Strategic Management Journal
Strat. Mgmt. J., **37**: 2545–2568 (2016)
Published online EarlyView 15 February 2016 in Wiley Online Library (wileyonlinelibrary.com) DOI: 10.1002/smj.2479
Received 25 May 2010; *Final revision received 25 October 2015*

CONSEQUENCES OF MISSPECIFIED MENTAL MODELS: CONTRASTING EFFECTS AND THE ROLE OF COGNITIVE FIT

DIRK MARTIGNONI,[1] ANOOP MENON,[2] and NICOLAJ SIGGELKOW[2]*
[1] *Institute of Management, Università della Svizzera italiana, Lugano, Switzerland*
[2] *Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania, U.S.A.*

***Research summary*: *Mental models, reflecting interdependencies among managerial choice variables, are not always correctly specified. Mental models can be underspecified, missing interdependencies, or overspecified, containing nonexistent interdependencies. Using a simulation model, we find that under- and overspecification have opposite effects on exploration, and thereby, performance. The effects are also opposite, depending on whether a manager controls all choice variables. The mechanism underlying our results is a feedback loop: misspecified mental models influence managerial learning about the effectiveness of choices; this learning guides how the environment is explored, which in turn, affects which information will be generated for future learning. We explore implications of these results for strategic management and introduce the notion of "cognitive fit" between the mental model of the decision-maker and the strategic environment.*

***Managerial summary*: *Managers often rely on mental models to guide their decision-making. These mental models, however, are often misspecified, that is, more or less complex than the situation managers are facing. Using a simulation model, we study the consequences of such misspecified mental models. We find that the performance implications of misspecified mental models crucially depend on whether the manager controls all choice variables. We identify situations in which simpler mental models are better than overly complex ones, and vice versa.*

INTRODUCTION

Managerial decision-making is often guided by mental models: models that reflect the expected consequences of chosen activities, and the assumed interdependence relationships either among these activities, or between these activities and variables not under the control of the decision-maker. For instance, activities along the value chain might interact with each other; likewise, some activities might be particularly valuable given certain environmental conditions. Managerial mental models, however, do not always accurately capture the true interdependence structure. In some cases, managerial mental models are cognitive simplifications of complex realities (Bettis and Prahalad, 1995; Eisenhardt and Sull, 2001; Schwenk, 1984), that is, mental models are underspecified in that they ignore some existing interdependencies; they contain causal "blind spots" (Zajac and Bazerman, 1991). In other cases, for example, through superstitious learning (Levinthal and March, 1981), managers' mental models may be overspecified, that is, containing relationships that do not actually exist.

Keywords: mental models; cognitive fit; simulation; interdependencies; learning
*Correspondence to: Nicolaj Siggelkow, 2000 SHDH, Wharton School, Philadelphia, PA 19104. E-mail: siggelkow@wharton.upenn.edu

It is not surprising that managers' mental models are misspecified: Even for researchers who can rely on large datasets, identifying and statistically validating interdependence relationships is a very difficult undertaking (Athey and Stern, 1998), and has led to many conflicting results, for instance, in the contingency literature (Donaldson, 2001). If mental models are not accurate, the question of what effects are caused by different kinds of misspecifications in mental models naturally arises. What are the consequences of having a mental model that is overspecified, that is, assuming interdependencies that do not exist? What are the consequences of having a mental model that is underspecified and blind to interdependencies that actually do exist? Under which conditions is over- and underspecification particularly costly? For instance, does it make a difference whether the interdependencies that are misperceived are *external* interdependency relationships, in which the value of an organizational choice is affected by a variable that is not under the control of the manager, or are *internal* interdependency relationships, in which the value of an organizational choice is affected by another variable that is under the control of the manager? To start addressing these questions, we build a simulation model that allows us to systematically analyze the performance implications of having different misspecifications within mental models.

To be more precise, we define *mental models* as comprising of two components, the *interdependence representation* that captures an individual's beliefs about the interdependencies among the elements in the mental model, and the *performance representation* that captures the assumed mapping from actions to performance outcomes. In the remainder of the article, for the sake of convenience, we use the phrase *under-/overspecified mental models* instead of *under-/overspecified interdependence representations within the mental models*.

A short example will clarify the types of situations we analyze. Consider a manager who contemplates how to best promote a new product introduction. The manager may have a number of alternatives: run a TV commercial, place an ad in a magazine, or distribute flyers in the streets. For a new product, the effectiveness of these different promotion choices may not be known to the manager at the outset. In more abstract terms, the performance landscape that the manager is facing is not known to him or her. To make a decision about which alternative to choose, the manager could decide to run some experiments. For instance, the manager could have flyers distributed and observe subsequent product sales. Given idiosyncracies on any given day, the signal the manager receives about the effectiveness of distributing flyers will be noisy. Over time, however, after running this experiment repeatedly, the manager would be able to reduce this noise. Likewise, the manager might run several TV commercials and obtain an estimate of their effectiveness. In this way, the manager will slowly create a *performance representation* of the actual performance landscape he or she is facing, that is, the manager develops a subjective performance landscape reflecting the manager's beliefs concerning the performance implications of the possible actions.

Now, consider an external contingency factor, the weather, for example. The manager might expect that certain interdependence relationships exist, for instance, that the weather has an impact on the effectiveness of distributing flyers in the street (sunny weather may improve mood, and hence, the likelihood of people looking at flyers). At the same time, the manager may believe that the weather does not have any impact on the effectiveness of TV ads. Thus, the manager possesses an *interdependence representation*, a cognitive representation of the assumed interaction structure between the various marketing alternatives and the (external) contingency factor "weather" that is not under his or her control.

Whether the weather actually has an impact on the effectiveness of TV ads or flyers is a different issue. We could imagine that in reality the weather has no systematic effect on any marketing alternative. In this case, the manager's interdependence representation (and mental model), which contains an interdependence between weather and distributing flyers, is overspecified—it includes an interdependence relationship that does not really exist. We could, however, also imagine that in reality the weather has an impact on all promotion activities. In this case, the manager's mental model, which does not contain the interdependence between weather and TV ads, would be underspecified.

For the set-up with internal interdependencies, imagine a slightly different situation in which the manager has control over two organizational choices: for instance, one decision among various marketing alternatives and a second decision of how to set the staffing level in the firm's call center. Again, the manager may have some beliefs about whether these choices interact or not. For instance,

DOI: 10.1002/smj

the manager might believe that the value of running a TV ad is higher when the call center is better staffed (due to a large increase in call volume), while the manager might believe that the value of distributing flyers is unaffected by the staffing of the call center. In contrast to the case of the weather, in this instance, the contingency variable ("staffing level in the call center") is under the manager's control.

Lastly, it should be noted that the terms *external* and *internal* do not necessarily refer to the boundaries of the firm. External interdependencies do not have to be external to the firm, such as the weather. The key aspect of *external* is that it concerns a variable that is beyond a decision-maker's control, while *internal* interdendencies relate to variables that are under the decision-maker's control whether or not they belong to the firm.

Directly measuring (the inaccuracy of) mental models is difficult for several reasons: The assessment of mental model inaccuracy requires a known true state of the world against which a mental model is compared (Edwards *et al.*, 2006). With the exception of experimental studies, this true state of the world is often not accessible. Instead, the accuracy of mental models is often assessed by comparing them to mental models of experts (Webber *et al.*, 2000). The simulation approach we pursue has the distinct advantage that we have complete control both over the true interaction pattern and over the mental models of our simulated managers. This allows us to systematically study the effects of over- and underspecified mental models.

In short, our results show that misspecified mental models can have an impact on the amount of exploration that a manager pursues and the subsequent performance that arises. The relationship between misspecification and exploration is, however, subtle. Unlike in many other studies of exploration, we find that (small) errors do not always lead to more exploration. We find that over- and underspecification have opposite effects, and the effects are different for external versus internal interdependencies. Overspecified mental models can lead to increased exploration for external interdependencies, while they tend to dampen exploration for internal interdependencies. Underspecified mental models have the reverse effects: They tend to decrease exploration for external interdependencies, and increase exploration for internal interdependencies, particularly for situations with high levels of complexity. Since exploration is beneficial in situations in which correctly specified mental models lead to quick lock-in, we can identify situations in which a misspecified mental model can even create higher performance than what is possible with a correctly specified mental model.

In the discussion section, we explore a number of implications of our results for strategic management. Our results suggest that managers who tend to have overspecified mental models—"complexifiers" or "differentiators"—might be a good fit lower down in the organizational hierarchy, while managers with more underspecified mental models—"simplifiers" or "generalizers"—might be better suited for the upper echelons. Our results also have an implication for how decision-making tasks might need to be (re-)structured to suit the cognitive styles (complexifier versus simplifier) of the managers involved in those decisions. When applied to the interorganizational and industry level, our results imply that strategic leadership with overspecified mental models might be better suited for industries with high levels of exogenous uncertainty (i.e., many variables of vital importance to firm performance that are outside the control of the firm's decision-makers) and vice versa, and also that within a given industry, large firms with significant influence on market structure might benefit from managers with underspecified mental models, while smaller, less influential firms might be better suited for managers who have overspecified mental models. Finally, we distill the more abstract point behind these implications by outlining the theoretical construct of "cognitive fit," that is, the degree of fit between the cognitive style of a manager and his or her strategic environment, a potentially useful concept to guide future research in the area.

EXISTING LITERATURE

The notion that the value of a particular activity is dependent on external factors or on how other activities are configured has been discussed in the literature under a variety of labels, including external contingencies (e.g., Lawrence and Lorsch, 1967), internal contingencies (e.g., Chandler, 1962), external or environmental fit (Miller, 1992; Siggelkow, 2001), internal fit (e.g., Miller and Friesen, 1984; Porter, 1996; Siggelkow, 2002), tight coupling (e.g., Levinthal and March, 1993), and complementarities (e.g., Milgrom and Roberts, 1995; Pettigrew, Woodman, and Cameron, 2001). Although invoking different labels, these contributions share the basic notion that the value of a specific choice or action is contingent on other—internal or external—factors.

There exists considerable empirical research trying to document the existence of interdependency relationships (for reviews see, for instance, Donaldson [2001], and Porter and Siggelkow [2008]). While some interdependencies may appear intuitive, showing their existence empirically is fraught with econometric difficulties (Athey and Stern, 1998). Unlike academic researchers, managers often have to rely on much smaller samples to identify interdependency relationships (March, Sproull, and Tamuz, 1991). As a result, it is not surprising that managers' mental models often do not accurately reflect the true existing interaction structure that is present. The mental models of managers may suffer from two forms of imperfections with respect to interdependencies: They may not contain existing interdependency relationships ("underspecification") or assume interdependency relationships that actually do not exist ("overspecification").

Prior research has provided evidence for both types of imperfections. For instance, Bettis and Prahalad (1995), Porac *et al.* (1995), Walsh (1995), Eisenhardt and Martin (2000), and Eisenhardt and Sull (2001) found that decision-makers often have simplified mental models, not capturing all complexities of the problems they are facing. These simplifications can arise from a number of sources. The reality may include too many interdependencies to keep track of, or managers may underestimate certain interactions (Heath and Staudenmayer, 2000). In other cases, managers may not be able to distinguish between various contingency states (e.g., is "misty rain" different from a "drizzle"?) that turn out to be relevant (Nelson, 2008). In this case, managers undercategorize, group certain interdependencies together, and end up with an underspecified mental model.

