A Contingency Theory of Representational Complexity in Organizations*

Felipe A. Csaszar
Ross School of Business, U. of Michigan
Ann Arbor, MI 48109
fcsaszar@umich.edu

James Ostler
The Wharton School, U. of Pennsylvania
Philadelphia, PA 19104
jostler@wharton.upenn.edu

October 25, 2019

Forthcoming in *Organization Science*

*The authors thank Gautam Ahuja, Jeho Lee, and Saerom Lee for helpful comments. The authors also thank the seminar participants at the University of Michigan, Seoul National University, the Organization Science Winter Conference 2017, the Utah/BYU Winter Strategy Conference, INSEAD, Purdue University, and New York University. Any errors are the authors' own responsibility.

A Contingency Theory of
Representational Complexity in Organizations

Abstract

A long-standing question in the organizations literature is whether firms are better-off by using simple or complex representations of their task environment. We address this question by developing a formal model of how firm performance depends on the process by which firms learn and use representations. Building on ideas from cognitive science, our model conceptualizes this process in terms of how firms construct a representation of the environment and then use that representation when making decisions. Our model identifies the optimal level of representational complexity as a function of (a) the environment's complexity and uncertainty and (b) the firm's experience and knowledge about the environment's deep structure. We use this model to delineate the conditions under which firms should use simple versus complex representations; in doing so, we provide a coherent framework that integrates previous conflicting results on which type of representation leaves firms better-off. Among other results, we show that the optimal representational complexity generally depends more on the firm's knowledge about the environment than it does on the environment's actual complexity. We also show that the relative advantage of heuristics vis-à-vis more complex representations critically depends on an unstated assumption of "informedness": that managers can know what are the most relevant variables to pay attention to. We show that when this assumption does not hold, complex representations are usually better than simpler ones.

Keywords: representation; complexity; heuristics

1 Introduction

1.1 The simple-versus-complex debate in the organizations literature

A foundational idea in the behavioral view of the firm is that firms make decisions based on a *representation* of the problems they face (see Csaszar 2018). A key characteristic of such representations is how complex they are—what we denote *representational complexity.* Firms have much latitude on how complex are the representations they use: firms can choose to make decisions on the basis of representations ranging from the very simple (e.g., by employing a manager who relies on a rule of thumb) to the overly complex (e.g., by hiring technical experts and using sophisticated decision-making models). Interestingly, there are conflicting views on what is the optimal representational complexity; that is, whether firms are better-off by using simple or more complex representations.

On the one hand, the literature on fast-and-frugal heuristics (Gigerenzer and Goldstein 1996) has argued that simple representations have several advantages over more complex representations, such as requiring less effort to learn and use, and being robust under a broad range of environments. In the context of organizations, the call for simple representations has found empirical support in research that shows that firms using simple decision processes (Fredrickson and Mitchell 1984) and simple rules (Bingham and Eisenhardt 2011) outperform similar firms using more complex decision processes and rules.

On the other hand, a different literature has argued that firms can improve their performance by using more complex representations, as these are better able to deal with the nuances of the business environment. Weick (1979:261) typifies this literature in exhorting managers to "complicate yourself!" This line of argumentation has also found empirical support in the context of organizations. For instance, McNamara et al. (2002) study 76 banks and show a positive relation between firm performance and the complexity of the representations managers use when categorizing their competitors.

The logic underlying the call for complex representations is that business environments are complex and, as argued by Ashby's (1956, chap. 11) law of requisite variety, responding successfully to a complex environment requires an equally complex system. Thus Ashby proposes an in-between, contextual recommendation: firms should use representations whose complexity matches the complexity of the environment. Yet Ashby's recommendation has not found much empirical support in the organizations literature, which explains why the debate over simple versus complex representations has persisted.

The fact that both the research on simple representations and on complex representations have empirical support suggests that both views are correct under certain situations. Hence the optimal complexity of representations may depend on particular contingencies, which means that simple representations, complex representations, and perhaps even representations of medium complexity may be preferable under the right circumstances. However, the organizations literature does not currently offer a theory capable of describing what these contingencies might be and how they co-determine the optimal complexity of representations. Our aim in this paper is to develop such a contingency theory.

1.2 Our approach

To develop such a theory, we present a parsimonious mathematical model of how firm performance is affected by the complexity of the representations it uses. Our model combines three ideas rooted in cognitive science: (i) understanding organizations in terms of the representations they use, (ii) modeling representations and their environment using Brunswik's (1952) lens model, and (iii) analyzing such model using ideas from statistical learning theory (the bias/variance trade-off and related ideas; Geman et al. 1992). The combination of these ideas allows us to examine in a unified manner theories that currently are disconnected. Under this unified view, theories advocating for different levels of complexity correspond to particular cases of a more general theory of representational complexity in organizations.

