Paper to be presented at the DRUID 2012

on

June 19 to June 21

at

CBS, Copenhagen, Denmark,

REINFORCEMENT LEARNING IN STRATEGIC DECISION-MAKING

Thorsten Grohsjean
Imperial College London
Innovation & Entrepreneurship Group
t.grohsjean@imperial.ac.uk

Tobias Kretschmer
University of Munich (LMU)
Institute for Strategy, Technology, and Organisation
t.kretschmer@lmu.de

Nils Stieglitz
University of Southern Denmark
Department of Marketing & Management
nst@sod.dias.sdu.dk

Abstract

The paper examines how firms learn over time to allocate resources to alternative strategic actions in a changing business environment. We draw on psychological models of reinforcement learning to develop theory about how market feedback influence managerial beliefs about the attractiveness of strategic actions, which in turn shape choices about resource allocation. We extend this simple model of reinforcement learning with considerations from the behavioral theory of the firm to understand under what conditions managers rely more on their current beliefs in allocating resources. We test our hypotheses on an unbalanced panel of 58 video game publishers releasing games in 11 genres between 1997 and 2008. Controlling for several firm, industry and genre specific effects as well as the influence of vicarious learning on resource allocation, we find broad support for our hypotheses.

Jelcodes:M00,-

REINFORCEMENT LEARNING IN STRATEGIC DECISION-MAKING

It is probably impossible to overstate the importance of organizational learning for firm strategy and industrial dynamics. For decades, an extensive literature in economics and management has studied how learning improves operational routines and capabilities (Argote & Miron-Spektor; Nelson & Winter 1982; Winter 2000) and its impact on industrial dynamics (Dobrev & Kim 2006; Klepper 1996). More recently, the focus has shifted toward understanding how learning contributes to the emergence and evolution of dynamic capabilities (Teece et al. 1997; Maurizio Zollo & Winter 2002). Dynamic capabilities shape how firms learn to create, extend, and modify their resource base. Subsequent work has therefore deepened insights into learning how to write contracts with external suppliers (e.g. Vanneste & Puranam, 2008), developing an alliance capability (e.g. Kale & Singh, 2007), or learning to acquire and to integrate other firms (e.g. Zollo & Singh, 2004).

Yet, despite the prominent role attributed to learning, it appears that strategy research has lost sight of other salient goals in the learning process. A persistent challenge for managers is to understand which strategic actions lead to superior performance. The task is challenging because it is often uncertain how strategic actions map onto performance outcomes. Competitive interactions and environmental changes complicate the task considerably as they alter the actions that lead to superior performance. The emphasis here therefore is not on learning how to do things, but rather which things to do (Makadok 2003). For example, the dynamic capability approach argues that firms must match existing and new productive resources to uncertain business opportunities (Helfat et al. 2007; Teece 2007; Teece et al. 1997), but it is not clear how firms assess and learn about the relative attractiveness of alternative opportunities. In a similar spirit, Lippman and Rumelt (2003) characterize strategy as the allocation of resources without the help of market prices. They call for research into performance advantages to "concentrate directly on the process of decision on the best 'use' of resources" and claim that "at a theoretical level, there remain deep questions about proper administrative processes and the treatment of heterogeneous beliefs [in resource allocations]." From this perspective, the central problem for managers is how to allocate resources to alternative strategic actions whose performance consequences might be highly uncertain (Gavetti & Rivkin 2007; Noda & Bower 1996).

In this paper we study how experience and competitive feedback shape strategic decision-making and the allocation of resources to strategic actions over time. We connect strategic decision-making to ideas drawn from the reinforcement learning literature. The simple intuition of reinforcement learning is that a decision maker reinforces an action that led to success, while failure prompts the withdrawal from an action (Denrell & March, 2001; Lave & March, 1975; Levinthal & March, 1993). Research in psychology (e.g. Yechiam & Busemeyer, 2005) and neuroscience (Daw et al. 2006; Dayan & Niv 2008) has produced strong empirical support that human decision makers exhibit behavior in many resource allocation tasks that is consistent with reinforcement learning principles. The question that we pursue here is whether reinforcement learning provides a useful theoretical and empirical lens to study strategic decision-making over time in a competitive context.

We argue that managers form fallible beliefs about the relative attractiveness of alternative actions from prior feedback. If feedback is positive an action is believed to be more attractive and this leads to higher resource commitments to the action. In contrast, negative performance feedback leads to a downward revision of beliefs about the action. The action receives relatively fewer resources than before. We elaborate on this basic model by also considering the role of goal attainment at the organizational level on the relationship between beliefs and resource allocations. While failure to meet organizational goals at the organizational level signals the inadequacy of current beliefs and prompts the problemistic search for new knowledge to update beliefs and improve organizational performance (Cyert & March, 1963; Greve, 2003), success signals the value of current beliefs for reaching goals, resulting in a reduced tendency to explore. Further, the availability of slack resources promotes exploration and thereby weakens the impact of beliefs on strategic actions (Nohria & Gulati, 1996; Singh, 1986). Finally, we also examine the role of prior resource commitments that create switching costs between actions (Dierickx & Cool 1989; Ghemawat 1991). We hypothesize that substantial prior resource commitments make behavior more path-dependent and current beliefs more salient for strategic actions.

We test our hypotheses on an unbalanced panel of 58 video game publishers releasing games in 11 genres between 1997 and 2008. The fast-paced, turbulent video game industry provides a strong test for our theory, since reinforcement learning is often associated with stable environmental conditions. Controlling for several firm, industry and genre specific effects as well as the influence of vicarious learning on strategic decision-making (i.e. the tendency to imitate other firms' successes (Denrell 2003; Greve 1998), we find broad support for our hypotheses.

Our research contributes to prior work in a number of ways. First, our findings strongly support earlier ideas of strategy as an iterated process of resource allocations regulated by reinforcement learning principles (Dierickx & Cool 1989; Noda & Bower 1996). Second, we offer a simple, but powerful theory of firm strategy and industrial dynamics over time. Managerial beliefs drive resource allocations. Resource allocations affect the relative competitive positions of all firms in the industries. Competitive positions distribute gains and losses among competitors. Gains and losses have a differential effect on the revision of beliefs, thereby often creating and maintaining heterogeneity in managerial beliefs, strategic actions, and performance outcomes. Third, reinforcement learning may provide a useful platform to organize and integrate diverse streams of literatures and to stimulate further work.

THEORY AND HYPOTHESES

Strategy as iterated processes of resource allocation under uncertainty

Strategy is about the creation and sustainability of superior business performance. The task of the strategist then is to identify and allocate resources to strategic actions that lead to superior performance outcomes. This general conceptualization underpins different approaches to strategy. For example, the resource-based view of strategy perceives that the central challenge of strategy is directing investment flows into resource stocks with uncertain ex ante value (Dierickx & Cool 1989; Makadok 2001; Peteraf 1993; Winter 2000). A similar structural problem occurs in the diversification into new markets (Bhardwaj et al. 2006; Mosakowski 1997; Noda & Bower 1996) or the entry into new market segments (Dobrev et al. 2003; King & Tucci 2002). In all of these situations resources are allocated to strategic actions (such as resource development, entry into new markets or segments) with uncertain performance outcomes. To effectively allocate resources, managers need to form beliefs about what actions are attractive for the firm. Given that firms are often characterized by heterogeneous resource bundles (Lippman & Rumelt, 2003) the extent to which they can draw on public sources of information and on vicarious learning from others tends to be limited (Nadkarni & Barr 2008). Rather, they will draw on their own past experience to assess the future performance outcomes of alternative actions.