Managers who make inappropriate generalizations also have underspecified mental models: These managers believe that certain conditions (be they choices under their control or not) are not relevant to the performance of a focal activity. Thus, prior good (or poor) experience with this action is thought to generalize to a broad range of other conditions. For instance, if a manager applies an inappropriate analogy, such an error occurs (Gavetti and Rivkin, 2007). Likewise, managers who mistakenly believe that certain actions are "best practices"—good practices regardless of the context in which they are embedded—have underspecified mental models. In all of these cases, managers have, in the terminology of Zajac and Bazerman (1991), causal blind spots. In this study, we call managers with underspecified mental models "simplifiers."

In contrast, managers might also be affected by superstitious learning (Denrell, Fang, and Levinthal, 2004; Levitt and March, 1988). When cause-effect relationships are difficult to ascertain, managers may create beliefs about interdependencies that actually do not exist, falling prey to "apophenia" (Conrad, 1958) or "patternicity" (Shermer, 2000), the inclination of seeing patterns in meaningless data. Managers may perceive "illusory correlations" (Hogarth and Kolev, 2013), that is, they perceive correlations among unrelated factors. Experimental studies by psychologists and experimental economists have documented the prevalence of errors of "covariation assessment" (e.g., Einhorn and Hogarth, 1986); people may fail to understand covariation for even the simplest 2×2 contingency tables. Likewise, managers with neurotic tendencies, being overly anxious that a particular interdependency exists, even though it actually does not, would have overspecified mental models. More generally, "differentiators," that is, managers who treat every instance as a special case, who refuse to generalize, and who insist on differentiating between every experience may have overspecified mental models. For ease of exposition, we call managers with overspecified mental models "complexifiers" as their mental model is more complex than the true performance landscape would require.

Prior work has found not only that mental models of managers can be misspecified, but also that significant heterogeneity of mental models can exist across managers in an industry (e.g., Barr, 1998; Barr, Stimpert, and Huff, 1992; Eggers and Kaplan, 2009; Gavetti and Rivkin, 2007; Tripsas and Gavetti, 2000). These findings have raised the question of why managers have different mental models. Over the years, researchers have attempted to study certain categories or "types" of such variation in cognition. For example, studies in cognitive science have found an intriguing pattern in how experts in a domain approach a problem differently (in a cognitive sense), compared to novices (Chase and Simon, 1973; Gentner and Gentner, 1983;

DOI: 10.1002/smj

McCloskey, Caramazza, and Green, 1980). Experts tend to quickly perceive the "deep" structure of a problem in their area of expertise (with greater focus on how the various elements of the problem link to each other), while novices tend to focus on the "superficial" properties. Other studies, found in the management literature, have documented how variation in situational and organizational factors, such as the length of tenure of the top management team, or the size and relative position of a firm within an industry, can also lead to systematic variations in how managers perceive their strategic environment (Porac *et al.*, 1995; Sutcliffe, 1994).

One way to think about systematic variation in the mental models of decision-makers is that the variation might be due to two different factors. First, managers have innate properties, different "cognitive types," that influence the mental model. For instance, one might think about personality dimensions such as the degree of neuroticism of the individual and how that might lead to systematic variations in the mental models they adopt. Second, there exist more consciously chosen elements that one might call a "cognitive approach." For example, Cho and Hambrick (2006) look at the shift in the cognitive orientation of top management teams in the airline industry from an engineering one to an entrepreneurial one, and its implications. Another example of a change in cognitive approach could be the adoption of a data-driven decision-making paradigm by the strategic leadership of a firm. (In general, courses or books on decision-making attempt to change the cognitive approach of decision-makers.) Cognitive types and cognitive approaches together lead to categories of "cognitive styles" that managers possess as they perceive and interpret their strategic environments. In our study, we look at two particular cognitive styles, overspecified mental models ("complexifiers") and underspecified mental models ("simplifiers").

While there is much consensus that managers may suffer from both types of imperfections (Gary and Wood, 2011; Hackner, 1991; Nair, 2001), the implications of these errors for learning and performance are less clear. There is some evidence that accurate mental models, that is, mental models without errors of over- or underspecification, lead to better decisions and improve organizational long-run performance (Barr *et al.*, 1992; Gary and Wood, 2011; Gentner, Loewenstein, and Thompson, 2003). Others, including Sutcliffe (1994:1374), however, suggest that errors in mental models may have some positive effects: "Misperceptions may be beneficial if they enable managers to overcome inertial tendencies and propel them to pursue goals that might look unattainable in environments assessed in utter objectivity." This, of course, raises the question of what type of misperceptions—errors of over- or errors of underspecification—may have some beneficial effects.

On the one hand, errors of underspecification, that is, mental models that are too simple, may have some benefits if they trigger improvisation and explorative efforts (Dess and Beard, 1984; Eisenhardt and Sull, 2001) or accelerate decision-making (Eisenhardt and Martin, 2000). Only simplified mental models may help transform complex information environments into tractable ones (Porac, Thomas, and Baden-Fuller, 1989; Simon, 1991). On the other hand, there is also evidence that errors of underspecification come at a cost. Models that are too simple may be associated with a loss of important information and may "blind" managers to important developments in their environment (Walsh, 1995; Zajac and Bazerman, 1991).

Ashby's (1956) law of requisite variety is often invoked to highlight the importance of (adequate) cognitive complexity: The complexity of the mental model must match or exceed the complexity of the decision environment. Citing Ashby's law of requisite variety, Weick (1979: 35) argues that "it takes a complex sensing system to register and regulate a complex object." In a similar vein, Mintzberg (1973: 185) contends that "it is the power of his (i.e., manager's) mental model that determines to a great extent the effectiveness of his decisions" (parenthetical statement added). Yet, there also exists empirical evidence that mental models that are too complex, that is, that are overspecified, can come at a cost. In their empirical studies, Hackner (1991) and Nair (2001) found a negative relationship between errors of overspecification and performance.

In sum, there is little doubt that mental models are important to structure managerial decision-making. There also exists ample empirical evidence that documents under- and overspecification in managerial mental models. Yet, prior research makes conflicting predictions with respect to the expected performance implications of these misspecifications. In our study, we propose an alternative, learning-based model that allows us to identify conditions under which different types of misspecifications are particularly costly or may even create some value.

DOI: 10.1002/smj

More broadly, our results imply that certain cognitive styles, and their resulting mental models, fit better with certain strategic environments, that is, create a degree of "cognitive fit" that also needs to be considered in strategic decision-making.

MODEL

Our model consists of five basic elements: (1) a performance landscape; (2) a cognitive representation of the interdependency structure of the performance landscape; (3) a cognitive representation of the expected payoffs of choices; (4) a process of updating the performance representation; and (5) a process of choosing alternatives based on the current performance representation (see Figure 1 for a high-level overview). We first describe the model with external interdependencies and then the model with internal interdependencies.

Performance landscape

Consider a manager who needs to choose among C alternatives (e.g., he or she needs to choose among C different marketing alternatives); moreover there are R different states an external contingency variable can take (e.g., R different weather conditions). A performance landscape then consists of a matrix with C columns and R rows, containing in each cell the mean performance, m_{rc}, that is achieved in this combination. We assume that if a manager chooses the alternative in column c and "nature" chooses the condition in row r, the manager observes a performance realization that is a random draw from a normal distribution with mean m_{rc} and variance 1. In all our simulations, we use $C = R = 5$. Thus, in our case, a performance landscape is a 5×5 matrix with 25 cells (see Figure 2 for two examples of performance landscapes).

The interaction structure of the true performance landscape, that is, which contingency state affects which choice alternative, is reflected in the performance matrix in whether cells within a column share the same value m_{rc}. For instance, if choice 1 is unaffected by all contingency states (e.g., the value of a TV ad is really independent of the weather), then all entries in the first column would be the same (i.e., $m_{11} = m_{21} = m_{31} = m_{41} = m_{51}$). In contrast, if, for instance, choice 2 is affected by all contingency states (e.g., the value of distributing flyers is affected in different ways on whether it rains, snows, hails, the sun is shining, etc.), then all entries in the second column would be different, i.e., $m_{i2} \neq m_{j2}\ \forall\ i \neq j$.

Landscapes thus differ in the number of interdependencies they contain. At one extreme, each choice alternative is independent from all contingency states, that is, entries within each column are identical and the matrix contains only five different performance values, one for each column (Figure 2[a]). At the other extreme, each choice alternative depends on all contingencies states, that is, each cell contains a different performance value. Drawing on Simon (1962) and Thompson (1967), we say that complexity within a landscape increases as the number of interdependencies increases. To study intermediate degrees of complexity, we proceed as follows. We start with a matrix with full independence, that is, a matrix that has identical entries within each column. For each column, we draw a contribution from a normal distribution with mean 0 and variance 1. We then randomly pick a number of cells and require these cells to contain a different value m_{rc}, again drawn from a normal distribution with mean 0 and variance 1. The parameter "*complexity*" equals this number of cells. Thus, for *complexity* = 0, we retain a landscape with no interdependencies. If *complexity* = 20, each cell is required to contain a different value. (Since we started already with five different values [one in each column], we only need to change 20 cells to have all cells with different values.)[1] Lastly, one should note that the highest performance that is achievable in each landscape is unaffected by the level of complexity.[2]

Figure 2(a) depicts a landscape with *complexity* = 0. In this case, when the manager picks alternative 1, he or she experiences a performance outcome, regardless of the contingency state, which is a random draw from a normal distribution with mean 1.3 (and variance 1). In contrast, Figure 2(b) depicts a landscape with considerably more interdependencies (*complexity* = 12). Here, for instance, choice 1 is affected by all contingency states. In this case, if the manager picks alternative 1 and contingency state 1 arises, then he or she would experience

[1] If the matrix has R rows, and $R-1$ entries in a column have already been randomly selected to contain a different value, then we do not allow the R^{th} entry in that column to be chosen.

[2] Regardless of complexity, the performance landscape matrix has five different, randomly drawn numbers in each row. Thus, for each contingency state (i.e., in each row), the expected performance of the best alternative is the maximum of five randomly drawn numbers, all from the normal distribution with mean 0 and variance 1.