An advantage of our approach is that it forces us to make explicit the implicit assumptions made by previous work. This increase in theoretical precision allows us to show that the different predictions of prior work stem from having conceptualized two key constructs—uncertainty and knowledge—differently. Previous research has interpreted the term "uncertainty" either as intrinsic unpredictability of the environment or as lack of knowledge about the environment (a.k.a. aleatory and epistemic uncertainty, respectively; Fox and Ülkümen 2011). Similarly, "knowledge" has been interpreted either as experience (i.e., having seen many observations) and as *informedness* (i.e., knowing what aspects matter the most). Our formal modeling approach shows that distinguishing between these different meanings is paramount, as they call for markedly different organizational responses. For instance, while lack of experience typically calls for simpler representations, lack of informedness calls for more complex representations.

The model we develop shows how firm performance is affected by the complexity of the representations it uses. Our model takes into account contingencies that are particularly relevant in organizational settings. These contingencies correspond to two key characteristics of the environment—its complexity and (aleatory) uncertainty—and to two key characteristics of the decision makers: their experience and informedness. In doing so, our model captures Simon's (1990:7) idea that organizational behavior is jointly shaped by "scissors" whose two blades are the environment's structure and the decision maker's cognitive capabilities.

We analyze the model by building on research on the bias/variance trade-off (Geman et al. 1992) and related explanations on why heuristics work (Gigerenzer and Gaissmaier 2011). Our work differs from this previous line of research by (i) exploring the effect of contingencies that matter in the context of organizations (i.e., the research building on bias/variance has focused on the effect of uncertainty and experience, to which we add the effect of complexity and informedness), (ii) exploring the full range representational complexity (i.e., research on heuristics typically focuses on simple representations), and (iii) illuminating questions in the organizations literature (e.g., the simple-versus-complex debate and the emerging research on how representations affect performance). Our approach is consistent with Vuori and Vouri's (2014) call to contextualize organizational decision-making models; in particular, to acknowledge that organizations face environments and employ decision makers that are different than what is typically assumed by the heuristics literature.

1.3 Our contribution

Our study contributes to the organizations literature along several fronts. First, we show that three theories in the organizations literature that speak to the issue of representational complexity—calling respectively for simple representations (i.e., Gigerenzer and Goldstein 1996, Bingham and Eisenhardt 2011), complex representations (i.e., Weick 1979), and complexity-matching representations (i.e., representations that match the complexity of the environment; Ashby 1956)—can be integrated into a coherent framework.

Second, in terms of specific results, we show that complex and complexity-matching representations lead to superior performance in one case each: complex representations are most valuable when managers are uninformed (i.e., when they do not have a priori knowledge about what matters most), irrespective of how experienced they are; and complexity-matching representations are most valuable when managers are both informed and experienced. In contrast, simple representations lead to superior performance under any of three circumstances: high uncertainty, informed but inexperienced managers, or high costs of learning and using complex representations. These results bear a number of managerial implications. We elaborate on these by discussing the types of managers, frameworks, and organizational structures that are likely to display the most adequate representational complexity in different environments.

Third, our work qualifies popular maxims about how complex should representations be. We show that, for managers tasked with making predictions, both Occam's razor and Ashby's law of requisite variety are not absolute truths but only useful under specific conditions. Paradoxically, managers who know little (i.e., have little experience and informedness) can be better off by using quite complex representations—being "anti-Occam" managers. The mechanism underlying our results is a relevance–accuracy trade-off: more complex representations are more likely to include all relevant aspects, yet they are also less able to accurately reflect the import of any given aspect. Which of these two elements matters the most depends on the situation; that is, the optimal representational complexity depends on the rate at which accuracy and relevance trade-off one another, which in turn depends in nuanced ways on the characteristics of the environment and the decision maker we study. This trade-off is an unavoidable consequence of learning a representation from limited and noisy observations.

Fourth, our work also contributes to the organization design literature by showing that cognitive contingencies (experience and informedness) can have a greater impact than the classic environmental contingencies (uncertainty and complexity) studied by this literature. For instance, uncertainty is the main contingency in the literature on simple rules (see, e.g., Davis et al. 2009) and environmental complexity is the main contingency in the literature on organizational adaptation (see, e.g., Levinthal 1997); our work suggests that both lines of research could significantly increase their explanatory power by incorporating cognitive contingencies.

Finally, our work contributes to better understanding how representations affect performance. The research on managerial cognition has developed a rich understanding of representational heterogeneity while remaining mostly silent on how representations affect performance (Gary and Wood 2011:570). By studying how representational complexity—a key characteristic of all representations—affects the quality of decisions made by organizations, this paper furthers the emerging research on the representation–performance link (Csaszar and Levinthal 2016, Csaszar and Laureiro-Martinez 2018).

The paper is structured as follows. The next section lays out a theoretical foundation for studying formally the complexity of representations in organizations. The subsequent section describes our model. We then present the results that emerge from this model and, building on those results, discuss the broader theoretical and managerial implications of our research.

2 Theoretical background

The concept of representation underlies our theoretical approach. To develop that theoretical foundation, this section proceeds in four steps: (i) it points out how three types of representations—internal, external, and distributed—affect decision making in organizations, (ii) it explains the Brunswikian approach to modeling representations, (iii) it explains how ideas from statistical learning theory can be used to analyze the performance of representations, and (iv) it shows how such ideas have been used by the heuristics literature and how the current work differs.