To illustrate, consider the problem of investing in different market niches. Prior research studying the entry into new market niches, pointed to the influence of prior operating experience

(Dobrev et al. 2003; Greve 2000; King & Tucci 2002), of dynamic capabilities (King & Tucci 2002), of imitative entry (Baum et al. 2000; Greve 1998), of strategic momentum (Amburgey & Miner 1992), and of competitive pressures (Baum & J. V. Singh 1994; Dobrev & Kim 2006). Yet, how firm resources are divided between niches has received virtually no attention. This is problematic, because treating niche entry as a binary event is misleading. The extent of a firm's resource commitment to a niche may differ substantially, and the resource allocations to a specific niche may change over time (Noda & Bower 1996; Adner & Levinthal 2004; Bhardwaj et al. 2006).

Noda and Bower (1996) drew attention to this problem in their case study of diversification moves by two US telecommunications companies in the 1980s. They found evidence that a company escalated its resource commitment to a new line of business when early results were positive. Negative results led to a de-escalation of resource commitment. This pattern is suggestive of a reinforcement learning process, in which a decision maker forms beliefs about the performance outcomes of actions (Levitt & March 1988; Denrell & March 2001). Positive feedback reinforces a belief about an action, and increases the likelihood of choosing the action again. Conversely, negative feedback weakens the belief about an action and thereby decreases the probability of choosing the action again. In the following, we draw on the reinforcement learning literature to understand how prior experience impacts managerial beliefs and resource allocations (Arthur 1991; Dayan & Niv 2008; Yechiam & Busemeyer 2005).

Reinforcement learning and the formation of beliefs

The traditional forum for problems of reinforcement learning is a task environment characterized by a set of possible actions $\{A_1, A_2, \ldots A_n\}$. The possible actions map onto unknown payoff outcomes $\{\Pi_1, \Pi_2, ... \Pi_N\}$ that are usually assumed to be stochastic. The set of possible actions and the associated performance outcomes may be stable or fluctuate over time (Daw et al. 2006). The problem for an economic agent is to identify and select an action A^* that maximizes payoffs over repeated trials (Sutton & Barto 1998)

Economic agents form beliefs about the payoff outcomes by sampling actions and observing the feedback (Yechiam & Busemeyer 2005). The feedback is then encoded into beliefs about the payoff properties of actions. A belief is defined in terms of two elements, namely the set of feasible actions $\{a_1, a_2, ... a_N\}$ and the estimated payoff outcomes $\{v_1, v_2, ... v_N\}$ for each action. The estimated payoff outcome for an action may diverge from the objective payoff consequence i.e. an economic agent may hold incorrect beliefs about the task environment. The idea of reinforcement learning is that agents form fallible beliefs based on feedback from their actions. Note, however, that the conceptualization of human beliefs in this framework is quite primitive and does not capture the richness of human cognition and beliefs (Gavetti & Rivkin 2007).

The most basic way to update belief estimates is to average over all prior payoff signals i.e. $v_N = (\pi_{N,1}, \pi_{N,2}, ... \pi_{N,k}) / k$, with k the number of payoff signals from action *n*. A higher payoff signal increases the estimate of the payoff, while a lower signal reduces the estimate. The reliability and accuracy of an estimate heavily depends on the number of payoff signals. A greater number signals a more accurate estimate of action-outcome linkages. Psychological research shows that human decision makers place more weight on recent feedback and discount more distant payoffs, even in stable task environments (Camerer et al. 2003; Erev & Barron 2005; Yechiam & Busemeyer 2005). That is, managerial decision-making is probably better represented by a recent weighted average rather than raw averaging over the payoff signals.

Updating is then assumed to be an exponentially weighted average such as (Sutton & Barto 1998):

$$(1) \qquad v_N^{t+1} = (1-\alpha)v_N^t + \alpha\pi_N^t \, .$$

The parameter α is the learning rate that regulates how salient prior payoff signals are for the formation of beliefs. A higher value of α discounts prior experience more heavily. In the case of α = 1 only the most recent payoff is considered in forming beliefs. In the limit case of α = 0 no learning and updating takes place and an initial belief is fixed forever.

The outlined model of reinforcement learning implies that managers in business firms often evaluate alternatives based on prior feedback. We propose that, ceteris paribus, they allocate more resources to actions in which they experienced success and less to actions in which they encountered failure. This implies the following hypothesis:

Hypothesis 1. The more positive the belief about a strategic action, the more resources will be allocated to this strategic action.

Belief-based resource allocation

Managerial beliefs about action-outcomes linkages become relevant for resource allocation in the choice process that considers how beliefs lead to the selection of actions. A simple choice rule is to (greedily) pick the action associated with the highest estimated payoff. A concern with "greedy" action selection is that beliefs about the task environment may be wrong, so the economic agent ends up with an inferior alternative. An economic agent may choose to select an apparently inferior alternative to learn more about its performance properties. This is the classic trade-off between the exploration of new knowledge and the exploitation of current knowledge about action-outcome linkages (Cohen, McClure, & Yu, 2007; Levinthal & March, 1993; March, 1991). Exploration is here defined as choosing an action currently believed to be inferior in the hope of increasing the accuracy and reliability of the estimate. Economic agents engage in exploration because they know that their beliefs are fallible and potentially wrong. (Cohen et al. 2007).

A “non-greedy” choice rule that admits the potential for exploration is the Softmax rule of probable choice. The idea is is that the choice of a non-greedy action is probabilistic i.e. there is a certain likelihood that a decision-maker engages in the exploration of apparently inferior actions. It has garnered some empirical support in laboratory experiments (Yechiam & Busemeyer 2005; Daw et al. 2006) and is often invoked in formal management models examining the tradeoff between exploration and exploitation in organizational learning (e.g. Denrell, Fang, & Levinthal, 2004; Fang & Levinthal, 2009; Posen & Levinthal, 2011). In Softmax choice, the probability of selecting an action A_N over alternative actions at time t based on current beliefs v is

$$(2) \qquad p_n^t = e^{v_n^t/\tau} / \sum_{b=1}^{n} e^{v_b^t/\tau} .$$

The basic relationship is that the more positive the current belief about an action, the higher is the probability of selecting that action. The critical parameter then is τ, which sets the agent’s tendency to explore. A higher tendency to explore softens the impact of current beliefs on action selection and thereby makes the choice of an inferior action more likely. The tendency to explore thereby acts as a moderator between managerial beliefs and resource allocations. In the following, we assume that beliefs are, at least to some extent, salient for behavior, so there is a strong correlation between beliefs and resource allocations. We consider contextual differences in the tendency to explore in the next section.

Moderators between managerial beliefs and resource allocations

The central parameter for allocation behavior in the outlined model of reinforcement learning is the tendency to explore (i.e. parameter τ in equation 2). Prior management research has developed detailed insights into the antecedents and processes of organizational exploration and exploitation. We draw on this literature to develop hypotheses on the moderators between beliefs and resource allocation. A positive moderator translates into a lower tendency to explore in equation (2) and a stronger correlation between beliefs and resource allocation. A negative moderator indicates a higher tendency to explore and softens the link between resource allocations.

Failure to attain organizational aspirations

A key insight from research in the tradition of the behavioral theory of the firm (Cyert & March 1963) is that organizations only search for new knowledge if they encounter a problem. Problemistic search is triggered by the failure to attain organizational aspirations ((e.g. Bromiley, 1991; Denrell, 2008; Greve, 2003; Shinkle, 2012 for a recent overview). Thus, firms operating below aspiration levels exhibit a higher tendency to explore, while firms operating above aspiration levels explore less. In the context of reinforcement learning, poor organizational performance is a strong indicator that the current beliefs of the organization are not well aligned to the reality of the task environment. Consequently, the response is to actively seek out new knowledge to update beliefs about action-outcomes linkages. In contrast, performance above aspirations signals the value of current beliefs for reaching goals and reduces the tendency to explore. This suggests that current beliefs become more consequential for firm behavior:

Hypothesis 2a. Organizations performing below aspiration level rely less on current beliefs in allocating resources.

