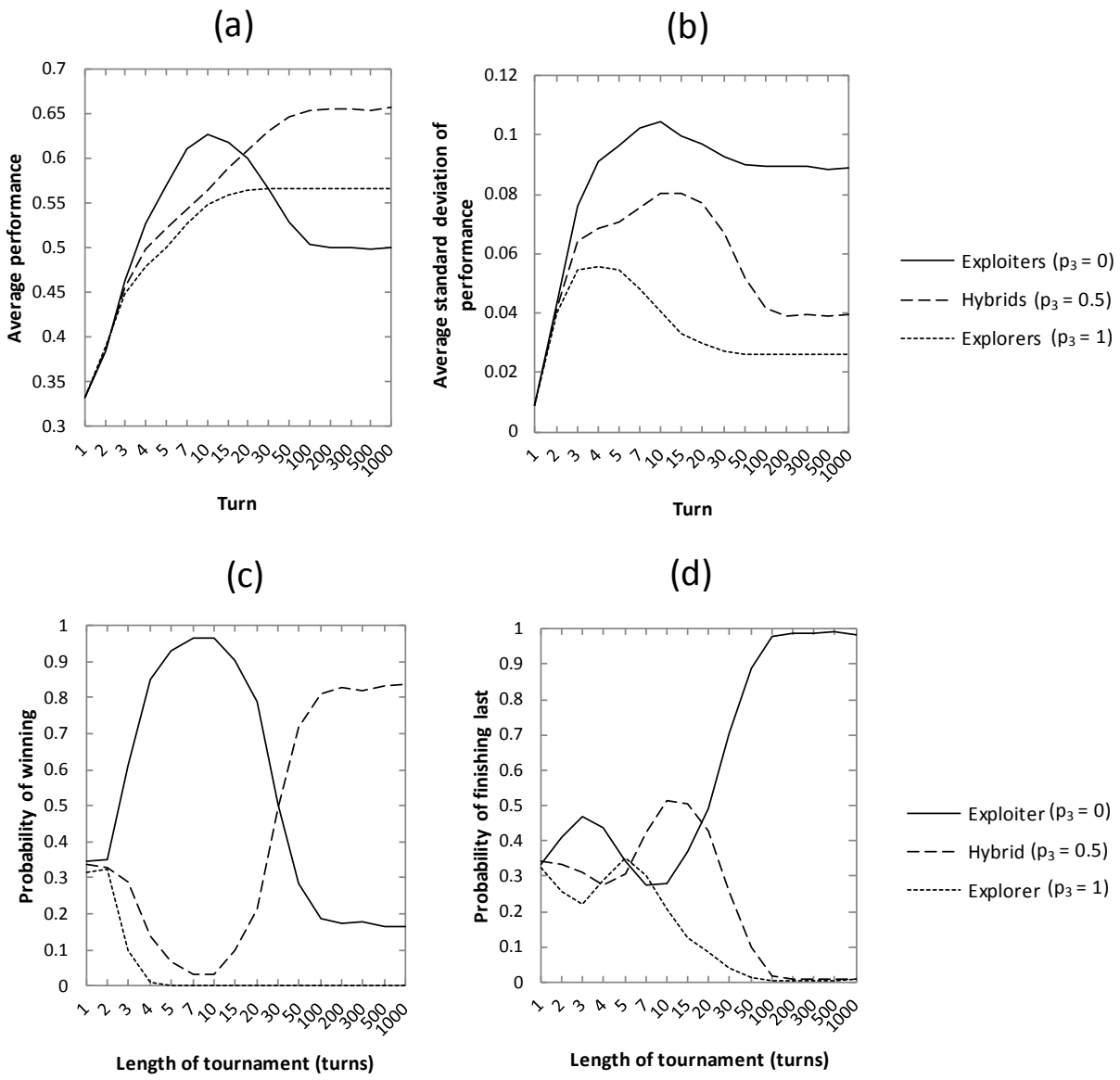

Figure 5. Average (a) performance, (b) standard deviation of performance, (c) probability of winning, and (d) probability of finishing last of different organization types in the open-system mutual learning model ($p_1 = 0.5$, $p_2 = 0.5$, $p_4 = 0.02$) with different turnover rates (p_3).