2.1 The role of representations in organization theory

A *representation* is a model that can be used to generate predictions (Craik 1943:61, Holland et al. 1986:12). The concept of representation plays a central role in cognitive science, whose central premise is that thinking can be best understood in terms of representations and the computational procedures operating on those representations (Thagard 2005:10–12).[1]

The concept of representation is also central to the organizations literature.[2] This centrality stems from the fundamental observation that managers must deal with problems whose complexity and multidimensional nature exceed managers' cognitive capabilities, from which it follows that dealing with such problems requires managers to use smaller-scale representations (Simon 1957:198–199). The key role of representations in both cognitive science and organization theory can be traced back to the seminal work of Simon in both fields (see Spender 2013 for an account of Simon's dual influence).

Given that the problems faced by organizations are typically much larger than their managers' ability to represent them, most problems can be represented in multiple ways. In fact, research on managerial cognition has documented a vast heterogeneity among managers' representations. Managers can differ in how they represent almost any aspect of business, including market uncertainty (Milliken 1990), competitors (Porac et al. 1989), product features (Benner and Tripsas 2012), technological

[1] The central premise of cognitive science is summarized by the maxim "thinking = representations + computation" (Thagard 2005:11), which mirrors the computer science maxim "programs = data structures + algorithms" (Wirth 1976).

[2] This literature has referred to representations under different labels, such as: schema, knowledge structure, mental model, cognitive map, dominant logic, interpretive scheme, thought world, and managerial lens. For an exhaustive list, see Table 1 in Walsh (1995).

opportunities (Eggers and Kaplan 2009), employees (DeNisi et al. 1984), and power relationships (Krackhardt 1990). To appreciate the pervasiveness of representations in organizations it is useful to distinguish between three types of representations: internal, external, and distributed. Distinguishing among these is also useful, as illuminates the levers that managers can use to change an organization's representations.

Internal representations. Internal representations are those help in the mind of an individual (and so are sometimes called mental representations). Most research in cognitive science and organization theory studies this type of representation. For instance, a look at the tables of contents of cognitive science textbooks (see, e.g., Lindsay and Norman 1977, Thagard 2005) reveals that most are organized around the internal representations that humans are postulated to use for accomplishing different tasks (e.g., representations used for vision, speech, reasoning). The organizations literature has similarly focused on internal representations. Examples include Porac et al. (1989), who study the internal representations used by managers in the Scottish knitwear industry; the chapters in Huff (1990) that describe various methods to map managers' internal representations; and all papers cited in the previous paragraph, which examine managers' internal representations. However, internal representations are not the only type of representation used by individuals. The last 20 years have seen growing interest in the study of representations that do not reside in an individual's mind. These other types of representations—called *external* and *distributed* representations—are commonly used by individuals and organizations.

External representations. External representations are those that are embedded in a given physical artifact (Norman 1991). External representations play a part in many cognitive tasks—as when one performs long division using paper and pencil, keeps a written "to do" list, uses a map to find a route, or analyzes data by graphing them (Zhang 1997). In cognitive science, research on external representations is relatively recent (see Kirsh 2010 for an overview) and much less developed than is the research on internal representations.

Although not usually acknowledged in this way, external representations are widely used in organizations—especially in the context of strategic decision making, where common artifacts include such frameworks as the BCG matrix, Porter's "five forces," and myriad 2×2 decision matrices. External representations are not only used in the context of strategic decisions, but also in operational and tactical decisions, such as when doctors and pilots use checklists, a clerk follows a written operating procedure, or a hiring officer sorts through candidates by comparing fields in a spreadsheet. Although the use of external representations is common in organizations, little research has explored how organizations are affected by this type of representation (for a notable exception, see Kaplan's 2011 analysis of how PowerPoint affects strategic decision making).

Distributed representations. Distributed representations are those that are spread over multiple individuals and possibly artifacts. For instance, the information required for an airliner's safe landing is distributed over the pilot's mind, the co-pilot's mind, and the dashboard instruments. No single one of the three parts in this system has all the information necessary for landing (Hutchins 1995b). Instead, it is the cockpit as a whole—the system—whose representation incorporates the necessary information. As with external representations, distributed representations are pervasive in organizations. For instance, they occur when the information needed to design a new product is spread over managers from different departments or when the information required to predict the next quarter's earnings is distributed over multiple employees and accounting systems.[3]

Research on distributed representation is an emerging area in cognitive science (for an overview, see Robbins and Aydede 2009). Because organizations are the quintessential holders of distributed representations, much of the organizations literature can be understood as implicitly examining issues of distributed representation. For instance, research on transactive memory (Wegner 1986), interpretive barriers (Dougherty 1992), organization design (Siggelkow and Rivkin 2005), and information aggregation (Csaszar and Eggers 2013) all explore core distributed representation issues (e.g., "who knows what" and "how is knowledge communicated").