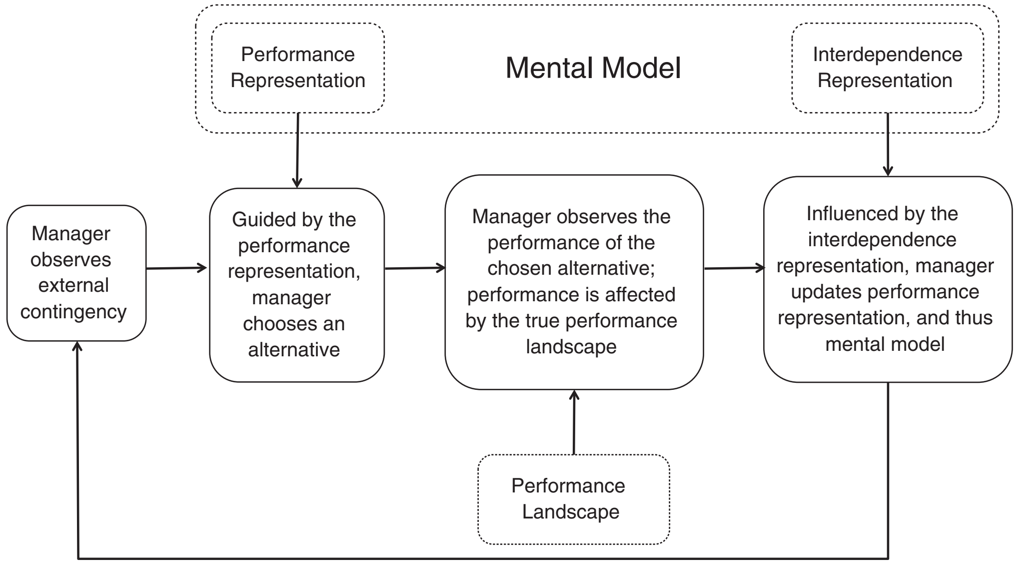

Figure 1. Elements of the simulation model

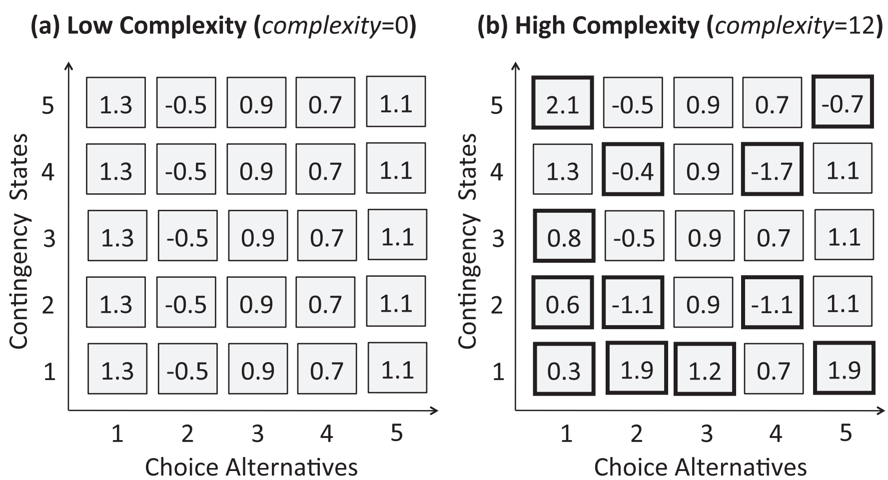

Figure 2. Illustration of (a) low complexity and (b) high complexity performance landscapes

a performance outcome that is a random draw from a normal distribution with mean 0.3 (and variance 1); if, however, the manager picks the same alternative and contingency state 2 arises, then the performance outcome is drawn from a distribution with mean 0.6.

Mental model of the performance landscape

Aspects of managerial cognition have been labeled in a variety of ways, including mental maps, mental models, frames of reference, mindsets, cognitive bases, schemata, representations, cognitive structures, and cognitive maps. In this article, we use three terms in very specific ways. We use the term *interdependence representation* to denote the beliefs that a manager holds with respect to the *interactions* that are present. Thus, the interdependence representation contains, for example, the beliefs "the effectiveness of TV ads is unaffected by the weather" and "for the effectiveness of distributing flyers it matters whether it rains or the sun is shining." We use the term *performance representation* to denote the beliefs that a manager holds with respect to the expected *payoffs* of actions given a particular contingency state (or, in

DOI: 10.1002/smj

the case of internal interdependencies, given a particular choice of the other activity). Thus, a performance representation contains, for example, the belief that "TV ads, when the sun is shining, yield on average a profit of 2.1" and "Flyers, when it is raining, yield on average a profit of −0.4." To use an analogy from the context of regressions, a manager's interdependence representation is akin to the choice of functional form ("What interaction effects are included in the model?"), while the performance representation is akin to the estimated parameters of the model ("What is the impact of a particular realization of a set of variables?"). A manager's "mental model" is comprised of both the interdependence representation and the performance representation.

Similar to the performance landscape, a manager's mental model can also be represented by an $R \times C$ matrix. The performance representation corresponds to the numerical entries in this matrix. The entry in column c, row r represents the manager's belief of the performance that would be achieved (on average) by choice c given contingency r. The manager's interdependence representation is reflected in how many different entries the mental model (i.e., the $R \times C$ matrix) allows for. For instance, if the manager believed that choice 1 is unaffected by all contingency states, then he or she would believe that all entries in the first column of his or her mental model contain the same value. In our figures of mental models, we represent the beliefs that certain choice/contingency combinations should share the same estimate by using the same color for these choice/contingency combinations. If the manager's mental model has a different color in each cell that contains a different value in the true performance landscape, then we say that the mental model is correctly specified. If the mental model allows for a larger number of different entries than are contained in the true performance landscape, then we say that a mental model is overspecified; if the mental model only allows for a smaller number of different entries than are contained in the true performance landscape, then we say that a mental model is underspecified.

By implication, when the true performance landscape has *complexity* = 0 and $C = R = 5$, a manager's mental model can only be overspecified, in the range of 0 to +20, containing up to 20 more assumed interdependency relationships than the real landscape warrants. (Since no interdependencies exist in this landscape, the manager cannot have fewer interdependencies in his or her mental model than in the true landscape; hence, underspecification is not possible.) For *complexity* = 10, misspecification runs from -10 to +10, encompassing both under- and overspecification. Lastly, for *complexity* = 20, only underspecification is possible, in the range 0 to -20. To create over- and underspecified mental models, we start with a correctly specified mental model and then either add or delete interdependencies.

At the beginning of each simulation (at $t = 0$), we assume that a manager has fairly little information about the performance landscape. In particular, we assume that the manager's mental model contains a 0 in each cell. In other words, while the manager's mental model is unbiased in aggregate (since all entries m_{rc} of the true performance landscape are drawn from a normal distribution with mean 0), in the beginning all choice/contingency combinations appear equally attractive. Only by experiencing a particular combination can the manager learn about the shape of the performance landscape and update his or her mental model.

In sum, in our figures depicting mental models, the interdependence representation of a manager is reflected in different cell colors, while the numerical entries within each cell correspond to his or her performance representation. Consider Figure 3. (To save space, we depict only three choice alternatives and three contingency states.) The left panel shows the mental model in period $t = 0$ of a manager who believes that each choice alternative is unaffected by the contingency states: All entries within each column share the same color. Thus, each choice alternative is believed to generate the same performance in each contingency state. At the beginning of the simulation run, the manager has not yet formed any particular beliefs about any choice/contingency combination, and his or her mental model contains an overall unbiased estimate of 0 for each combination. In contrast, the left panel of Figure 4 shows a mental model of a manager who believes that every contingency matters for each alternative. As a result, his or her mental model allows for different entries in each cell.

Updating the mental model

In the process of updating the mental model, the interdependence representation influences a manager's interpretation of information and mediates

DOI: 10.1002/smj

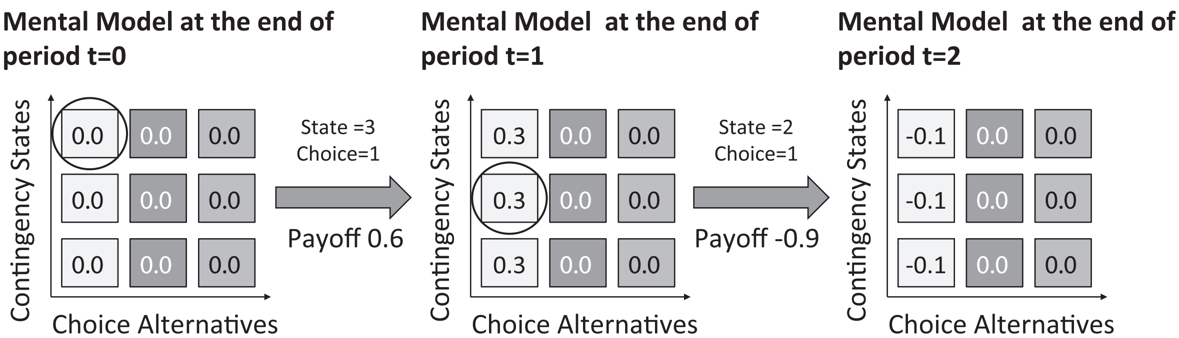

Figure 3. Illustration of the updating of a mental model given a no-interaction interdependence representation

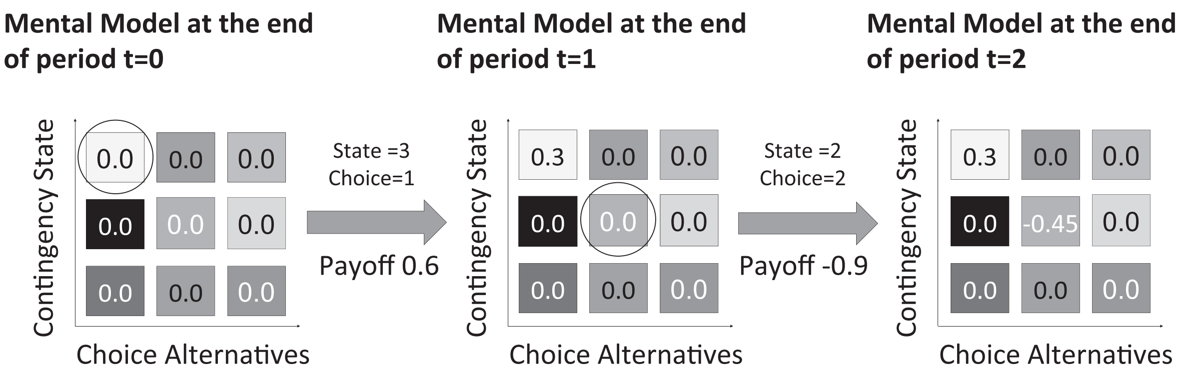

Figure 4. Illustration of the updating of a mental model given a complex interdependence representation

the manager's learning process. This is easiest seen with an example (see Figure 3 and Figure 4). Suppose that at the beginning of period 1, the manager observes that contingency state 3 has arisen; he or she decides to try alternative 1 and experiences a payoff of 0.6. Since the manager's interdependence representation contains the belief that alternative 1's value is unaffected by the external contingency, the manager updates the entire first column of the mental model at the end of period 1 (see middle panel of Figure 3): Since the manager believes that the contingency does not influence the value of the alternative, the manager generalizes his or her experience with this alternative, derived in one contingency state, and updates the mental model concerning this alternative for all contingency states. In contrast, if the manager's interdependence representation stipulates that alternative 1 is affected by all contingency states, then the manager would update the mental model only for the case "Alternative = 1, Contingency state = 3" (see middle panel of Figure 4). We assume that managers update their beliefs by taking the average over all observations in the particular (group of) cells. In this case, since the manager's prior estimate was 0, he or she replaces the 0 with 0.3 (= (0 + 0.6)/2) in the respective cells.[3]

Choosing actions

The performance representation guides the manager's decision-making. In particular, the manager is assumed to pick the alternative that he or she believes will yield highest performance given the contingency state, which is observable before an action is taken. If more than one alternative has highest performance, then the manager chooses randomly among them. Consider again the example depicted in Figures 3 and 4. In period 1, the manager is completely indifferent among all alternatives.