Hypothesis 2b. Organizations performing above aspiration level rely more on current beliefs in allocating resources.

Availability of slack resources

Beyond problemistic search prompted by failure to meet aspirations, the behavioral theory of the firm points to slack resources as a second source of organizational exploration. Slack resources are defined as excess resources beyond those required for current operations. They buffer the organization from environmental pressures and thereby allow more room for experiment and exploration (Bourgeois III, 1981; Nohria & Gulati, 1996; Singh, 1986; Voss, Sirdeshmukh, & Voss, 2008). Organizations with lager stocks of slack resources are thus assumed to engage in more exploration. Hence, slack resources should weaken the tight relationship between current beliefs and future resource allocation, leading to more exploration of apparently inferior alternatives:

Hypothesis 3. Organizations with higher stocks of slack resources rely less on current beliefs in allocating resources.

Irreversible investments into action-specific assets and capabilities

In many managerial settings, irreversible investments in a strategic action create durable assets and capabilities that are specific to the particular action (Dierickx & Cool 1989, Ghemawat 1991; Winter 2000). For the firm, irreversible investments create switching costs, since the asset or capability erodes if no additional resources are allocated to keep it in operation (Dierickx & Cool 1989, Ghemwat 1991; Teece et al. 1997). An important element of irreversible investment is that they give rise to path-dependencies (Dixit 1989) and a rational lock-in to a specific action (Ghemawat 1991; Kraatz & Zajac 2001). This suggests that irreversible investments such as into brand reputation, specialized machinery, or vertical relations tend to make firms less inclined to explore new alternatives and focused more on exploiting current knowledge. We therefore propose that firms with higher stocks of specific assets and capabilities tend to have a lower tendency to explore and therefore a stronger link between current beliefs and the allocation of (fungible) resources.

Hypothesis 4. Organizations with higher stocks of specific resources rely more on current beliefs in allocating resources.

Reinforcement learning in a competitive context

For business organizations, the learning and resource allocation process described above unfolds in a competitive setting (Herriott, Levinthal, & March, 1985; Levitt & March, 1988). The choice of a resource allocation influences the competitive position of a given firm and the ensuing competitive process distributes gains and losses among the firms in an industry (Jacobides & Winter 2005). The differential gains and losses provide feedback to update beliefs about action-outcome links. The updating of beliefs then may induce changes in resource allocation.

Note that an organization obtains feedback on two distinct levels in the present framework. First, feedback is valuable to form beliefs about the relative attractiveness of alternative actions. Second, an organization also receives higher-level feedback on the attainment of organizational aspirations and goals. The latter is a complementary indicator of the quality of current beliefs. Failure to attain organizational goals indicates that current beliefs about action-outcome link are flawed and that the organization needs to actively explore new knowledge to update beliefs. The tendency to explore new knowledge therefore depends on performance feedback from the organizational level. The model predicts three distinct sources of change in firm strategy. First, changes may be induced by the failure to attain organizational goals and the attainment gap prompts problemistic search (Cyert & March 1963; Denrell 2008). Second, changes could emanate from experimentation buffered by slack resources. These two sources figure prominently in the behavioral theory of the firm. The third, new source of change is surprise that leads to a revision of prior beliefs and changes in the pattern of resource allocations.

METHODS

Research Setting

Testing our predictions requires a research setting in which firms have to constantly allocate resources to different activities in different domains with an unknown performance outcome. This requirement is met in the case of product introductions. Prior to launching new products firms have to invest considerable resources in research and development activities with unknown performance outcomes. Further, products can be classified into different market domains or product niches in which firms could have already experience. While product introductions are crucial for the success and the survival of firms in all industry this is even more pronounced in industries in which products have short lifecycles, like the video game industry.

Video games make most of their revenues in the first months, weeks or even days after their release. We thus use the product introduction decisions of video game publishers as a test bed for our theory. A video game publisher is a firm that publishes video games whose development they initially financed. Games can either be developed by an in-house developer studio, i.e. a studio owned by the publisher or by an independent third-party developer. While developer studios are usually specialized in different genres, publishers maintain relationships with different developers so that they can publish games in several genres. Like books or movies, video games can be classified into different genres. Each genre displays idiosyncratic characteristics in terms of storytelling, gameplay and technical requirements so that a genre can be considered as a different product niche.

Data and Sample

Our study is based on data collected from three different sources. First, data on the revenues, the publisher and the genre of the game is taken from the NPD database. NPD is an international market research company that monitors amongst others soft- and hardware sales in the video game industry in the US and Canada from 1995 until the present day. Second, financial information on video game publishers is taken from the January 2012 Internet version of Osiris by Bureau van Dijk. Osiris is a global database that provides balance sheets and income statements on over 45,000 companies from over 140 countries. Third, information on the developer of the game, which is used to calculate the share of games that a publisher developed in-house, comes from MobyGames, the world's largest and most detailed video game documentation project, containing comprehensive information on more than 63,000 games published between 1972 and 2012.

While the MobyGames database is fairly comprehensive and covers almost the whole industry from its early days, our final sample includes only observations between 1997 and 2008. As mentioned, NPD did not record sales data before 1995 and we needed a two year window to calculate our belief measure. The second constraint comes from the fact, that Osiris only provides financial information on firms, which have to report this information and in what detail. Hence, our sample may be biased towards large firms in counties with demanding reporting standards .However, the survivorship bias in our sample should be limited as Osiris provides not only information on active but also on dissolved firms.

Measures

Dependent variable. The variable *number of games* denotes the number of newly released games in a genre in a year. We rely on the genre classification of the NPD and use the following 11 genres in our analysis: action, adventure, arcade, children's and family entertainment, fighting, flight, racing, role-playing, shooter, sport games and strategy.

Independent variables. To test our first hypothesis on the positive relationship between the feedback signal and the resource allocation decision we include the variable *belief.* To capture the idea of a reinforcement learning process that is nonstationary we follow Sutton and Barto (1998) and Yechiam and Busmeyer (2005) and build the variable as follows:

$$(3)\ B_{g,t+1} = B_{g,t} + \propto \left(R_{g,t} - B_{g,t}\right),$$

where B denotes the belief about the performance in a genre, R is the revenue per game in a genre, g is the genre, t is a time subscript and α is a step-size parameter between zero and one, that gives a constant weight to recent information, i.e. the difference between the belief and the performance in the previous period. This empirical strategy to elucidate the unobservable beliefs of decision-makers is often used in psychological research to fit observed behavior to alternative specifications of belief updating. It is also similar to empirical approaches to measuring unobservable organizational aspirations in the management literature (e.g. Greve 2003).

To find the appropriate value of α we conducted a grid search, i.e. we calculated the variable *belief* for each value of α between 0.01 and 0.99 before we ran our full regression model with the different specifications of α. We obtained the best overall model fit for a value of $\alpha = 0.28$, indicating a rather low weight for recent information.

Social comparison. The decision how many games to release in a genre might not only be affected by the belief about the success in a specific genre, but also by the overall success of the firm. As a measure of overall success of the firm we include the variable *social comparison*. Following the literature on performance feedback (Greve 2003) we define *social comparison* as the difference between the social aspiration level, i.e. the average performance of all other firms, and the performance of the focal firm, where performance is measured as the annual return on assets. The variable is above zero if a firm outperforms its rivals, and negative otherwise. As firms react differently to positive and negative performance feedback (Cyert &March 1963; Greve 2003) we split the variable social comparison into two parts. The variable *neg. social comparison* takes on the values of social comparison if the performance of the firm is below the average performance of rival firms and zero otherwise. Contrary, *pos. social comparison* takes on the values of social comparison if the performance feedback is positive i.e. if the firm performs better that rival firms, and zero otherwise.