In sum, representations are deemed internal, external, or distributed depending on the system that holds them. For internal representations, this system is the manager's mind; for external representations, the system is the manager plus the set of physical artifacts being manipulated. For distributed representations, the system is the set of managers (and possibly artifacts) that are part of the decision-making process.

[3] Some authors group external and distributed representations together—as, for example, in the research on situated cognition (Robbins and Aydede 2009) and socially distributed cognition (Hutchins 1995a). It is also worth mentioning that our use of the term "distributed representation" should not be confused with its use in the neural networks literature, where it denotes the type of sparse memory system implemented by those networks (Hinton et al. 1986).

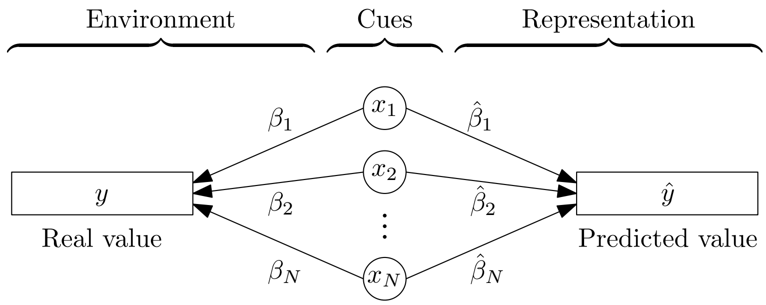

Figure 1: Brunswik's lens model.

2.2 A Brunswikian understanding of representations

So far we have shown how representations are at the core of how individuals and organizations make decisions. We now move on to explain Brunswik's (1952:16–21) lens model—a framework that formalizes the concept of representation, which allows us to build our model. Much research on decision making is based on Brunswik's lens model (see, e.g., Karelaia and Hogarth 2008 survey of 249 lens-model studies). The task captured by Brunswik's model is predicting a value when presented with cues or characteristics of the environment.

Brunswik's model conceptualizes the environment and its representation in a symmetric fashion: the environment is a function that connects cues to a *real* value; and the representation is a function that connects those cues to a *predicted* value. These relationships are usually depicted as shown in Figure 1 and can be illustrated by the following toy example. The profits of a software company evaluating which smartphone "app" to develop may depend on the app's market size (x_1), the number of competing apps (x_2), and the difficulty of development (x_3) according to the expression $y = 0.8x_1 - 0.4x_2 - 0.2x_3$. However, the firm's CEO may believe that profits vary according to some *other* relationship. Perhaps she underestimates the effect of competing apps and does not consider difficulty of development, believing that profits are instead given by $\hat{y} = 0.8x_1 - 0.3x_2$. The first equation (describing y) corresponds to the environment, and the second equation (describing $\hat{y}$) corresponds to the CEO's representation of it.[4]

The lens model can be used to describe internal, external, or distributed representations. For instance, $\hat{y} = x_1 + x_2$ could correspond to a manager who pays equal heed to two cues (i.e., an internal representation), a framework that weighs two aspects equally (an external representation), or an

[4]It is customary to draw *all* the arrows in the lens model's graphical depiction, but this does not imply that every cue affecting the environment also affects the individual. Cues that affect one side but not the other have their corresponding β or $\hat{\beta}$ set to 0. It is also customary to illustrate the environment and representation functions with linear functions, but these functions could have any functional form (in fact, Brunswik 1952 does not assume a functional form; empirical work following Hursch et al. 1964 has typically used linear forms, but for an exception see Hamm and Yang 2017).

organization that gives equal say to two managers (a distributed representation).

Under Brunswik's framework, the performance of a given representation in a given environment depends on a function that relates y to $\hat{y}$. Suppose, for instance, that the CEO in our previous example was evaluated according to her ability to predict what will happen; in that case, her performance would be a function of the (absolute) difference between her prediction and what actually occurred (i.e., $|\hat{y} - y|$). A more realistic measure of performance would reflect her contribution to profits. So if the CEO decided that the firm should enter a market only when positive profits are expected (i.e., when $\hat{y} > 0$), then performance would be equal to 0 under no entry or equal to y (the actual profit or loss) under entry (such contribution to performance can be written as $y\mathbb{1}[\hat{y} > 0]$).

The literatures on adaptive decision making (Payne et al. 1993) and "fast and frugal" heuristics (Gigerenzer and Gaissmaier 2011) both rely heavily on the Brunswik lens model, by using it as a framework within which to study the effectiveness of various heuristics (in this context, a heuristic is just a simple representation, one that depends on few x's). For instance, the "single variable" heuristic (Hogarth and Karelaia 2007) corresponds to a representation in which the coefficient of the most relevant cue is set to 1 and all the other coefficients are set to 0.

2.3 Using statistical learning theory to study representations

Apart from providing a formal understanding of representations and of their interplay with the environment (i.e., Simon's scissors), a critical benefit of using Brunswik's framework to study representations is that it allows borrowing ideas from statistical learning theory to understand what drives the effectiveness of different representations. In a nutshell, because the quality of a representation depends on how well it approximates the environment (i.e., how closely $\hat{y}$ matches y), one can study representations with the same tools used to understand the problem of learning a function from data.