[3] We checked the robustness of our results to our assumption of flat priors, that is, all zero entries in the mental model at the beginning of the simulation. Replacing the flat priors with random draws from N(0,1), that is, the same distribution from which the true performance landscape was constructed, did not affect the results qualitatively. In general, though, the value of exploration due to misspecified mental models declined. Second, in our main model, the entry of "zero" in each cell of the mental model is included in the performance updates once a manager receives information about a cell. Not including this assumed (but not experienced) entry of zero did not affect the results in any significant way.

DOI: 10.1002/smj

Thus, the manager chooses randomly among the three alternatives and picks, for example, alternative 1. In period 2, the manager's decision will be affected by his or her interdependence representation since different interdependence representations lead to different updated performance representations. If the manager believed that alternative 1 is affected by all contingencies (as in Figure 4), then only if contingency state 3 arises again does the manager have a clear favorite action, namely, alternative 1. Should any other contingency state arise in period 2, the manager will be indifferent among the alternatives (since his or her performance representation for those states still contains only zeros) and will again choose randomly among the alternatives. In contrast, if the manager's interdependence representation states that alternative 1 is unaffected by the contingency states (as in Figure 3), then regardless of which contingency state the manager observes in period 2, the manager will pick alternative 1 (since the entry of 0.3 is higher than any other entries in his or her mental model). Thus, the interdependence representation has two effects: First, it influences which parts of the mental model will be updated given any new experience. Second, through the performance updates, it affects which alternatives will be chosen, and thus, which part of the performance landscape will be experienced, and consequently, which new data will be generated to create further updates of the performance representation.

To recap, in period 0, the manager's performance representation contains all zeros. In each subsequent period, the manager observes the contingency state (each contingency state has the same probability of occurring), chooses an action based on his or her performance representation, observes the outcome, and updates his or her performance representation (see again Figure 1). (For the relationship between our model and prior work using *n*-armed bandit models, please see Appendix S1.)

Internal interdependencies

The model with internal interdependencies is very similar to the one with external interdependencies. Now, however, the "contingency states" are not played by nature, but represent a second choice variable under the control of the manager. Each choice variable has its own performance landscape, resulting in a "global" performance landscape that contains the sum of the two individual performance landscapes. Likewise, the manager has two mental models, that is, two interdependence representations and two performance representations, one for each choice variable, and a resulting "global" performance representation, which contains the sum of the performance estimates of the two choices. We assume that the manager is able to observe the performance of each individual choice, and thus, is able to update his or her two performance representations concerning each individual choice. The manager picks the choice combination that has the highest performance estimate in his or her global representation, that is, the manager attempts to maximize the overall performance that the two choices generate.

We should note that we purposefully do not model managers who also update their beliefs about interactions, that is, their interdependence representation. We made this modeling decision for two reasons. First, in order to make headway on the issue of the effects of misspecified mental models, it seemed sensible to start by keeping interdependence representations fixed. Only once we understand better the effects of interdependence representations on performance, is a study of the changes of interdependence representations likely to yield fruitful results. Second, it seemed to make sense to have fixed interdependence representations rather than fixed performance representations because interdependence representations tend to be more deeply rooted within managers' minds than beliefs about particular payoffs. Consider, for instance, the experiences of the United States and West European car manufacturers when they first attempted to catch up to the Japanese manufacturers with respect to productivity (e.g., Jaikumar, 1986). Managers were willing to copy various new practices, for example, just-in-time delivery or investments in more flexible machinery. Once they implemented these practices, they revised their estimates of the benefits of these practices. Unfortunately, managers operated with outdated interdependence representations that did not take into account that the interdependencies in the Japanese lean manufacturing system were very different from those in the more buffered, traditional mass manufacturing system. It took many years before managers understood and revised their beliefs about the interdependencies among the practices that they were implementing (and thus, fully understood the performance implications of the various new practices). In this sense, within mental models, beliefs about payoffs, that is, elements within performance

DOI: 10.1002/smj

representations, appear to change faster than beliefs about interdependencies, that is, interdependence representations.

RESULTS

Our analysis is organized into two parts. First, we look at the context of external interdependencies and investigate the effects of misspecifications within mental models on short-term and long-term performance. Second, we examine the performance implications of misperceiving internal interdependencies. Since we have stochastic elements in our model (for example, the noise in the performance feedback), we average for each parameter setting our results over 40,000 replications.

Results for external interdependencies

We first focus on how different degrees of over- and underspecification in the mental model affect short- and long-run performance. In Figure 5, we show the effects of misspecifications on short-run performance (the simulation runs for 20 rounds; $t = 20$) and long-run performance ($t = 200$) for settings of *complexity* of 0, 5, 10, 15, and 20. At $t = 200$, performance for all managers has reached a plateau, equaling the performance at $t = 1{,}000$. Thus, we treat results in period 200 as the "long-run" performance. In contrast, for most situations, managers have not reached a performance plateau at $t = 20$; thus, we treat these results as "short-run" performance.

Each line in Figure 5 represents an environment of different complexity. The points on the vertical line of "*degree of misspecification* = 0" represent the performance achieved by a manager with a correctly specified mental model. Each point to the right of this vertical line corresponds to the performance achieved by overspecified mental models; points to the left of the line correspond to underspecified mental models. As noted in the model section, different degrees of complexity allow for different ranges of misspecification. For instance, in an environment with no interdependencies, a mental model cannot be underspecified. (In general, *degree of misspecification* + *complexity* is bounded by 0 and 20.) In total, Figure 5(a and b) contains the results of 105 different models (all possible levels of misspecification for the five degrees of complexity chosen), each of which was run 40,000 times, that is, summarizes the results of 4.2 million simulation runs.

Figure 5(a) contains the short-run results. Most results are as one might expect: (1) both under- and overspecification is costly as compared to having a correctly specified mental model (each line has its maximum at the vertical line at *misspecification* = 0); (2) performance declines monotonically with the degree of misspecification; and (3) performance declines with complexity.[4] We can also note that costs of over- and underspecification are fairly symmetrical, that is, the same degree of misspecification leads to similar performance penalties.

The cost of underspecification is driven by managers' making incorrect inferences across contingency states: managers use experiences (concerning the efficacy of an alternative) garnered in one contingency state and apply it in another state, even though such a generalization is not appropriate. In contrast, overspecification affects short-term performance because it slows down performance improvement. Given that each choice/contingency state combination is treated separately, the mental model is updated only cell by cell, requiring a longer time to create a fuller picture of the entire landscape. Thus, in this case, appropriate generalizations are not made, slowing down the process of finding the optimal response given a contingency state.

The long-run performance results, shown in Figure 5(b), are more subtle. First, we observe that in the long-run, underspecification tends to be much more harmful than overspecification. The short-run costs of overspecification, treating every case as a special case, decrease substantially in the long run. Thus, for external interdependencies, it appears to be better for managers to be complexifiers than to be simplifiers. However, the degree of overspecification, even in the long run, has some effects. Indeed, our second observation is that sometimes some overspecification can even be helpful. As the results for *complexity* = 0 (solid line) show, performance initially increases as misspecification increases from a correctly specified mental model. In this case, maximum performance is achieved when managers are slightly overspecified; as overspecification increases to very high levels, performance declines back to the level achieved by a correctly specified mental model. Thus, in this case,

[4] Recall, in our set-up, complexity does not affect the highest possible outcome; thus, in this case, complexity only makes it more difficult to find the best alternative given any external state.

DOI: 10.1002/smj

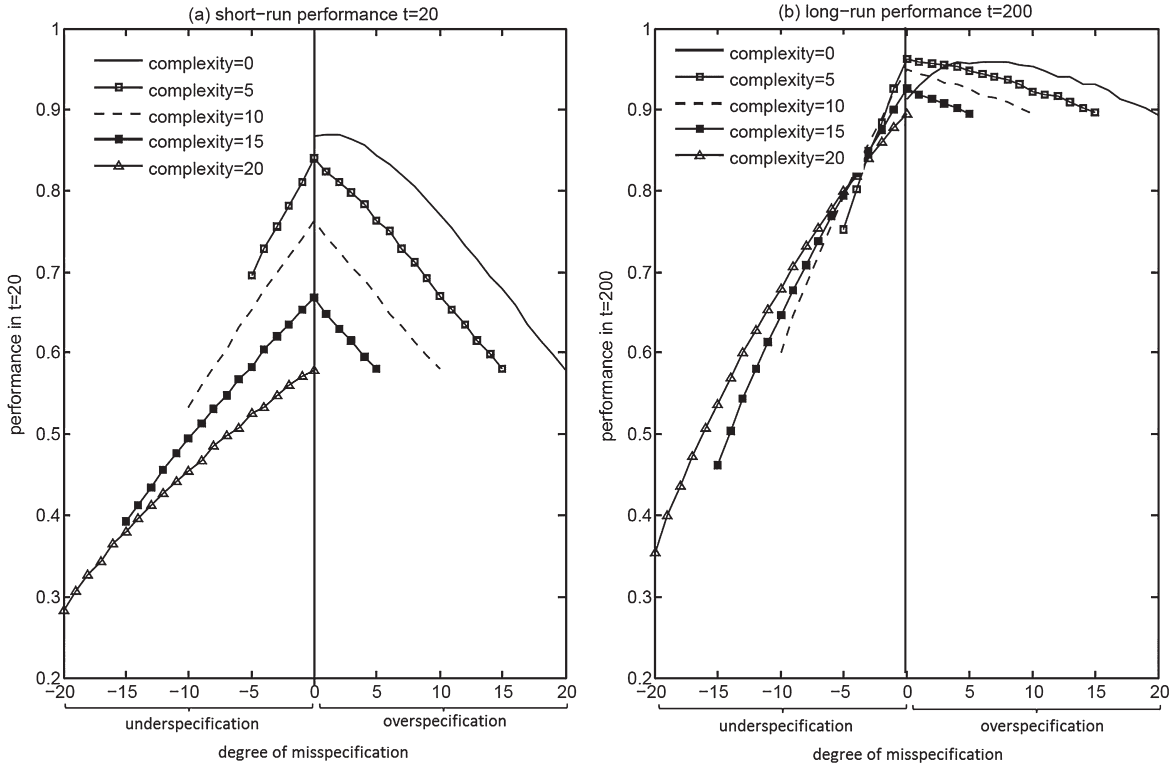

Figure 5. The effect on short-run performance (a) and long-run performance (b) of mental models that misspecify external interdependencies

being slightly overspecified can actually be beneficial. Third, when we look at the performance achieved by correctly specified mental models (i.e., the points along the vertical axis at *misspecification* = 0), we see that performance first increases with complexity, and then decreases.