Slack resources. According to the Behavioral Theory of the Firm, slack resources enable firms to explore new possibilities. Exploring new possibilities can moderate the relationship between the belief about the success in a niche and the amount of resources devoted to further exploit the niche. We thus include the variable *slack resources,* which is measured as the quick ratio, i.e. the ratio of cash and cash equivalents divided by current liabilities (Mishina et al., 2010; Combs & Ketchen, 1999).

Internal developer resources. As mentioned before, some publishers own in-house developer studios specialized in a small set of genres. These in-house developer resources can be interpreted as specific resources that amplify the impact of the belief on the resource allocation. We build the variable *internal developer resources* as the share of games developed in-house.

Control and indicator variables. To control for firm, industry and genre specific effects that might influence the resource allocation decision across the different genres, we include a set of control and indicator variables on all three levels.

First, as larger firms might release more games than smaller firms we include the variable *firm size* measured as the turnover of a firm in a given year. Further, the decision to launch games might also be influenced by the past success of the firm. We thus include the variable *firm performance* measured as return on assets of the firm. To control for time invariant factors not captured by the size and the performance of the firm we include 57 publisher dummies.

Second, the resource allocation decision might also be driven by the overall volume of the video game industry. Hence, we include the variable *industry size* measured as the sum of revenues of all games sold in the U.S. and Canada in a year.

Third, as some genres attract more gamers but also more publishers we include 10 genre dummies to control for genre specific characteristics that influence the number of games launched in a genre.[1]

Estimation procedure

To model the number of newly released games in a genre in a year we adopt a zero-inflated, negative binomial model with Huber-White standard errors. This approach is suitable for the following reasons. First, the nature of the dependent variable as a non-negative count variable requires a Poisson model as linear regression models may result in inefficient, inconsistent, and biased estimates if the dependent variable is not continuous (Long 1997). A likelihood-ratio test of alpha being equal to zero reveals that a negative binomial model is preferable to a true Poisson

[1] We also included year dummies to capture overall growth trends in the industry. However, none of them have been significant.

model (chibar2= 3739.54; pr>z =.0000). In contrast to the true Poisson model a negative binomial model does not assume equality of the conditional mean and variance of the underlying distribution (Barron, 1992). Second, as about half of our observations have a zero entry for the dependent variable we further investigate whether a zero-inflated negative binomial model is preferable to a negative binomial model. The former type of model assumes that zeros and positive counts can be generated by a Poisson process, but zero counts occur with a certain probability from a separate process, which can be estimated by using a probit function. Put differently, zero-inflated Poisson models estimate in the first step the likelihood of the occurrence of a zero and in a second step the occurrence of positive counts. To decide which model to adopt we follow the procedure suggested by Vuong (1989) and test the null whether two models are as close to the actual model. As the null can be rejected (z= 10.96; pr>z =.0000) we favor the zero-inflated negative binomial model. Third, to correct for potential interdependence of observations we used Huber-White standard errors and adjusted the standard errors for intrafirm correlation (Wooldridge, 2002).

As zero-inflated negative binomial regression models, like all count data models, implicitly assume an exponential form we logged all variables in the binomial regression to obtain a log-log specification. The log-log specification is suitable as we assume a linear relationship between the dependent and the independent variables. Further, to avoid problems of reverse causality, we lagged all independent variables by one year. Finally, to check for multicollinearity in our analysis we estimated the variance inflation factors of the independent variables. As all variance inflation factors are below two we cannot detect any multicollinearity issues.

RESULTS

FIGURE 1
The Basic Model of Reinforcement Learning

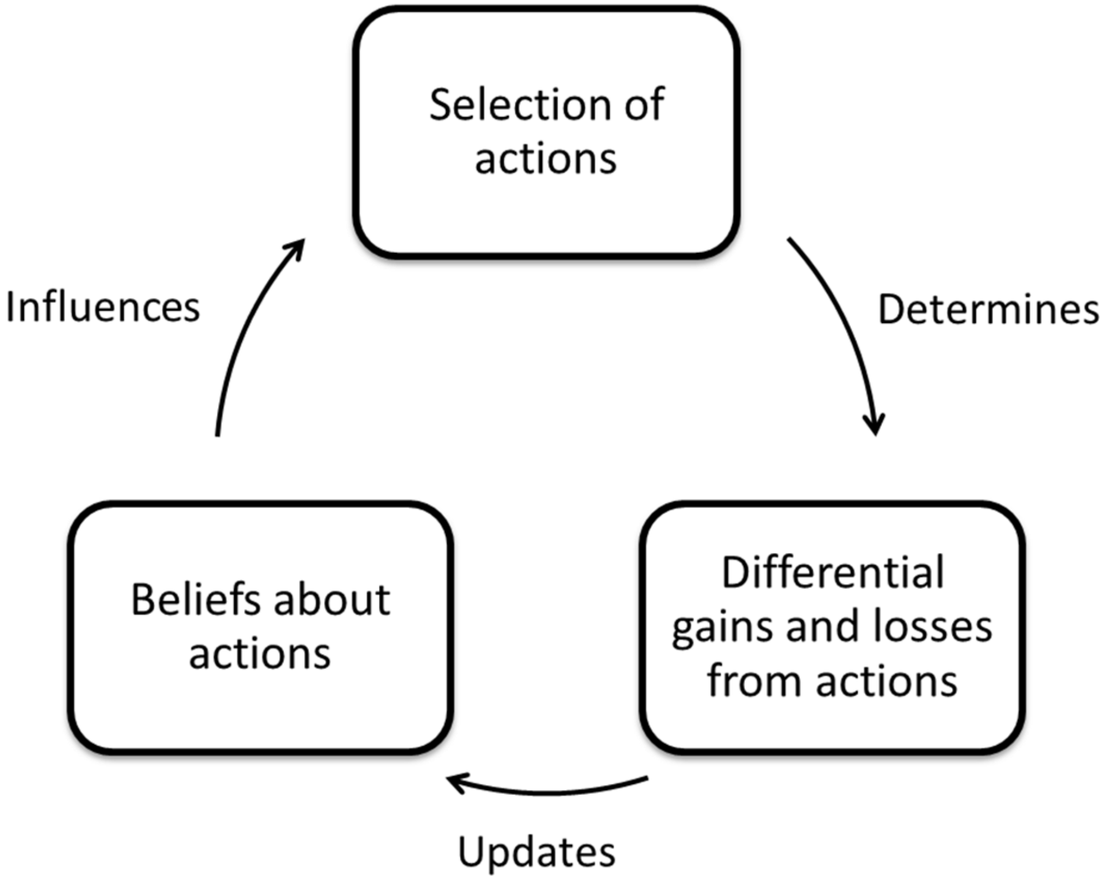

Table 1 provides paired correlations and descriptive statistics of the variables used in the study. While a publisher releases on average about 3 games in a genre in a year the number of games ranges from 0 to 95.

INSERT TABLE 1 ABOUT HERE

We present the basic results of the zero-inflated negative binomial regression in Table 2. Among our control variables only *firm size* has a positive and significant (p<.05) effect on the number of new releases in a genre indicating that larger firms release more games. *Industry size* is only positive and significant (p<.01) in column 1 before we introduce our main variables.

INSERT TABLE 2 ABOUT HERE

Hypothesis 1 states that more positive feedback from prior activities in a domain leads to more resource allocation in the domain in the future. Throughout the model we find a positive and significant (p<.01) effect of *belief* on the dependent variable supporting Hypothesis 1.