The canonical problem of learning a function from data is this: given some noisy observations, what is the best estimate of the data-generating process that produced such observations? For instance, if one has seen the black dots in Figure 2(a), is it better to infer that the data was produced by a polynomial of degree 1, 2, or 20? (shown as colored curves in panels (a), (b), and (c), respectively). As the degree of the polynomial increases, the curves fit the data better. But that does not mean that the best estimate of the data-generating process is the polynomial with the highest degree. Although such polynomial would fit past observations very well, it would probably not fit additional observations

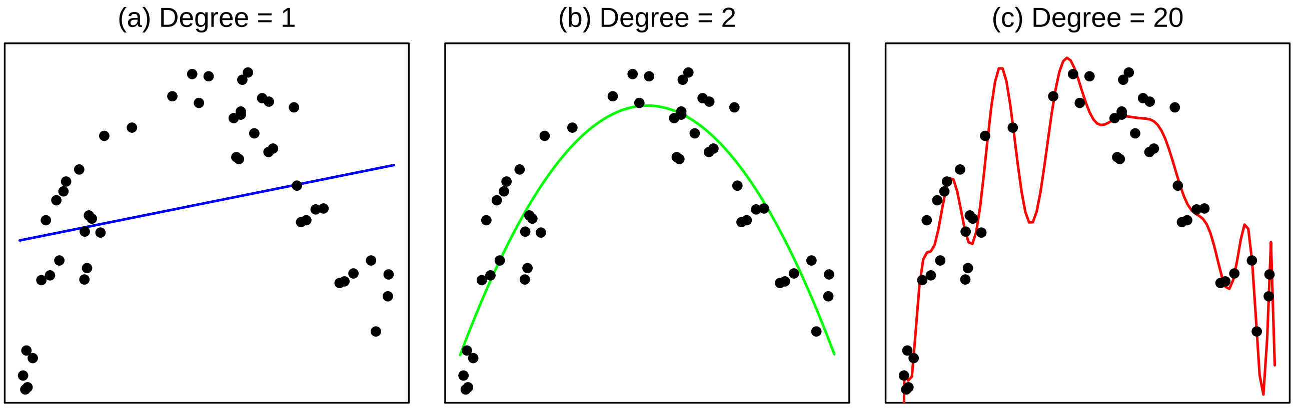

Figure 2: Approximating a data-generating process with successively more complex polynomials.

appropriately. In this example, the actual data-generating process is a second degree polynomial plus noise ($y = -x^2 + x + \varepsilon$). So, using a polynomial of a higher degree *overfits* the data—is good at fitting the noise but not the underlying data-generating process.

Statistical learning theory has analyzed the problem of overfitting in terms of two sources of error: variance and bias (Geman et al. 1992). In a dartboard analogy, two ways in which a dart thrower could miss the bulls-eye are: (a) by consistently hitting one wrong spot (low variance, high bias) and (b) by hitting random points centered around the bulls-eye (high variance, low bias). Both bias and variance contribute to the total error.

In the context of learning a model from data, there is a trade-off between these two sources of error. If one repeatedly estimates a model using different samples and looks at the prediction error for a *given* new point, the following happens. Very complex models will typically have low bias (as they will not exhibit any consistent error when estimating the new point) but will have high variance (as sometimes they will learn random things, such as the big zigzags in Figure 2(c)). In contrast, very simple models will have high bias (as they will consistently learn the wrong pattern; for example, linear models facing data like the one in Figure 2(a) will consistently underestimate the topmost point) but will have low variance (because they will not be as dependent on the exact sample used to estimate the model and, hence, the error when estimating a given point will be relatively stable). The bias/variance trade-off applies to any problem of learning a function from data—not just polynomials, but any learning system that has a tunable degree of model complexity.[5] In sum, the bias/variance trade-off establishes that

[5]It is out of the scope of this paper to provide a detailed exposition of the bias/variance trade-off. The seminal description of this trade-off is Geman et al. (1992). Clear expositions of it (listed in increasing level of detail) appear in Domingos (2012), Alpaydin (2014), and Hastie et al. (2009).

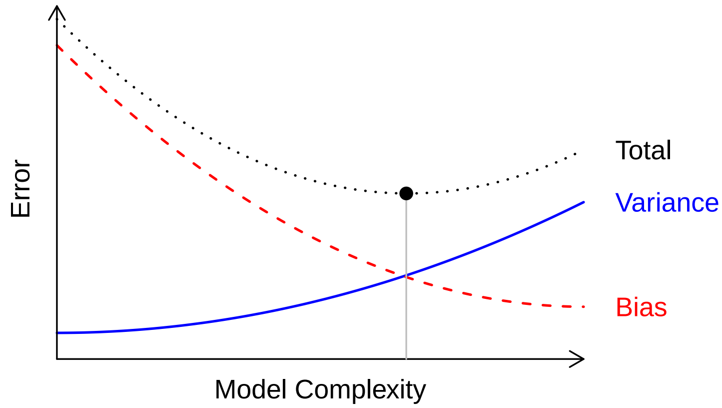

Figure 3: Illustration of the bias/variance trade-off.

increasing model complexity, decreases bias but increases variance (see Figure 3).