What explains these results? Let us first start with the costs of underspecification. In the long-run, underspecification can lead to premature lock-in. Imagine a world with high complexity and a manager who believes that no interdependencies exist. Say contingency 1 arises in the first period, the manager picks alternative 1 and gets a return of 1. Since the manager does not believe that interdependencies exist, the manager would update his or her belief of the efficacy of alternative 1 to be 0.5 for all possible contingencies, that is, in the entire first column of his or her performance representation. (The manager's starting belief was 0; thus, averaging yields 0.5). In the next period, regardless of which external contingency arises, the manager would pick this alternative (since 0.5 is larger than 0 in all other cells of his or her performance representation). As long as this alternative does not lead to very negative performance in a contingency state (creating an overall negative estimate of the value of this alternative), the manager remains stuck with this alternative. In Figure 6, we provide more systematic evidence of this intuition. In this figure, we plot a measure of exploration (the number of different choice/contingency combinations that were chosen by $t = 200$) as a function of underspecification and complexity of the environment. As the figure shows, the higher the degree of underspecification, the less exploration a manager engages in.

Modest overspecification, in contrast, can generate more exploration. In situations of low complexity, a correctly specified mental model can lead to quick lock-in on the first good solution that is found. This happens because the manager assumes (correctly) that this solution will work in all contingencies. At the same time, however, this assumption effectively stops search for even better solutions that might exist. In contrast, with a slightly overspecified mental model, the manager might (wrongly) assume that a previously found good solution will not work in a new contingency state. As a result, the manager will continue to

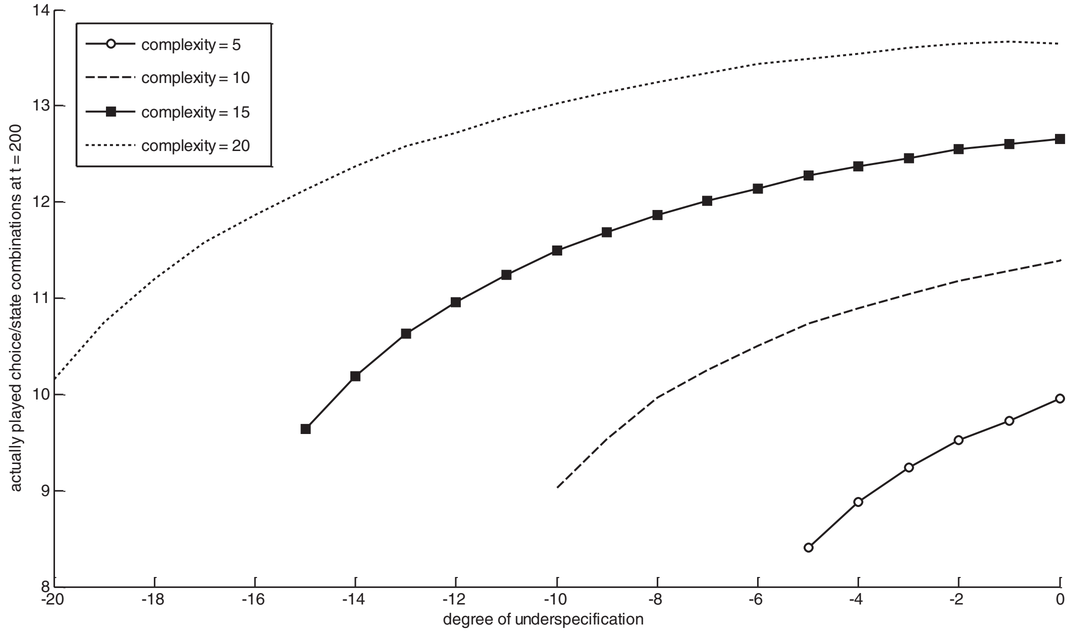

Figure 6. The effect of underspecification on exploration for external interdependencies

search should this state arise, thereby increasing the likelihood of finding a potentially even better solution (see Appendix S2 for an example of this mechanism).

At higher levels of complexity, too much overspecification harms performance again, however. Since each case is treated as a different one, learning about the efficacy of a choice in one contingency state is not transferred to another. As a result, once an above-average choice has been found for each contingency state, search effectively stops and other choice-contingency combinations are never tried and assessed. Figure 7 provides evidence of this intuition. Here, we plot the degree of exploration as measured by the updated entries in the performance representation as a function of overspecification for different levels of complexity. This measure of exploration tracks well the eventual performances achieved, including the inverse U-shape in the case of *complexity* = 0.[5]

Results for internal interdependencies

In the previous sections, we sought to determine the value of understanding external interdependencies, that is, interdependencies with variables that are beyond the control of the manager. In this section, we study the value of understanding the interaction effects among internal activities. As in the introduction, assume that a manager faces two decisions: a choice among marketing strategies and how richly to staff the calling center. Thus, while previously the second "decision," that is, the contingency state, was not controlled by the manager, now the manager chooses a combination of two activities. As before, we compare the effects of over- and underspecification on short-run ($t = 20$) and long-run ($t = 200$) performance for different levels of complexity. The only difference in model set-up is that each of the two individual payoffs has a variance of 0.5, rather than 1. Since the manager uses the sum of the individual payoffs to make decisions (his or her combined performance representation), the variance of this sum is thus again 1.

As seen in Figure 8, for both short- and long-term performance, it is less detrimental to be underspecified than to be overspecified. Moreover, for long-term performance, some underspecification can actually lead to higher performance than having the correct mental model. These results are in stark contrast to the results for external interdependencies, for which we found that underspecification was generally more harmful than overspecification.

[5] We included the case *complexity* = 1 in this figure to show that the inverse U-shape is not unique to *complexity* = 0. As complexity increases, the line slowly bends upward.

DOI: 10.1002/smj

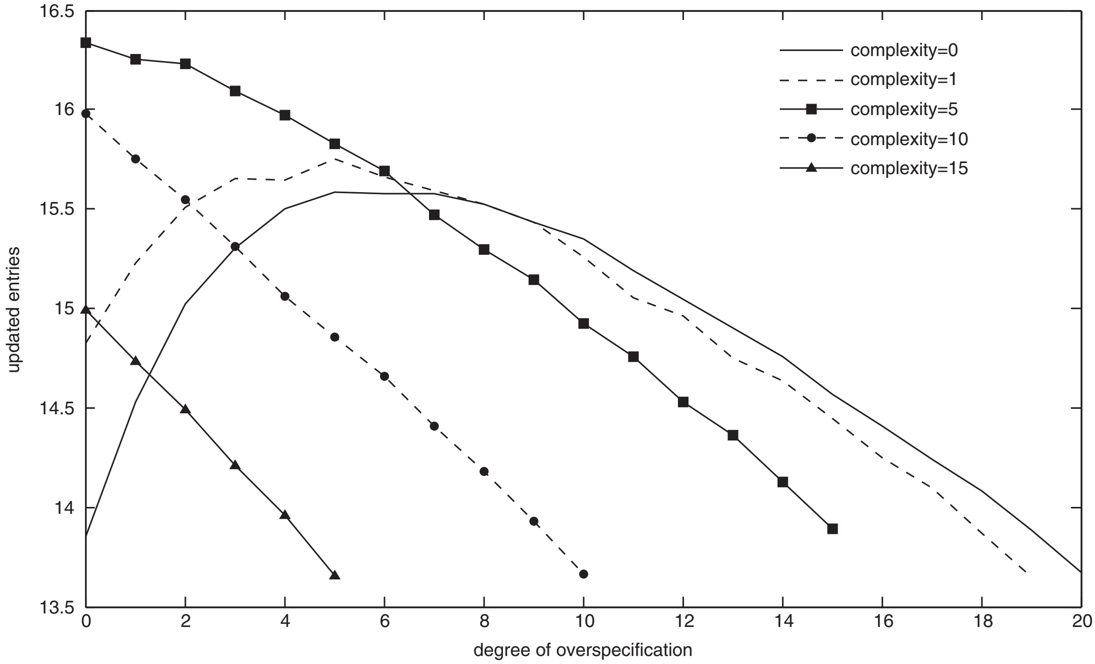

Figure 7. The effect of overspecification on exploration for external interdependencies

For internal interdependencies, overspecification is quite costly, particularly when complexity is low (see the solid line in Figure 8[b]). In the case of internal interdependencies, overspecification leads to rapid lock-in and a reduction of exploration. In the extreme case of a fully overspecified mental model, the manager would stick to the first combination of choices that generates above-average performance (see Figure 9 for systematic evidence that overspecification reduces exploration for internal interdependencies).

It is helpful to contrast this situation with the case of external interdependencies. The benefit of overspecification with external interdependencies is created because "nature" might put the manager into a contingency state for which he or she does not have yet formed an opinion. In that case, he or she would explore anew. With internal interdependencies, however, the manager has full control over which states (or precisely combination of activities) he or she experiences, since both activities are under his or her control. Thus, in this case, overspecification cannot lead to dislodging since it does not force the manager into situations for which he or she has not formed an opinion. The manager can simply decide to avoid these combinations.

Underspecification, in contrast, can lead to increased exploration in the case of internal interdependencies. Underspecification can yield increased exploration if one activity provides a positive return, the other a negative return, and the sum of both is positive. With a correctly specified mental model, the manager would stick to this combination of activities, since his or her joint return is high. With an underspecified mental model, the manager would wrongly assume that the high performance of the first activity could also be achieved if the other activity is differently configured. This wrong, generalizing assumption prompts the manager to stick with the chosen alternative for the first activity choice and try a different alternative for the second activity, thereby exploring a new activity combination. (For systematic evidence that underspecification increases exploration for internal interdependencies, see Figure 10.)

In sum, over- and underspecification of mental models have opposite effects on exploration (and performance); and these effects are again opposite for external and internal interdependencies. Overspecified mental models can in some cases increase exploration in the case of external interdependencies, but decrease exploration for internal interdependencies. In contrast, underspecified mental models decrease exploration for external interdependencies, but increase exploration for internal interdependencies.