Before we turn to our hypotheses on the moderating effect of negative and positive performance feedback separately, we test the effect of the overall social comparison measure as a moderator. As shown in Column (3) we find no significant effect on the number of games released.

However, if we split the variable into negative and positive performance feedback we find a negative and significant (p<.05) coefficient of the interaction term of *belief* and *neg. social comparison*. We further analyze the interaction effect by looking at the marginal effect of *belief* at a low and a high value of the moderating variable, where the low (high) value is one standard deviation below (above) the mean of *neg. social comparison*. Each calculated marginal value has a standard error and implied z-statistic value. As shown in Table 3 the relationship between *belief* and the number of new games is less positive at higher values of *neg. social comparison* but both marginal effects are highly significant ($p<.01$). This indicates that, holding fixed the value of all model variables at their sample mean value, higher values of *neg. social comparison* reduce the impact of *belief* on the number of games released. Taken together, these results broadly confirm Hypothesis 2a.

INSERT TABLE 3 ABOUT HERE

Hypothesis 2b predicts that positive performance feedback at the organizational level reinforces the positive effect of beliefs on the number of releases in a genre. In line with the hypothesis we find a positive interaction effect between *pos. social comparison* and *belief* but it is only significant in Columns (5) and (6). The marginal effect of *belief*, as shown in Table 3, is significant for both low and high values of belief ($p<.01$) and it is increasing in the value of *pos. social comparison*. In sum, we find mixed empirical evidence for Hypothesis 2b.

Hypothesis 3 states that larger amounts of slack resources reduce the positive effect of the belief about the success in a genre on the number games launched in a genre. Indeed, we find a positive and significant ($p<.01$) coefficient of the interaction effect between *belief* and *slack resources*. The z-statistic in Table 3 shows that the marginal effect for a high and a low value of *slack resources* is significant ($p<.01$) and that higher levels of slack resources go along with a lower marginal effect of belief on the number of games. These results provide a strong support for Hypothesis 3.

Finally, in our last hypothesis we postulate that firms with larger stocks of specific resources rely more on their *beliefs*. Column (6) in Table 2 shows a negative but only marginally significant ($p<.01$) interaction effect of *belief* and *internal developer resources*, our measure of a specific resource, on the number of releases. As shown in Table 3 the z-statistic reveals that both marginal effects are significant ($p<.01$) and lower values of internal developer resources lead to smaller levels of the belief variable. Summarizing, Hypothesis 4 is also supported.

Altogether, our results suggest that more positive feedback from an action in a specific domain leads to higher resource allocations in this domain. We further find that the marginal effect of the positive feedback from the action is decreasing for higher values of negative performance feedback at the organizational level and for higher values of slack resources. Contrary, the marginal effect of the positive feedback from an action is increasing in both the levels of positive performance feedback at the firm level and the amount of specific resources.

DISCUSSION

One of the central questions in management research is how firms allocate resources to strategic actions to gain superior performance over time. This task is especially challenging if the outcome of the action is uncertain and subject to changes over time. Following the literature on reinforcement learning, we develop a simple model that leads to our first hypothesis that managers allocate more resources to a strategic action if prior activities in the domain have been successful but reduce the amount of activities if the performance feedback is negative. We combine our basic model on reinforcement learning with features from the behavioral theory of the firm and the resource-based view. From the former we take the idea that firms only search for new idea if they encountered a problem or if they are well endowed with slack resources. From the resource based view we borrow the idea that prior resource commitments create switching costs reducing a manager's willingness to search for new alternatives.

Using an unbalanced panel of 4,422 publisher-genre year observations and controlling for genre, firm and industry specific characteristics we find good support for our hypotheses. Throughout all models we find a strong support for our reinforcement learning hypothesis indicating that greater performance feedback from prior activities in a specific domain lead to more resource allocation in the domain in the future. We further find evidence that the positive relationship is moderated by positive and negative performance feedback at the organizational level, by slack resources and by prior resource commitments. While we only find a weak support for the idea that positive performance feedback at the organizational level and specific resource commitments reinforce the positive effect of reinforcement learning, we find a strong support that both negative performance feedback at the firm level and for slack resources weaken the relationship between beliefs and resource allocations and promote the exploration of inferior alternatives.

Contributions

Our research contributes to prior theoretical and empirical work in a number of ways. First, our empirical findings strongly support earlier ideas of strategy as an iterated process of resource allocations regulated by reinforcement learning principles (Dierickx & Cool 1989; Noda & Bower 1996). Even in a fast-changing business environment such as the video gaming industry, managers rely on prior feedback to form beliefs about the attractiveness of alternative strategic actions and base their resource allocation decisions on these beliefs. Importantly, beliefs are not based on most recent feedback, but rather reflect a process of much slower updating of beliefs also based more distant feedback signals. This empirical finding substantiates theoretical work on organizational adaptation that argued that slow learning about actions is an appropriate response to more uncertainty in the environment (Herriott et al. 1985; Denrell & March 2001).

Furthermore, our empirical findings confirm some of the basic tenets of the behavioral theory of the firm about the role of problemistic and of slack search in bringing about changes in firm strategy. These findings are well-supported in prior empirical work (Argote & Greve 2007). What is new in this context is that the proposed model of reinforcement learning points to a third distinct source of organizational and strategic changes that is notably absent from recent contributions to the behavioral theory of the firm (e.g. Greve 2003; Argote & Greve 2007). The model suggests that feedback induces a revision of beliefs, which in turn will change the observed pattern of resource allocation. Belief-based changes occur independent of problemistic search prompted by the failure to attain goals and of slack search enabled by excess resources in the organizations.

Second, we offer a simple, but powerful theory of firm strategy and industrial dynamics over time. Managerial beliefs drive resource allocations of firms into strategic actions. The resource allocations affect the relative competitive positions of all firms in the industries. The relative strengths of these competitive positions distribute gains and losses among competitors. Importantly, the performance properties of a strategic action often depends on what other firms do and these competitive interactions change performance properties over times. The gains and losses from competitive interactions have a differential effect on the revision of beliefs, thereby often creating and maintaining heterogeneity in managerial beliefs, strategic actions, and performance outcomes. While this analysis is suggestive and very broad, it nevertheless points to an interesting avenue for further research in strategic management in general and the resource-based view in particular. One the vexing questions there is where strategic resources come from that sustain performance differentials over time (Ahuja & Katila 2004; Gavetti & Rivkin 2007; Coen & Maritan 2011). Our analysis conceptualizes resource development as a process of iterative resource allocations that is guided by managerial beliefs and the feedback that changes them over time. Analyzing the interrelationship between feedback, belief formation, and resource allocation holds the promise of making further progress in understanding the origins of strategy and competitive advantages.

Third, reinforcement learning may provide a useful platform to organize and integrate diverse streams of literatures and to stimulate further work. Reinforcement learning is an active field of research in psychology, neuroscience, and economics and it thereby provides a firm foundation to model bounded rationality in managerial decision-making. Its insights may help researchers to think about and integrate a set of issues relating to managerial beliefs, learning, and resource application. For example, further research could look into how firms explore new actions. The reinforcement learning model suggests that the exploration of new knowledge is nevertheless guided by prior beliefs and that organizations tend to explore actions that they already know something about rather than actions that are completely unknown to them. This idea may be useful to analyze how firms respond to radical environmental changes, how they manage slack resources in exploration activities, and how they select alliance partners and targets for diversification.