The machine learning literature has extensively used the bias/variance trade-off as a means to understand the effect of increasing model complexity (see, e.g., Hastie et al. 2009, Alpaydin 2014). But note that the bias/variance trade-off just establishes what its name says—that there is a trade-off between these two sources of error—however, it says nothing about what is the optimal model complexity under specific conditions. In terms of Figure 3, the optimal model complexity (marked by a black dot) could fall *anywhere* depending on the specific shape of the bias and variance curves. For example, if the variance curve grew at a slower rate, the optimal would move to the right; and if the bias curve was almost flat, the optimal would move to the left. In general, the optimal model complexity depends on specifics of the problem: what is the sample size, what is the data-generating process, and what is the family of functions used to approximate it.

The bias/variance trade-off is a qualitative relationship that applies to any learning system. In some situations it is possible to quantify the trade-off using what is called a bias/variance decomposition. Such decomposition is a mathematical formula that splits the total error into bias, variance, and irreducible error components. This decomposition, however, is only well-established for prediction problems that use a square loss (i.e., that minimize $(y - \hat{y})^2$). There are no well-established decompositions for other types of problems (e.g., classification) and loss functions (e.g., absolute error, 0/1 error, asymmetric errors). For attempts at developing bias/variance decompositions for other cases—and the problems of doing so—see Domingos (2000), James (2003), and references therein. Using a bias/variance decomposition is further complicated because such decompositions keep the data-generating process and the point being estimated fixed (in this paper we vary both elements, which correspond to varying environments and incoming projects). For these reasons, many works (including this paper) look at measures that are closely related to bias and variance, but that are not a bias/variance decomposition.[6]

2.4 The heuristics literature

The heuristics literature has used the bias/variance trade-off to explain when heuristics work well. For example, in a survey of the heuristics literature, Gigerenzer and Gaissmaier (2011:459) explain that "both too few and too many parameters can hurt performance" (because, as shown in Figure 3, the total error follows a U-shape). The heuristics literature has also invoked the bias/variance trade-off to argue that more experience (i.e., more observations) calls for more complex models and that higher uncertainty (i.e., more noise in the environment) calls for simpler models. Or, as summarized by Artinger et al. (2015:37), "the larger the sample size and the smaller the noise, the better the complex decision strategies perform." From this logic it follows that heuristics are most useful in novel and uncertain situations.

Most research on heuristics has focused on describing different heuristics and determining their relative performance in various environments. This literature has its roots in work showing that the predictive accuracy of "improper" linear models (i.e., models that do not use optimally derived coefficients but instead use coefficients that are noisy or severely constrained) is surprisingly similar to the performance of optimally-derived models (Dawes and Corrigan 1974, Einhorn and Hogarth 1975, Dawes 1979). The ensuing literature on heuristics studied simpler (i.e., even more "improper") decision rules, including: tallying pros and cons, deciding based on a single variable, familiarity-based rules (e.g., recognition, fluency, and similarity heuristics) as well as lexicographic rules (such as take-the-best, which sequentially adds cues in order of their validity until no ties are found).

This research has established conditions under which some heuristics are preferable over others by benchmarking different heuristics in real-world environments (see, e.g., Hogarth and Karelaia 2007, Martignon and Hoffrage 2002) and by developing analytical results (see, e.g., Katsikopoulos and Martignon 2006). For example, research on take-the-best has shown that it outperforms other lexicographic models in 20 real-world datasets (Martignon and Hoffrage 2002:46), that it is more robust to changes in cue correlation and validity than the equal weights and single variable rules (Hogarth and

[6] In fact, even if Hastie et al. 2009 spend many pages developing the mathematics of the standard bias/variance decomposition, they arguably only use this concept in a qualitative manner, as in the context of their figures 2.4, 2.11, 7.1 and 7.2.

Karelaia 2007:49), and that it outperforms tallying if cues are conditionally independent (Katsikopoulos and Martignon 2006:491). In their survey of the heuristics literature, Gigerenzer and Gaissmaier (2011:474–475) note that the research that has studied the applicability of different heuristics to different environments is fragmented and incomplete, as currently there is no systematic theory to describe heuristics and the environmental structures they exploit.

Although we draw inspiration from the heuristics literature, the current paper differs from it in three main ways. First, we add two contingencies that are relevant to the organizations literature: environmental complexity and informedness. Including environmental complexity is important, as this a key characteristic of environments as studied by the organizations literature (see, e.g., Dess and Beard 1984, Burton and Obel 2004, Levinthal 1997). Similarly, including informedness (in addition to experience), acknowledges the two main modes of learning—vicarious and experiential—studied by the organizational learning literature (see, e.g., Denrell 2003, Posen and Chen 2013, Argote 2013). Second, while the heuristics literature typically compares the performance of simple decision rules, we explore the whole range of representational complexity (i.e., from minimal to maximal representational complexity rather than focusing on heuristics, which are clustered at the low end of the spectrum of representational complexity). Third, as mentioned in the Introduction, we use our model to illuminate questions that are relevant to the organizations literature and that have not been studied by the heuristics literature (i.e., the simple-versus-complex debate, the representation–performance link, the conflicting interpretations and effects of uncertainty, and the relative importance of cognitive vis-à-vis environmental characteristics).