DOI: 10.1002/smj

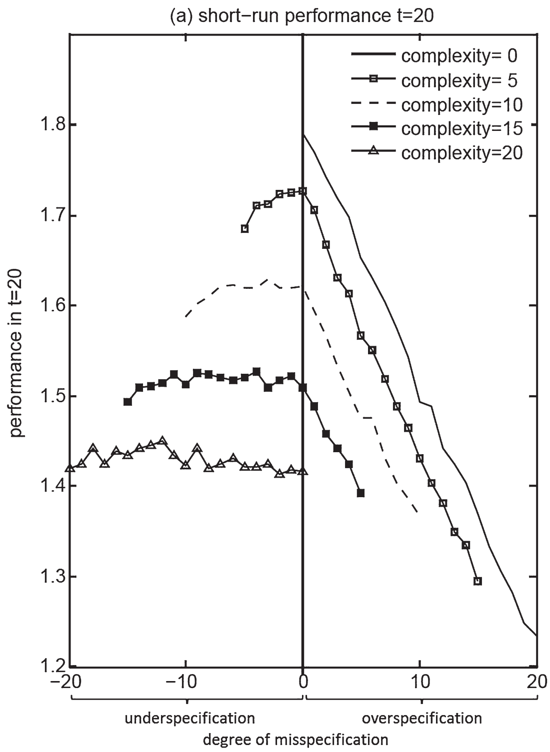

Figure 8. The effect on (a) short-run performance and (b) long-run performance of mental models that misspecify internal interdependencies

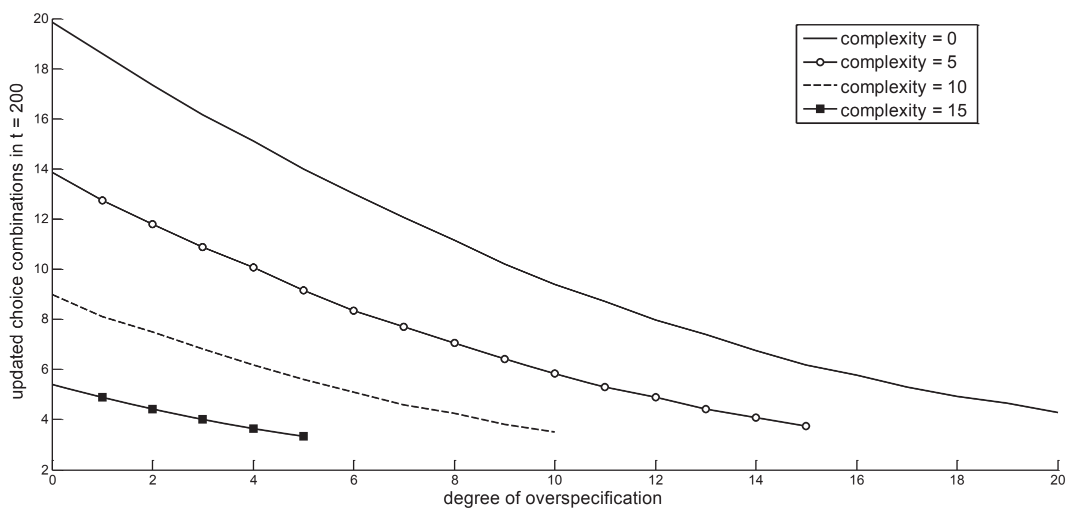

Figure 9. The effect of overspecification on exploration for internal interdependencies

Interactions with other sources of exploration

As with any model that focuses on a particular mechanism that creates (or inhibits) exploration, it is important to note that many other elements of a firm influence exploration, for example, its incentive system or organizational design (Siggelkow and Rivkin, 2005). These different sources of exploration can interact with each other. In supplementary simulations (results available from the authors), we show, for instance, that a lack of exploration induced by misspecified mental

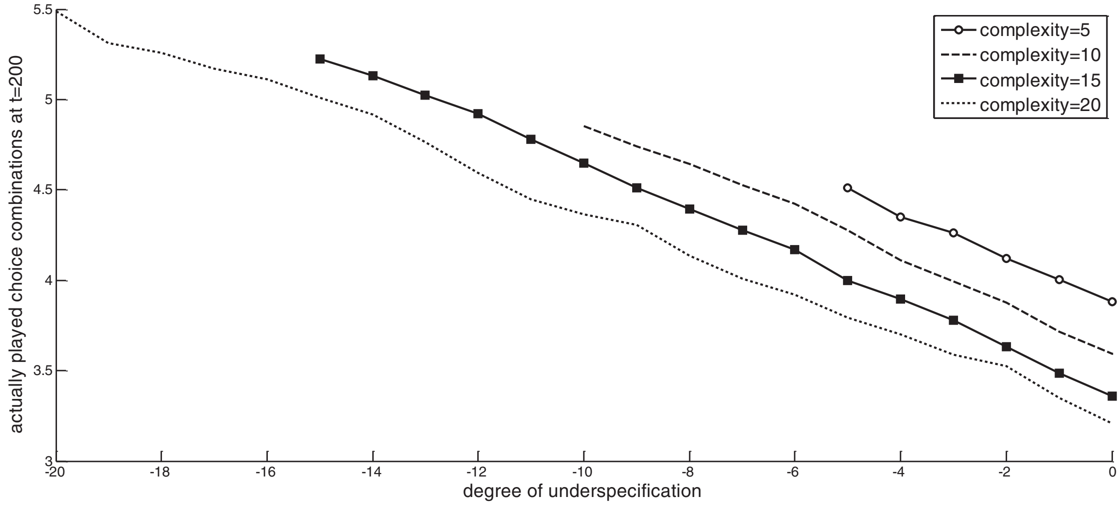

Figure 10. The effect of underspecification on exploration for internal interdependencies

models can in part be compensated for when managers "naturally" explore by not always picking the alternative with the highest performance.[6] These actions have short-term costs, but can lead to higher long-term outcomes. We find that exploration induced through the choice rule and exploration induced by misperceptions of the mental models are substitutes. Whenever misspecified mental models create exploration, they tend to reduce the value of direct exploration. Likewise, whenever misspecified mental models reduce exploration, they increase the value of direct exploration.

DISCUSSION AND CONCLUSION

Implications for research on mental models

With the help of our simulation model, we investigated the consequences of misspecifications in mental models on the amount of exploration, and in turn, performance. As noted in our earlier overview of the prior research in this arena, while there is agreement that managers do not always have a correctly specified mental model, there is much less agreement about the effects of these misspecifications. For instance, the literature on "simple rules" supplies compelling arguments for the benefits of simplified, underspecified mental models, whose benefits should be seen especially in complex environments in which decision-making can easily get bogged down (e.g., Eisenhardt and Sull, 2001). At the same time, however, there are equally compelling arguments that underspecified mental models, especially in situations of high complexity, can be harmful. For example, Ashby's Law of requisite variety (1956) would imply that as the degree of complexity of the situation increases, the complexity of the mental model should increase as well.

Our learning-based model suggests that there is some truth to both views. By focusing on three dimensions of the problem that had not previously been clearly separated simultaneously, namely, (1) the type of interdependence (internal versus external interdependencies); (2) the type of misspecification (under- versus overspecification); and (3) the extent of the complexity of these interdependencies (ranging from the absence of any interdependencies to highly interdependent choices), we were able to identify boundary conditions under which different kinds of misspecifications may promote or suppress exploration (and in turn, performance). With underspecified mental models, for instance, we find benefits in the case of internal interdependencies, but harmful consequences in the case of external interdependencies. Thus, more generally, our study contributes to existing theory by developing a more nuanced, contingent view of the effect of mental

[6] In particular, we assumed that the manager's choice follows a softmax choice rule (Erev and Barron, 2005; Fang and Levinthal, 2009; Sutton and Barto, 1998). With a softmax choice rule, the probability that alternative r (of N possible alternatives) given the contingency state c is chosen is given by $p_{rc} = [\exp(m'_{rc}/\tau)]/[\Sigma\exp(m'_{ic}/\tau)]$ where m'_{ic} is the manager's estimate of the value of choice i in contingency state c, the sum is taken over all N alternatives, and τ is a measure of the agent's proclivity to explore.

DOI: 10.1002/smj

	External Interdependencies	**Internal Interdependencies**
Overspecification	Slight overspecification can be beneficial for low complexity. Overspecification is in general not very harmful in the long run.	Overspecification is very harmful for all levels of complexity.
Underspecification	Underspecification is very harmful for all levels of complexity.	Underspecification can be beneficial, especially for high complexity.

Figure 11. Summary: The effects of misspecified mental models on performance

model misspecifications. Our findings, summarized in Figure 11, can be seen as new hypotheses that can guide future empirical work in this area.

Our results have a number of implications for research on mental models. First, we see that the exploration and performance implications of misspecifications in mental models depend crucially on the moderating variables mentioned above. The results reverse when we move across these moderators, for example, from external to internal interdependencies. As our results show, without drawing the distinction between these two types of interdependencies, different studies might come to different conclusions, possibly even contradictory ones, if the mix of internal and external interdependencies is different across them. Moreover, if one combined misperceptions concerning both types of interdependencies, one might draw wrong inferences with respect to the overall effect of misspecified mental models. For instance, in our model, had we not distinguished between these two types of interdependencies—effectively pooling our results from Figures 5 and 8—we would have identified a consistently negative performance effect of both under- and overspecification, thereby missing the possible positive effect of underspecified mental models in the context of internal interdependencies.

Likewise, over- and underspecification are often combined into one measure such as mental model accuracy (e.g., Gary and Wood, 2011). Yet, as our study shows, over- and underspecified mental models might have quite different effects on exploration and performance as mentioned above. Had we not made a distinction among different types of misspecification in our analysis, we would again have come to different conclusions. In this case, had we just measured "misspecification" and pooled our observations regardless of whether the misspecification was an over- or an underspecification (i.e., averaging the points that are equidistant to the vertical $misspecification = 0$ line in both Figures 5 and 8), we would have found a negative monotonic relationship between the degree of misperception and performance. Thus, we would have missed the inverted U-relationship for overspecified mental models in the case of external interdependencies and the positive relationship for underspecified mental models in the case of internal interdependencies.

Similar concerns arise regarding the results for the third moderator, namely, the complexity of the problem managers are facing. For instance, the positive effect of overspecified mental models appears to arise primarily in situations in which few true interdependencies exist. Likewise, the positive effect of underspecified mental models arises in particular in situations of high internal interdependency, that is, high complexity. Again, pooling our results across different levels of complexity would have obscured these effects. In sum, if these three moderators—type of interdependency, type of misspecification, and degree of complexity—are not taken into account, one might miss effects that misspecified mental models create. More troublesome, if empirical settings differ in the degree or mix of these moderators, and these moderators are not controlled for, then it is quite likely that inconsistent findings will emerge.

Besides pointing toward the potential importance of these moderators for future empirical research, we also believe that both conceptual and empirical analyses would benefit from making a distinction between what we called the decision-maker's "interdependence representation" and his or her "performance representation," that is, to make a distinction between beliefs about the presence and absence of interdependencies, and beliefs about the payoffs of particular choices. Prior research has often derived measures of mental model complexity from causal maps. Yet, causal maps quite frequently include both beliefs about the presence of interdependencies and the effect size (and sign) of these interdependencies (Barr *et al.*, 1992). Without this conceptual separation, measures of accuracy conflate misperceptions about the nature of the interdependencies and misperceptions of the effects of these interdependencies. For instance, at a particular point in time, a causal map might accurately reflect the presence of interdependencies (interdependence representation), but might be quite wrong with regards to the performance implications (performance representation). In this case, learning over time might correct the errors in the causal map since the interdependence representation is correct. Conversely, a causal map that contains misspecified interdependencies, but fewer (initial) errors in the performance representation, might over time degrade as the learning process is misguided by the misspecified interdependence representation.

Another point to note is that in some instances misspecified mental models can result in performance improvements over correctly specified ones. For instance, underspecified mental models resulted in higher performance than correctly specified ones when internal interdependencies are involved, and this benefit increased as the complexity of the situation increased. Thus, in some situations, instead of imposing constraints on the decision-maker, these "cognitive simplifications" actually helped improve performance as compared to having the correct model, through the increased exploration that it triggered.

It should be emphasized that the results obtained here are not simply the result of "(small) errors triggering search," which in turn, leads to improved performance, a common finding for search processes on highly rugged performance landscapes (e.g., in simulated annealing approaches). Instead, the *type* of error matters crucially. For instance, in a setting with internal dependencies, errors of overspecification can have the opposite effect, stifling exploration and performance. Similarly, in a setting with external dependencies, errors of underspecification can also stifle exploration and performance.