Managerial Implications

Our study does not offer any managerial implications, since we are primarily interested in understanding how firms behave and learn to allocate resources over time rather than whether these resource allocation lead to superior performance. Still, the empirical finding that reinforcement learning is salient for strategic behavior points to clear path to developing managerial implications by drawing on the normative work that has studied how boundedly rational managers should behave in iterative resource allocation tasks. For example, Denrell and

March (2001) show how slow learning enhances the ability to identify superior actions. Stieglitz et al. (2009) and Posen and Levinthal (2011) show that firms may be better off by relying more on current beliefs the more unpredictable the business environment becomes. That is, unpredictable turbulence calls for less exploration, not more. Fang and Levinthal (2009) and Rahmandad (2008) analyze effective managerial behavior in situations with delayed feedback. Delayed feedback increases the danger of settling with an inferior alternative and this may be compensated by choosing a higher tendency to explore.

Limitations

Despite its attractive features, theoretical models of reinforcement learning also have some important limitations that restrict their application to problems of managerial decision-making and industry dynamics. First, the strategic actions available to the manager are usually assumed to be given. While the arrival of new actions could be thought of a stochastic process – new actions emerging in the business environment – the question of where new actions come from cannot be really addressed in a reinforcement learning framework. Given the strong interest of where new actions come from in strategy and in entrepreneurship other approaches are better suited to understand how entrepreneurs and firms recombine resources to stake out new competitive positions and create new strategic actions such as markets, niches, products, and capabilities (Becker et al. 2006; Denrell et al. 2003; Hsieh et al. 2007). Ideas from reinforcement learning only come into play the moment a new action is discovered (McMullen & Shepherd 2006). Another limitation that may be a fruitful area for further research is how beliefs based on prior experience interact with forward-looking cognition and the imagination of the future. That is, managers do not only rely on past experience to make sense out of the future, but may also try to imagine how the world will change by drawing on mental models or pattern recognition (Gary

& Wood 2011; Nadkarni & Narayanan 2007; Gavetti & Rivkin 2007). These important issues are beyond the scope of our research.

Our empirical study also has limitations. First, it is conceivable that the rate of changes in the business environment has a large impact on how beliefs map onto resource allocations over time. Our present study does not allow for a comparison of different rates of changes and their ramifications for learning over time. Second, our sample is biased toward large firms and small and large firms – due to differences in internal decision-making processes (Noda & Bower 1996) – may differ in the extent to which beliefs shape resource allocations. Third, we do not test whether learning rates are heterogeneous across firms.

REFERENCES

Ahuja, G. & Katila, R., 2004. Where do resources come from? The role of idiosyncratic situations. *Strategic Management Journal*, 25(8-9), pp.887–907.

Amburgey, T.L. & Miner, A.S., 1992. Strategic Momentum: The Effects of Repetitive, Positional, and Contextual Momentum on Merger Activity. *Strategic Management Journal*, 13(5), pp.335–348.

Argote, L. & Greve, H.R., 2007. A Behavioral Theory of the Firm--40 Years and Counting: Introduction and Impact. *Organization Science*, 18(3), pp.337–349.

Argote, L. & Miron-Spektor, E., 2011. Organizational Learning: From Experience to Knowledge. *Organization Science*, 22(5), pp.1123 –1137.

Arthur, W.B., 1991. Designing Economic Agents that Act like Human Agents: A Behavioral Approach to Bounded Rationality. *The American Economic Review*, 81(2), pp.353–359.

Baum, J.A.C., Li, S.X. & Usher, J.M., 2000. Making the Next Move: How Experiential and Vicarious Learning Shape the Locations of Chains' Acquisitions. *Administrative Science Quarterly*, 45(4), pp.766–801.

Baum, J.A.C. & Singh, J.V., 1994. Organizational Niches and the Dynamics of Organizational Mortality. *American Journal of Sociology*, 100(2), pp.346–380.

Becker, M.C., Knudsen, Thorbjorn & March, James G., 2006. Schumpeter, Winter, and the sources of novelty. *Ind Corp Change*, 15(2), pp.353–371.

Bhardwaj, G., Camillus, J.C. & Hounshell, D.A., 2006. Continual Corporate Entrepreneurial Search for Long-Term Growth. *Management Science*, 52(2), pp.248–261.

Bourgeois III, L.J., 1981. On the measurement of organizational slack. *Academy of Management review*, pp.29–39.

Bromiley, P., 1991. Testing a Causal Model of Corporate Risk Taking and Performance. *The Academy of Management Journal*, 34(1), pp.37–59.

Camerer, C., Ho, T. & Chong, K., 2003. Models of Thinking, Learning, and Teaching in Games. *The American Economic Review*, 93(2), pp.192–195.

Coen, C.A. & Maritan, C.A., 2011. Investing in Capabilities: The Dynamics of Resource Allocation. *Organization Science*, 22(1), pp.99 –117.

Cohen, J.D., McClure, S.M. & Yu, A.J., 2007. Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1481), pp.933 –942.

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

Daw, N.D. et al., 2006. Cortical substrates for exploratory decisions in humans. *Nature*, 441(7095), pp.876–879.

Dayan, P. & Niv, Y., 2008. Reinforcement learning: The Good, The Bad and The Ugly. *Current Opinion in Neurobiology*, 18(2), pp.185–196.

Denrell, J., 2008. Organizational risk taking: adaptation versus variable risk preferences. *Industrial and Corporate Change*, 17(3), pp.427 –466.

Denrell, J., 2003. Vicarious Learning, Undersampling of Failure, and the Myths of Management. *Organization Science*, 14(3), pp.227–243.

Denrell, J., Fang, C. & Levinthal, D.A., 2004. From T-Mazes to Labyrinths: Learning from Model-Based Feedback. *Management Science*, 50(10), pp.1366–1378.

Denrell, J., Fang, C. & Winter, S.G., 2003. The Economics of Strategic Opportunity. *Strategic Management Journal*, 24(10), pp.977–990.

Denrell, J. & March, James G., 2001. Adaptation as Information Restriction: The Hot Stove Effect. *Organization Science*, 12(5), pp.523–538.

Dierickx, I. & Cool, K., 1989. Asset Stock Accumulation and Sustainability of Competitive Advantage. *Management Science*, 35(12), pp.1504–1511.

Dobrev, S.D. & Kim, T.-Y., 2006. Positioning among Organizations in a Population: Moves between Market Segments and the Evolution of Industry Structure. *Administrative Science Quarterly*, 51(2), pp.230–261.

Dobrev, S.D., Kim, T.-Y. & Carroll, G.R., 2003. Shifting Gears, Shifting Niches: Organizational Inertia and Change in the Evolution of the U.S. Automobile Industry, 1885-1981. *Organization Science*, 14(3), pp.264–282.

Erev, I. & Barron, G., 2005. On Adaptation, Maximization, and Reinforcement Learning Among Cognitive Strategies. *Psychological Review*, 112(4), pp.912–931.

Fang, C. & Levinthal, D., 2009. Near-Term Liability of Exploitation: Exploration and Exploitation in Multistage Problems. *Organization Science*, 20(3), pp.538–551.

Gary, M.S. & Wood, R.E., 2011. Mental models, decision rules, and performance heterogeneity. *Strategic Management Journal*, 32(6), pp.569–594.

Gavetti, G. & Rivkin, J.W., 2007. On the Origin of Strategy: Action and Cognition over Time. *Organization Science*, 18(3), pp.420–439.

Ghemawat, P., 1991. *Commitment*, Free Press.

Greve, H.R., 1998. Managerial cognition and the mimetic adoption of market positions: what you see is what you do. *Strategic Management Journal*, 19(10), pp.967–988.

Greve, H.R., 2000. Market Niche Entry Decisions: Competition, Learning, and Strategy in Tokyo Banking, 1894-1936. *The Academy of Management Journal*, 43(5), pp.816–836.