In sum, this section has shown how representations underlie decision making in organizations, that Brunswik's lens model can be used to conceptualize representations and their relationship with the environment, and that ideas from statistical learning theory and the heuristics literature can be used to understand the effectiveness of representations. Building on these concepts, the next section develops a formal model that sheds light on the optimal complexity of representations as a function of characteristics of the environment and the firm that are relevant in organizational settings.

3 Model

Our model describes a firm[7] that has to *screen* projects; that is, approve good projects and reject the rest. The firm infers a representation of the environment by observing past projects; it then uses that representation to predict the profits of new projects, approving those with positive predicted profits. Finally, firm performance depends on how profitable the approved projects actually are.[8]

The model is used to study how firm performance depends on the complexity of the representation used, contingent on characteristics of the environment and the firm. The contingencies that characterize the environment are its complexity K (i.e., the extent to which multiple, interacting cues matter) and its uncertainty U (the extent to which the environment exhibits unpredictable variation). The contingencies describing the firm are experience E (measured as the number of past projects observed) and informedness I (i.e., whether the decision maker already knows which aspects of the environment matter the most or, instead, must discover those aspects "from scratch"). The model's main independent variable is the representation's complexity, denoted K'. Thus our model investigates the function Performance$(K'; K, U, E, I)$.

The rest of this section describes more formally the elements just outlined. Toward that end, the presentation is divided into three parts: (i) the environment, (ii) the representation, and (iii) how performance emerges from the interplay between environment and representation.

3.1 The environment

The environment determines the types of projects that the firm ends up screening. We follow Brunswik in supposing that each project is described by a set of cues (denoted $x_1, \ldots, x_M$) and an outcome (denoted y). Take, for instance, the context of a software company evaluating which smartphone apps to develop; in this case, the cues are the apps' attributes (e.g., market size, target platform, development costs, number of competing apps) and the project outcomes are the profits or losses accruing to the firm if it developed those apps.

Following Brunswik, we conceptualize different environments as different functions that produce

[7]This model views the firm as a unitary decision maker; hence we use the terms "firm", "manager", and "decision maker" interchangeably depending on explanatory convenience.

[8]Screening projects is a common task in many settings, such as deciding whether to hire an employee, acquire a firm, or launch a new product. The firm in our model is assumed to screen a stream of exogenously given projects, so this is not a model of search (à la Levinthal 1997) but rather a model of screening (along the lines of those developed by Sah and Stiglitz 1986 and Csaszar 2013).

the outcome y. Rather than studying some arbitrary environment (described by, say, $y = 2x_1 + x_2$), we study environments that vary in their complexity and uncertainty—which are the main environmental contingencies highlighted in the organizations literature (see, e.g., Dess and Beard 1984 as well as Burton and Obel 2004, chap. 6).

Both complexity and uncertainty can be expressed in mathematical terms that conform with Brunswik's framework. The complexity of a system, as defined by Simon (1962:468), increases with the number of parts and interactions in the system. Hence, for example, one can say that environment $y = 6x_1$ is less complex than environment $y = 6x_1 + 4x_2 + 2x_1x_2$ because the former has fewer parts (i.e., cues) and interactions than does the latter.

In turn, uncertainty corresponds to the extent to which an environment exhibits unpredictable variation (see Priem et al. 2002:725–27 and references therein). Translated into Brunswikian terms, this characteristic is the extent to which random noise affects the environment function—for example, environment $y = x_1 + \varepsilon$, where $\varepsilon \sim \text{Normal}(0, \sigma)$ becomes more uncertain as σ increases.

Our model creates environments with given levels of complexity (K) and uncertainty (U) via the following procedure. Starting from a multilinear polynomial on M variables (this includes one intercept $+\, M$ main effects $x_1, \ldots, x_M +$ all two-way interactions, totaling $1 + M + \frac{M(M-1)}{2}$ terms), we eliminate all *except* K randomly chosen terms. We then set all the polynomial's coefficients (i.e., its β values) to random values drawn from a $\text{Normal}(0, 1)$ distribution (to make results independent of any specific random draw, the reported results are averaged over many simulations). Finally, we multiply the resulting polynomial by $1 + \varepsilon$, where $\varepsilon \sim \text{Normal}(0, U)$.[9] Note that the β's remain constant within a given environment whereas ε is a random variable drawn for each new project in that environment.

In sum, the described procedure: (i) creates more *complex* environments as K increases, since this produces environment polynomials with more main effects and interactions; and (ii) creates more *uncertain* environments as U increases, since then the relationship between cues and outcome becomes increasingly aleatory.