Finally, on a fundamental and conceptual level, we would like to re-emphasize that a mental model has two effects on the learning and decision-making processes. First, it influences how incoming data from the world is filtered and interpreted, which in turn, determines which parts of the performance representation will be updated given any new information. The decision-maker's belief about the problem structure — the interdependence representation — serves as a lens during this process, shaping how raw data gets mapped onto the performance representation, the subjective performance landscape. Second, through the updated performance representation (and in conjunction with the current external conditions), the mental model affects which alternatives will be chosen. This choice, in turn, determines the part of the true performance landscape that will be sampled and the type of new data that will be generated for a further round of updating of the performance representation. These dual effects of mental models are captured in Figure 12. We can see how these two effects, working in tandem, are at the heart of the main results in the article. The mental model leads to biased updating of the perceived performance landscape, which leads to skewed choice patterns, which in turn, leads to skewed sampling of the world, which drives further updating, and so on. This is the core mechanism underlying the contrasting long-term exploration and performance outcomes. While these dual effects of mental models and their interaction have been documented before (e.g., Tripsas and Gavetti, 2000), we believe that our model is the first to explore it in a formal setting.

Implications for strategic management

The results above produce a number of implications when translated into the context of strategic management. We see three broad domains of implications.

Investments in exploration

Our results point to those situations in which misspecified mental models are particularly detrimental: underspecification for external interdependencies and overspecification for internal interdependencies. In both of these cases, performance is poor due to a significant lack of exploration. If the cognitive approach of managers cannot

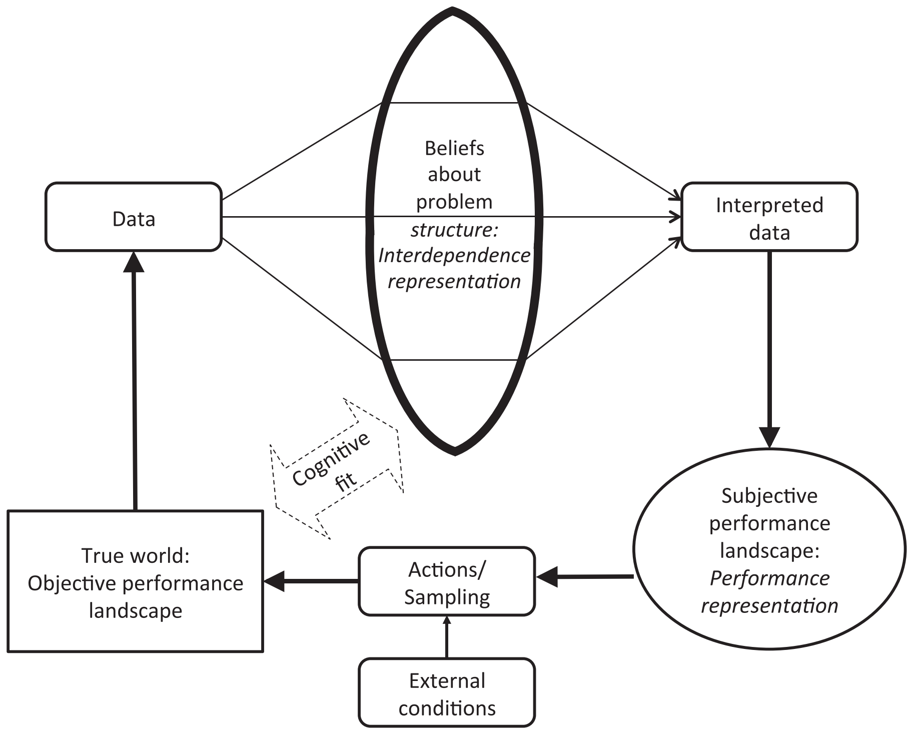

Figure 12. The dual effects of mental models. The interdependence representation works as a lens in the interpretation of incoming data. The interpreted data is then used to update the performance representation. Both external conditions and the performance representation influence which actions are being taken. These actions subsequently create new data, which is generated by the objective performance landscape. In this sense, actions determine which part of the objective performance landscape is sampled. The mental model of a decision-maker is composed of the interdependence representation and the performance representation

be changed, then the firm needs to seek other ways to increase exploration in these cases, for instance, through changes in organizational structure or by hiring external help to create additional explorative efforts.

Investments in understanding interdependencies

The asymmetric costs of misspecified mental models also have implications on how firms should allocate their resources in their attempts to better understand the true interdependencies they are facing. Firms should rather overinvest in identifying external contingencies since errors of underspecification are very costly in this case. This is particularly challenging given that managers often tend to seek evidence that a generalization is legitimate (Gavetti and Rivkin, 2007). If, however, the problem is one of internal coordination, a firm should rather invest in testing whether interdependencies really exist; here, errors of overspecification are particularly costly.

Matching cognitive approach with decision context

The third broad domain of application focuses on the locus of control over the various decisions involved. As noted earlier, in our set-up, "external" interdependencies do not have to be external to the firm. The key aspect of "external" is that it is a variable that is beyond the control of the focal decision-maker. Thus, a contingency variable could also be a different activity choice within the firm that is not under the control of the manager. Similarly, if a manager has control over a variable that is outside the boundary of the firm, that would be an "internal" interdependency. Looking at the results in Figure 11 through this lens allows us to identify a number of different organizational settings where these findings have implications of interest to strategists.

Inside the firm. We first focus on a number of settings where variations in the locus of control happen within a firm:

DOI: 10.1002/smj

Hierarchy: Consider how the span of control over decisions varies as we move up the management hierarchy from line managers to the CEO. Generally, one would expect that as one proceeds up the hierarchy, the span of the decision control increases for the manager. Thus, a larger proportion of the variables in the firm become "internal" to the manager, as per our earlier point about locus of control. In other words, as we move up the hierarchy, we are moving along the interdependence dimension of Figure 11 from external to internal. At the same time, it can be argued that the complexity of the decision tasks is increasing as well. Given these variations, we can see an interesting pattern in the consequences of over- versus underspecification of mental models as we move up the organizational hierarchy. At lower levels, when the decision environment mostly contains external interdependencies and when the task complexity is (relatively) low, underspecified mental models can be particularly damaging for performance, while overspecified ones are not only less damaging, but can, in fact, outperform correctly specified ones when complexity is very low. On the other hand, at the level of the CEO, the decision environment (with respect to the variables inside the firm) mostly contains internal interdependencies and the task complexity is (relatively) high. Here, overspecified mental models can be especially detrimental to performance, while underspecified ones can actually result in better performance compared to correctly specified ones, especially as the task complexity increases. These observations lead to direct implications on what cognitive styles would create a good fit for certain positions within the firm. While the line managers at the lower levels of the hierarchy might benefit from being complexifiers with slightly overspecified mental models, at the level of the CEO, it can be beneficial to have a generalizer, someone who tends to think in terms of cognitive simplifications, with underspecified mental models. We certainly do not want to imply that an underspecified mental model should be a primary selection criterion for CEOs. At the same time, it is however helpful to realize, for instance for the board of directors, that should the CEO have "complexifying" tendencies, stronger other mechanisms for exploration have to be employed than if the CEO tends toward being a "generalizer."

Task integration versus separation: If the cognitive style of a manager is fixed, one can ask how the task environment might be changed to better suit the cognitive style of a manager. Specifically, how might we change the mix of internal versus external interdependencies of the task to fit the individual? The implication of the results would be that if the individual is someone who tends to complexify with an overspecified mental model, it might be best to break down the task into very small pieces such that each task will have a significant number of external interdependencies (and very few, if any, internal), and are also of low complexity. On the other hand, managers with underspecified mental models should be given aggregated or integrated tasks that have high levels of internal interdependencies, as well as relatively higher levels of complexity. These implications provide a different approach to questions of optimal task design than is usually considered in the purely rational, agency theoretic perspectives (Bolton and Dewatripont, 2005).

Outside the firm. We now focus on a number of settings where the variation in the locus of control happens outside the firm:

Industry properties: Industries vary significantly in terms of inherent (irreducible) uncertainty, or the inverse, the degree of control firms have on the industry outcomes. Translating this idea to the terms of our model, one can think of industries as varying along a dimension of controllability from being very unpredictable in outcomes (with a lot of moving parts, most of which are outside the control of firms) to very docile and controllable (still with many moving parts, but with most of them being controlled by the firms), which translates to task environments that vary from high external interdependencies to high internal interdependencies. This has immediate implications for the preferred types of mental models among the strategic leadership of firms within these different industries. In industries where many variables that are vital to firm performance are outside the control of the firms, underspecified mental models can be very dangerous, while slightly overspecified ones might be beneficial. But in more sedate industries, where firms have many more levers of control over industry level outcomes, but also have high levels of complexity, the strategic leaders might be better served by having underspecified mental models, or in other words, simplified representations of their strategic environments. As with the intra-firm case, this might also be a criterion to determine which individuals are better suited for which industries, based on the fit between their cognitive style and the industry properties.

Relative firm properties: Even within a given industry, firms vary widely in terms of their size and market power, and thus, their ability to control important levers in the industry. The "gorillas" in an industry might be thought of as being in environments with high levels of internal interdependencies, while the "minnows" would be largely in high external interdependency environments. This, in turn, implies that larger, more powerful firms would want their strategic leadership to have more underspecified mental models compared to the smaller players in the industry. As before, this cognitive style is an intriguing property to consider when thinking about the fit of a CEO to a firm.

Cognitive fit

All the implications of the results that were explored above relate to a notion of "fit" between the cognitive style of the decision-makers — namely, the degree of over- or underspecification of their mental models — and some properties of the environment, be it inside the firm or outside. At a more abstract level, this could be thought of as arguing for the importance of a new type of fit, "cognitive fit," between the cognitive style of the strategic decision-maker and the decision environment (as seen in Figure 12). Cognitive fit captures the notion that performance is improved when there is a good fit between some cognitive properties of the decision-maker and some environmental properties. Whether it be the choice of a manager with a certain cognitive style to fit a certain task environment, or the alteration of the strategic environment via task integration or separation to fit the style of the manager(s), each of these examples was attempting to achieve a higher degree of cognitive fit, and through this, to improve performance. In this particular setting, the leverage for this concept is coming from the effect of over- or underspecification of mental models and its effect on exploration and performance. Cognitive fit is, however, a more general concept. Other sets of cognitive properties and environmental properties might exist such that performance is improved only under certain combinations, or "fit," between them. Hopefully, focusing attention on this concept will help open a window for future work to identify such sets of properties.

Extensions

Further extensions of the model suggest themselves readily. We outline three avenues that seem particularly worthwhile. First, as noted previously, one might allow managers to update their interdependence representations within their mental models. For instance, a manager who initially believes that the external factors 1 and 2 do not have an effect on choice 1, but then subsequently, observes that the performances achieved by alternative 1 in the presence of these two factors differ significantly, could start refining his or her mental model and treat the two factors as separate contingencies. Likewise, a manager who initially believes that two external factors have an impact on the payoff of an action and then observes very similar results may combine these two action/contingency "cells" in his or her mental model. In this extension, two key model considerations would be the trigger for splitting or merging cells in the manager's mental model and the type of information that a manager keeps track of. For instance, if the manager initially believes that weather has no impact on the payoff of an action, does he or she actually keep track of the weather? (For an interesting discussion of refining mental models, see Nelson [2008]).