Greve, H.R., 2003. *Organizational learning from performance feedback*,

Helfat, C.E. et al., 2007. *Dynamic Capabilities: Understanding Strategic Change in Organizations*, Oxford: Blackwell.

Herriott, S.R., Levinthal, D. & March, James G., 1985. Learning from Experience in Organizations. *The American Economic Review*, 75(2), pp.298–302.

Hsieh, C., Nickerson, J.A. & Zenger, T.R., 2007. Opportunity Discovery, Problem Solving and a Theory of the Entrepreneurial Firm. *Journal of Management Studies*, 44(7), pp.1255–1277.

Jacobides, M.G. & Winter, S.G., 2005. The Co-Evolution of Capabilities and Transaction Costs: Explaining the Institutional Structure of Production. *Strategic Management Journal*, 26(5), pp.395–413.

Kale, P. & Singh, Harbir, 2007. Building firm capabilities through learning: the role of the alliance learning process in alliance capability and firm-level alliance success. *Strategic Management Journal*, 28(10), pp.981–1000.

King, A.A. & Tucci, C.L., 2002. Incumbent Entry into New Market Niches: The Role of Experience and Managerial Choice in the Creation of Dynamic Capabilities. *Management Science*, 48(2), pp.171–186.

Klepper, S., 1996. Entry, Exit, Growth, and Innovation over the Product Life Cycle. *The American Economic Review*, 86(3), pp.562–583.

Kraatz, M.S. & Zajac, E.J., 2001. How Organizational Resources Affect Strategic Change and Performance in Turbulent Environments: Theory and Evidence. *Organization Science*, 12(5), pp.632–657.

Lave, C.A. & March, J. G, 1975. *An introduction to models in the social sciences*, New York: Harper & Row.

Levinthal, D.A. & March, James G., 1993. The Myopia of Learning. *Strategic Management Journal*, 14, pp.95–112.

Levitt, B. & March, James G., 1988. Organizational Learning. *Annual Review of Sociology*, 14, pp.319–340.

Lippman, S.A. & Rumelt, R.P., 2003. A Bargaining Perspective on Resource Advantage. *Strategic Management Journal*, 24(11), pp.1069–1086.

Makadok, R., 2003. Doing the Right Thing and Knowing the Right Thing to Do: Why the Whole Is Greater than the Sum of the Parts. *Strategic Management Journal*, 24(10), pp.1043–1055.

Makadok, R., 2001. Toward a Synthesis of the Resource-Based and Dynamic-Capability Views of Rent Creation. *Strategic Management Journal*, 22(5), pp.387–401.

March, James G., 1991. Exploration and Exploitation in Organizational Learning. *Organization Science*, 2(1), pp.71–87.

McMullen, J.S. & Shepherd, A., 2006. Entrepreneurial action and the role of uncertainty in the theory of the entrepreneur. *Academy of Management Review*, 31(1), p.132.

Mosakowski, E., 1997. Strategy Making under Causal Ambiguity: Conceptual Issues and Empirical Evidence. *Organization Science*, 8(4), pp.414–442.

Nadkarni, S. & Narayanan, V.K., 2007. Strategic schemas, strategic flexibility, and firm performance: the moderating role of industry clockspeed. *Strategic Management Journal*, 28(3), pp.243–270.

Nadkarni, Sucheta & Barr, P.S., 2008. Environmental context, managerial cognition, and strategic action: an integrated view. *Strategic Management Journal*, 29(13), pp.1395–1427.

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

Noda, T. & Bower, J.L., 1996. Strategy Making as Iterated Processes of Resource Allocation. *Strategic Management Journal*, 17, pp.159–192.

Nohria, N. & Gulati, R., 1996. Is Slack Good or Bad for Innovation? *The Academy of Management Journal*, 39(5), pp.1245–1264.

Peteraf, M.A., 1993. The Cornerstones of Competitive Advantage: A Resource-Based View. *Strategic Management Journal*, 14(3), pp.179–191.

Posen, H.E. & Levinthal, D.A., 2011. Chasing a Moving Target: Exploitation and Exploration in Dynamic Environments. *Management Science*. Available at: http://mansci.journal.informs.org/content/early/2011/10/14/mnsc.1110.1420.abstract [Accessed February 3, 2012].

Rahmandad, H., 2008. Effect of Delays on Complexity of Organizational Learning. *MANAGEMENT SCIENCE*, 54(7), pp.1297–1312.

Shinkle, G.A., 2012. Organizational Aspirations, Reference Points, and Goals. *Journal of Management*, 38(1), pp.415 –455.

Singh, J.V., 1986. Performance, Slack, and Risk Taking in Organizational Decision Making. *The Academy of Management Journal*, 29(3), pp.562–585.

Stieglitz, N., Knudsen, Thorbjørn & Becker, M.C., 2009. Strategic Focus and the Quest for Temporary Advantage. *Academy of Management Proceedings*, pp.1–6.

Sutton, R.S. & Barto, A.G., 1998. *Reinforcement learning: An introduction*, The MIT press.

Teece, D.J., 2007. Explicating dynamic capabilities: the nature and microfoundations of (sustainable) enterprise performance. *Strategic Management Journal*, 28(13), pp.1319–1350.

Teece, D.J., Pisano, G. & Shuen, A., 1997. Dynamic Capabilities and Strategic Management. *Strategic Management Journal*, 18(7), pp.509–533.

Vanneste, B.S. & Puranam, P., 2008. Repeated Interactions and Contractual Detail: Identifying the Learning Effect. *Organization science*, p.orsc.1080.0402.

Voss, G.B., Sirdeshmukh, D. & Voss, Z.G., 2008. The Effects of Slack Resources and Environmental Threat on Product Exploration and Exploitation. *Academy of Management Journal*, 51(1), pp.147–164.

Winter, S.G., 2000. The Satisficing Principle in Capability Learning. *Strategic Management Journal*, 21(10/11), pp.981–996.

Yechiam, E. & Busemeyer, J.R., 2005. Comparison of basic assumptions embedded in learning models for experience-based decision making. *Psychonomic bulletin & review*, 12(3), p.387.

Zollo, M. & Singh, H., 2004. Deliberate learning in corporate acquisitions: post-acquisition strategies and integration capability in US bank mergers. *Strategic Management Journal*, 25(13), pp.1233–1256.

Zollo, Maurizio & Winter, S.G., 2002. Deliberate Learning and the Evolution of Dynamic Capabilities. *Organization Science*, 13(3), pp.339–351.