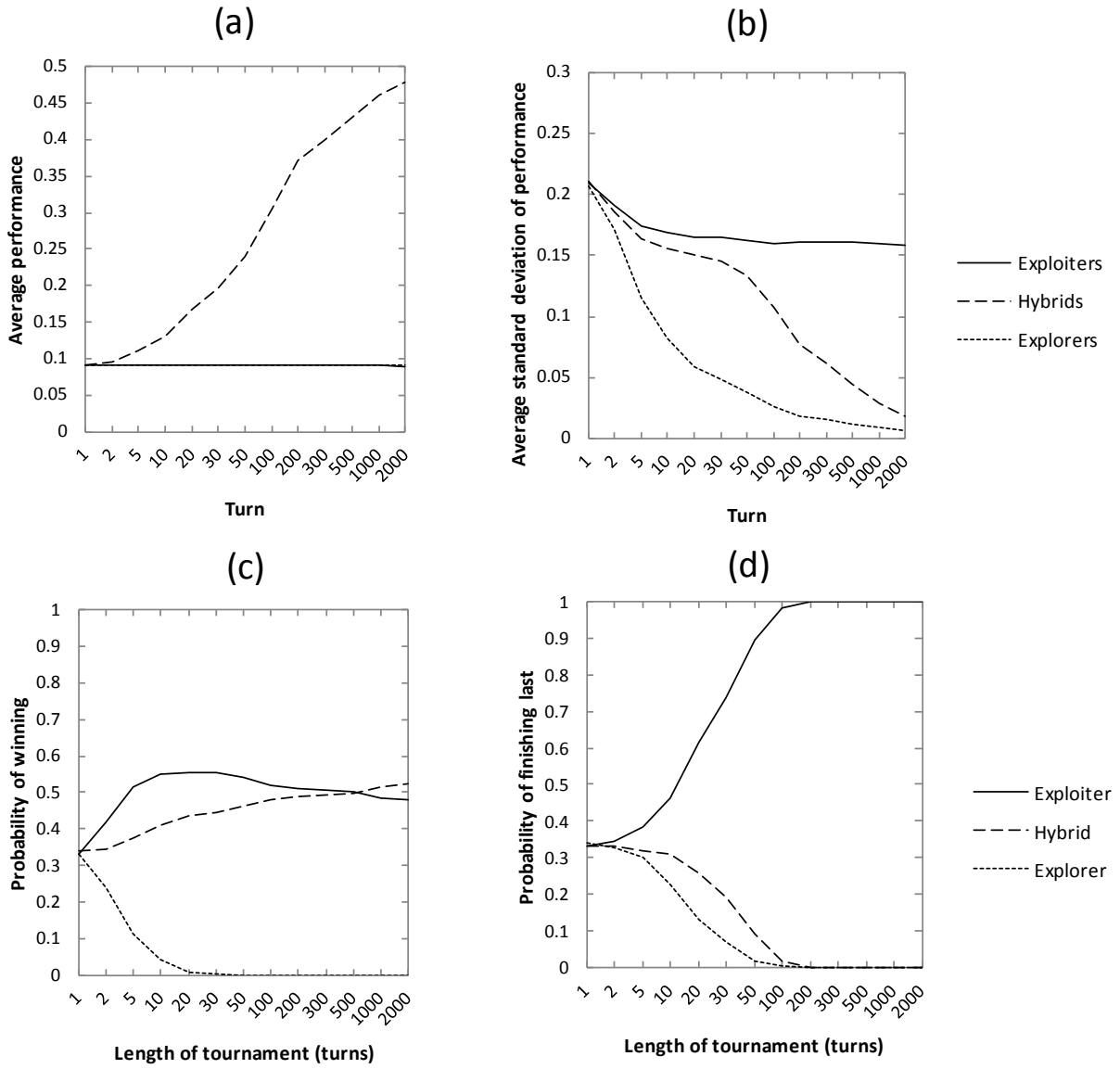

Figure 6. Average (a) performance, (b) standard deviation of performance, (c) probability of winning, and (d) probability of finishing last in a 100-armed bandit.