A second extension, involving internal contingencies, would be to split the two actions between two separate managers. (For related work along these lines, see Puranam and Swamy, 2015.) In this extension, managers might try to optimize their own payoffs, rather than total firm payoffs as in our model. In this case, further coordination mechanisms might be necessary.

A third extension would introduce nonstationarity in the payoffs (Posen and Levinthal, 2012). Thus, managers do not operate in stable environments anymore. In this case, different updating rules for performance representations may prove to be important. The attractive feature of the model set-up that we presented is that all of these extensions are, at least from a modeling perspective, relatively straightforward. The interpretation of results, though, will certainly be less straightforward.

CONCLUSION

We would like to conclude by returning to our result that underspecification of mental models in the presence of internal interdependencies among actions can be helpful. We find this result particularly intriguing. The current work on interdependencies using NK-models tends to paint a picture of highly rugged landscapes with managers (and firms) getting stuck very quickly on low, local peaks. While a number of mechanisms, such as different organizational designs, have been described to help firms dislodge themselves from low peaks (e.g., Rivkin and Siggelkow, 2003), our results indicate that simplified managerial mental models may create endogenously helpful exploration, allowing firms to avoid getting stuck too quickly. More broadly, our study contributes to the literature that examines the mental models of management and how representations about the performance landscape may be developed. In particular, we show how mental models can act as a mediator in the process of learning about the shape of the actual performance landscape. Consequently, mental models influence how managers organize and make sense of information. Differences in these processes are likely to be important to understand how organizations in the same setting may develop different expectations about the performance consequences of particular strategies.

ACKNOWLEDGEMENTS

The authors would like to thank two anonymous reviewers and Jeff Reuer as well as participants at the Academy Meeting 2009, the EGOS conference 2009, the Organization Science Winter conference 2010, and seminar participants at the Copenhagen Business School, London Business School, University of Maryland, University of Auckland Business School, University of Southern Denmark, and the Wharton School for helpful comments. The authors also gratefully acknowledge the Mack Institute for Innovation Management at the University of Pennsylvania and the Swiss National Science Foundation (SNSF) for generous funding.

REFERENCES

Ashby WR. 1956. *An Introduction to Cybernetics*. Chapman and Hill: London, UK.

Athey S, Stern S. 1998. An empirical framework for testing theories about complementarity in organizational design. NBER Working paper No. 6600, NBER, Cambridge, MA.

Barr PS. 1998. Adapting to unfamiliar environmental events: a look at the evolution of interpretation and its role in strategic change. *Organization Science* **9**: 644–669.

Barr PS, Stimpert JL, Huff AS. 1992. Cognitive change, strategic action, and organizational renewal. *Strategic Management Journal* **13**: 15–36.

Bettis RA, Prahalad CK. 1995. The dominant logic: retrospective and extension. *Strategic Management Journal* **16**(1): 5–14.

Bolton P, Dewatripont M. 2005. *Contract Theory*. MIT Press: Cambridge, MA.

Chase WG, Simon HA. 1973. Perception in chess. *Cognitive Psychology* **4**(1): 55–81.

Chandler AD Jr. 1962. Strategy and Structure: Chapters in the History of the American Industrial Enterprise. MIT Press: Cambridge, MA.

Cho TS, Hambrick DC. 2006. Attention as the mediator between top management team characteristics and strategic change: the case of airline deregulation. *Organization Science* **17**(4): 453–469.

Conrad K. 1958. *Die Beginnende Schizophrenie. Versuch einer Gestaltanalyse des Wahns*. Thieme: Stuttgart, UK.

Denrell J, Fang C, Levinthal DA. 2004. From T-mazes to labyrinths: learning from model-based feedback. *Management Science* **50**(10): 1366–1378.

Dess G, Beard D. 1984. Dimensions of organizational task environments. *Administrative Science Quarterly* **29**: 52–73.

Donaldson L. 2001. *The Contingency Theory of Organizations*. Sage: Thousand Oaks, CA.

Edwards BD, Day EA, Arthur W, Bell ST. 2006. Relationships among team ability composition, team mental models, and team performance. *Journal of Applied Psychology* **91**: 727–736.

Eggers JP, Kaplan S. 2009. Cognition and renewal: comparing CEO and organizational effects on incumbent adaptation to technical change. *Organization Science* **20**(2): 461–477.

Einhorn HJ, Hogarth RM. 1986. Decision making under ambiguity. *Journal of Business* **59**(4): 225–250.

Eisenhardt KM, Martin JA. 2000. Dynamic capabilities: what are they? *Strategic Management Journal*Special Issue **21**: 1105–1121.

Eisenhardt KM, Sull DN. 2001. Strategy as simple rules. *Harvard Business Review* **79**(1): 107–116.

Erev I, Barron G. 2005. On adaptation, maximization and reinforcement learning among cognitive strategies. *Psychological Review* **112**: 912–931.

Fang C, Levinthal DA. 2009. The near-term liability of exploitation: exploration and exploitation in multi-stage problems. *Organization Science* **20**(3): 538–551.

Gary MS, Wood RE. 2011. Mental models, decision rules, strategies, and performance heterogeneity. *Strategic Management Journal* **32**(6): 569–594.

Gavetti G, Rivkin JW. 2007. On the origin of strategy: action and cognition over time. *Organization Science* **18**: 420–439.

Gentner D, Gentner DR. 1983. Flowing water or teeming crowds: mental models of electricity. In *Mental Models*, Gentner D, Stevens AL (eds). Erlbaum: Hillsdale, NJ; 99–130.

Gentner D, Loewenstein J, Thompson L. 2003. Learning and transfer: a general role for analogical encoding. *Journal of Educational Psychology* **95**: 393–408.

Hackner YER. 1991. Integrated complexity and profitability. Working paper, Case Western Reserve University, Cleveland, OH. Presented at the Academy of Management Conference, Miami, FL.

Heath C, Staudenmayer N. 2000. Coordination neglect: how lay theories of organizing complicate coordination in organizations. *Research in Organizational Behavior* **22**: 155–193.

Hogarth RM, Kolev GI. (2013). The Ombudsman: The "Wicked" Environment of CEO Pay. *Interfaces*, **43**(6), 596–598.

Jaikumar R. 1986. Postindustrial manufacturing. *Harvard Business Review* **64**(6): 69–76.

Lawrence PR, Lorsch JW. 1967. *Organization and Environment*. Harvard Business School Press: Boston, MA.

Levinthal DA, March JG. 1981. A model of adaptive organizational search. *Journal of Economic Behavior and Organization* **2**(4): 307–333.

Levinthal DA, March JG. 1993. The myopia of learning. *Strategic Management Journal* **14**: 95–112.

Levitt B, March JG. 1988. Organizational learning. *Annual Review of Sociology* **14**: 319–340.

March JG, Sproull LS, Tamuz M. 1991. Learning from samples of one and fewer. *Organization Science* **2**: 1–13.

McCloskey M, Caramazza A, Green B. 1980. Curvilinear motion in the absence of external forces: naive beliefs about the motion of objects. *Science* **210**(4474): 1139–1141.

Milgrom PR, Roberts J. 1995. Complementarities and fit: strategy, structure, and organizational change in manufacturing. *Journal of Accounting and Economics* **19**: 179–208.

Miller D. 1992. Environmental fit versus internal fit. *Organization Science* **3**: 159–178.

Miller D, Friesen P. 1984. *Organizations: A Quantum View*. Prentice-Hall: Englewood Cliffs, NJ.

Mintzberg H. 1973. *The Nature of Managerial Work*. Harper and Row: New York, NY.

Nair KU. 2001. Adaptation to creation: progress of organizational learning and increasing complexity of learning systems. *Systems Research and Behavioral Science* **18**(6): 505–521.

Nelson RR. 2008. Bounded rationality, cognitive maps, and trial and error learning. *Journal of Economic Behavior and Organization* **67**(1): 78–89.

Pettigrew AM, Woodman RW, Cameron KS. 2001. Studying organizational change and development: challenges for future research. *Academy of Management Journal* **44**: 697–713.

Porac JF, Thomas H, Baden-Fuller C. 1989. Competitive groups as cognitive communities: the case of Scottish knitwear manufacturers. *Journal of Management Studies* **26**(4): 397–416.

Porac JF, Thomas H, Wilson F, Paton D, Kanfer A. 1995. Rivalry and the industry model of Scottish knitwear production. *Administrative Science Quarterly* **40**(2): 203–230.

Porter ME. 1996. What is strategy? *Harvard Business Review* **74**(6): 61–78.

Porter ME, Siggelkow N. 2008. Contextuality within activity systems and the sustainability of competitive advantage. *Academy of Management Perspectives* **22**: 34–56.

Posen H, Levinthal DA. 2012. Chasing a moving target: learning in dynamic environments. *Management Science* **58**(3): 587–601.

Puranam P., Swamy M. 2015. How initial representations shape coupled learning processes. *Organization Science*. Forthcoming.

Rivkin JW, Siggelkow N. 2003. Balancing search and stability: interdependencies among elements of organizational design. *Management Science* **49**: 290–311.

Schwenk C. 1984. Cognitive simplification processes in strategic decision making. *Strategic Management Journal* **5**: 111–128.

Shermer M. 2000. *How We Believe*. W.H. Freeman: New York, NY.

Siggelkow N. 2001. Change in the presence of fit: the rise, the fall, and the renaissance of Liz Claiborne. *Academy of Management Journal* **44**: 838–857.

Siggelkow N. 2002. Evolution toward fit. *Administrative Science Quarterly* **47**: 125–159.

Siggelkow N, Rivkin JW. 2005. Speed and search: designing organizations for turbulence and complexity. *Organization Science* **16**: 101–122.

Simon HA. 1962. The architecture of complexity. *Proceedings of the American Philosophical Society* **106**: 467–482.

Simon HA. 1991. *Models of My Life*. MIT Press: Cambridge, MA.

Sutcliffe KM. 1994. What executives notice: accurate perceptions in top management teams. *Academy of Management Journal* **37**(5): 1360–1378.

Sutton RS, Barto ML. 1998. *Reinforcement Learning: An Introduction*. MIT Press: Cambridge, MA.

Thompson JD. 1967. *Organizations in Action*. McGraw-Hill: New York, NY.

Tripsas M, Gavetti G. 2000. Capabilities, cognition, and inertia: evidence from digital imaging. *Strategic Management Journal* **21**(10/11): 1147–1161.

Walsh JP. 1995. Managerial and organizational cognition: notes from a trip down memory lane. *Organization Science* **6**: 280–321.

Webber SS, Chen G, Payne SC, Marsh SM, Zaccaro SJ. 2000. Enhancing team mental model measurement with performance appraisal practices. *Organizational Research Methods* **3**: 307–322.

Weick KE. 1979. *The Social Psychology of Organizing*. Addison-Wesley Publishing Company: Reading, MA.

Weick KE. 1989. Mental models of high reliability systems. *Industrial Crisis Quarterly* **3**: 127–142.

Zajac EJ, Bazerman MH. 1991. Blind spots in industry competitor analysis: implications of interfirm (mis)perceptions for strategic decisions. *Academy of Management Review* **16**(1): 37–56.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article:

Appendix S1. Relationship to prior work using *n*-armed bandit models

Appendix S2. How mental model overspecification can create exploration