FIGURE 1
The Basic Model of Reinforcement Learning

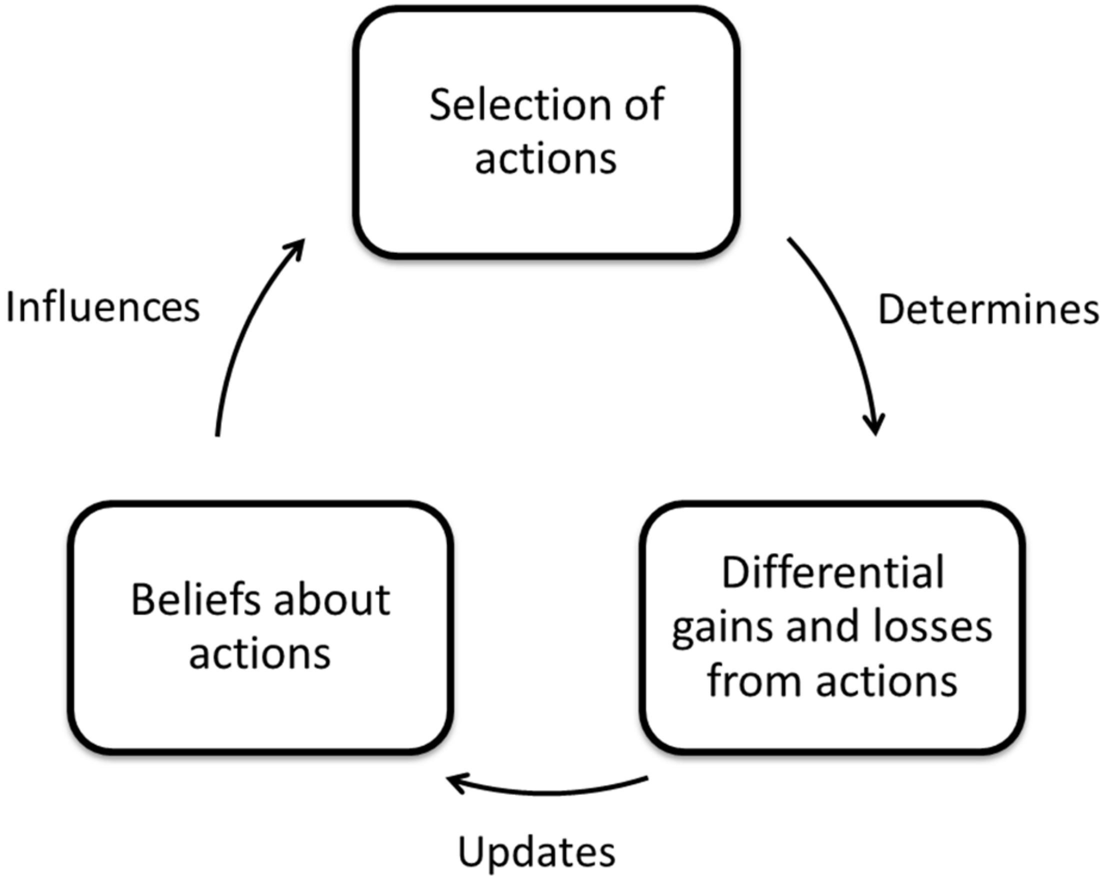

TABLE 1
Descriptive Statistics and Correlations [a]

Variable	Mean	S.D.	Min	Max	1	2	3	4	5	6	7	8	9
1 Number of games	3.06	6.50	0	95	1								
2 Belief (α=0.28)	0.07	0.18	0	4.08	0.12*	1							
3 Social comparison	0.31	0.59	-7.02	1.63	0.04*	0.03	1						
4 Neg. social comparison	-0.08	0.43	-7.02	0	0.05*	0.03	0.86*	1					
5 Pos. social comparison	0.39	0.31	0	1.63	0	0.02	0.70*	0.23*	1				
6 Slack resources	0.84	1.09	0.00	9.36	-0.01	0.25*	0.09*	0.07*	0.07*	1			
7 Internal developer resources	0.38	0.31	0.00	1	0.01	0.07*	0.06*	0.10*	-0.02	0.14*	1		
8 Firm size	12.30	2.73	0	18.30	0.19*	0.24*	0.14*	0.13*	0.08*	-0.03	0.24*	1	
9 Firm performance	-0.03	0.52	-7.335	2	0.04*	0.04	0.89*	0.90*	0.43*	0.08*	0.13*	0.13*	1
10 Industry size	6.18	1.75	2.59	9.47	0.01	0	0.38*	0.07*	0.62*	0.05*	-0.11*	0.07*	-0.05*

[a] n= 4,488 observations. * denotes significance at the 1% level.

TABLE 2
Results of Zero-Inflated Negative Binomial Regression [a]

	(1)	(2)	(3)	(4)	(5)	(6)
BINOMIAL REGRESSION: NUMBER OF GAMES[b]						
Belief (α=0.28)		3.15***	3.18***	2.53***	3.49***	2.76***
		(0.29)	(0.38)	(0.47)	(0.65)	(0.73)
Belief x social comparison			-0.12			
			(0.78)			
Belief x neg. social comparison				-2.75***	-2.43**	-2.49**
				(1.05)	(1.06)	(1.07)
Belief x pos. social comparison				1.84	2.09*	2.09*
				(1.22)	(1.22)	(1.24)
Belief x slack resources					-1.18***	-1.29***
					(0.41)	(0.41)
Belief x internal developer resources						2.24*
						(1.23)
Social comparison	0.02	0.00	0.01			
	(0.10)	(0.10)	(0.12)			
Neg. social comparison				0.16	0.14	0.15
				(0.16)	(0.16)	(0.16)
Pos. social comparison				-0.11	-0.18	-0.17
				(0.25)	(0.25)	(0.25)
Slack resources	-0.02	-0.06	-0.06	-0.05	0.06	0.07
	(0.08)	(0.08)	(0.08)	(0.08)	(0.09)	(0.09)
Internal developer resources	0.15	0.17	0.17	0.15	0.17	0.00
	(0.21)	(0.22)	(0.22)	(0.22)	(0.22)	(0.24)
Firm size	0.42**	0.46***	0.46***	0.48***	0.49***	0.48***
	(0.16)	(0.17)	(0.17)	(0.17)	(0.17)	(0.17)
Industry size	0.25***	0.10	0.10	0.07	0.07	0.06
	(0.09)	(0.10)	(0.10)	(0.11)	(0.11)	(0.11)
Constant	-4.36**	-1.22	-1.26	-0.49	-0.56	-0.37
	(2.11)	(2.15)	(2.17)	(2.53)	(2.54)	(2.53)
Publisher and genre dummies	Included	Included	Included	Included	Included	Included
PROBIT REGRESSION: INFLATE						
Firm Performance	0.70***	0.79***	0.80**	0.80**	0.80**	0.80**
	(0.25)	(0.31)	(0.31)	(0.33)	(0.33)	(0.33)
Slack resources	-0.27***	-0.33**	-0.33**	-0.33**	-0.32**	-0.32**
	(0.10)	(0.13)	(0.13)	(0.13)	(0.13)	(0.13)
Firm size	-0.53***	-0.54***	-0.54***	-0.54***	-0.54***	-0.54***
	(0.12)	(0.13)	(0.13)	(0.13)	(0.14)	(0.14)
Industry size	0.00***	0.00**	0.00**	0.00**	0.00**	0.00*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Constant	3.45***	3.59***	3.59***	3.61***	3.62***	3.62***
	(1.09)	(1.31)	(1.31)	(1.32)	(1.34)	(1.35)
Publisher and genre dummies	Included	Included	Included	Included	Included	Included
Observations	4,422	4,422	4,422	4,422	4,422	4,422
Chi 2	2458	2681	2682	2683	2694	2695
Log likelihood	-7620	-7532	-7532	-7529	-7524	-7522

[a] Huber-White robust standard errors are in parentheses.
[b] All variables in the binomial regression are logged.
*** p<0.01, ** p<0.05, * p<0.1 (two-tailed tests).

TABLE 3
Moderating Effect Analysis of Different Variables on the Marginal Effect of Belief on the Number of Games [a]

Level of moderating variable	Value of moderating variable	Marginal effect of belief	z-statistic
Low level of neg. social comparison	-0.33	13.77	8.18***
High level of neg. social comparison	0.00	10.61	9.16***
Low level of pos. social comparison	0.09	9.31	6.20***
High level of pos. social comparison	0.53	12.43	8.45***
Low level of slack resources	0.06	13.71	6.35***
High level of slack resources	0.93	9.11	9.72***
Low level of internal developer resources	0.08	8.81	5.54***
High level of internal developer resources	0.52	12.96	8.32***

[a] n= 4,488 observations. *** $p<0.01$, ** $p<0.05$, * $p<0.1$ (two-tailed tests). The high (low) value of the moderating variable is its value one standard deviation above (below) its sample mean.