[9] We use multiplicative (rather than additive) noise so that the effect of uncertainty will scale with the number of terms in the environment polynomial. This is a simple way of making the effect of uncertainty comparable across environments that do not have the same number of terms (otherwise, increasing K would decrease the effect of uncertainty). An alternative specification that achieves the same is to use additive noise whose standard deviation grows with the number of terms in the environment polynomial (e.g., $\varepsilon \sim \text{Normal}(0, \sqrt{\#\text{terms}} \times U)$). All results are robust to using this alternative definition of uncertainty.

3.2 The representation

Following Brunswik, we conceptualize the firm's representation as the function that produces the predicted outcome $\hat{y}$. Continuing with our example, the software company's representation of the apps' profits might be given by $\hat{y} = 0.8x_1 - 0.7x_2 + 0.6x_1x_3$.

Analogously to how we defined environmental complexity K, we define representational complexity K' as the number of terms in the representation's polynomial. For instance, the representation in the previous paragraph has $K' = 3$ and we have $K' = 2$ for the representation $\hat{y} = x_1 + x_2$.

The extent to which a given representation is good at predicting outcomes depends on the degree to which the representation (a) captures relevant aspects of the environment (i.e., does not miss relevant cues or interactions and does not include irrelevant ones) and (b) weighs these aspects sensibly (i.e., uses $\hat{\beta}$ values that resemble the true β values). Firms would ideally prefer a representation that deviates minimally from the actual environment, since that state of affairs would endow firms with almost perfect foresight: $\hat{y}$ would be identical to y except for the unpredictable variation ε. However, the real world seldom permits one to acquire such a perfect representation. So to develop a realistic model of the role played by representations, we must account for how they are acquired. Our model focuses on representations acquired by learning from observations of the environment (we elaborate on other ways of acquiring representations, such as by hiring, in the Discussion section).

The Brunswikian framework, which views the decision maker as a "natural statistician," suggests a clear way of modeling how representations are learned. Learning a representation is akin to running a regression—namely, estimating coefficients (the $\hat{\beta}$'s) of the representation based on previously observed projects. We therefore model the process of learning a representation as running an ordinary least-squares (OLS) regression.[10]

We make this learning process a function of two contingencies: experience E and informedness I. We define experience as the number of previously observed projects (i.e., the number of "rows" of data available to run the regression). This contingency captures the idea that some managers (e.g., veteran managers) are more experienced than others (e.g., management trainees). Using experience as a contingency is consistent with previous research on learning, which uses experience as its main independent variable.[11]

[10]The effect of using an estimation method less efficient than OLS is akin to decreasing the number of observations available for estimation (parameter E, described next).

[11]The customary definition of organizational learning is "change in the organization's knowledge that occurs as a

The next contingency, informedness (I), captures our idea that learning a representation depends not only on experience but also on whether the manager is informed about the structure of the task environment. In terms of running a regression, this knowledge affects which regressors are included in the regression formula (i.e., the "columns" of data that are fed into the regression). We model two informedness conditions: uninformed ($I = 0$) and informed ($I = 1$). The gist of this contingency is that informed managers have knowledge that leads their representation to incorporate what is most relevant. Both informed and uninformed managers must estimate their representation using their experience, but the informed manager estimates the effect of the K' most relevant coefficients whereas the uninformed manager estimates the effect of K' randomly chosen coefficients.

The *informed* condition is representative of situations in which managers know what matters the most. Management education, such as the MBA program, arguably provides students with this type of knowledge for well-known industries. For instance, an MBA-trained investment banker analyzing the possible acquisition of a supermarket will know that some cues (such as the supermarket's location) are more relevant than other cues (such as the supermarket's décor). We model the informed condition by forcing the representation to pick the K' regressors that have the largest absolute magnitudes in the environment's polynomial. For example, a $K' = 2$ representation of the environment $y = 2+3x_1+4x_2-5x_1x_3$ could only take the form $\hat{y} = \hat{\beta}_2 x_2+\hat{\beta}_{1,3}x_1x_3$, as x_2 and x_1x_3 are accompanied by the two β's with the largest absolute magnitudes.

The *uninformed* condition is representative of situations in which there is no well-established knowledge about what determines performance. For example, in the 1990s it was not clear how firms could profit from the Internet and so managers held radically different representations of how best to evaluate Internet businesses (e.g., there was even debate on whether cash flows mattered; Desmet et al. 2000). Another example is that, because the process of manufacturing high-end computer chips is not well understood, Intel expands its production by "copying exactly" its successful facilities (i.e., replicating every characteristic of its successful plants, even their wall colors; Winter and Szulanski 2001). These two examples are similar in that uninformed managers have no reliable way to identify what are the relevant terms to include in their representations. We model the uninformed condition by picking the K' regressors in the representation randomly from all the $(1 + M + \frac{M(M-1)}{2})$ possible regressors. For instance, a $K' = 1$ representation of an $M = 3$ environment is equally likely to take

function of *experience*" (Argote 2013:31; emphasis added).

any of the following forms: $\hat{y} = \hat{\beta}_0$, $\hat{y} = \hat{\beta}_1 x_1$, $\hat{y} = \hat{\beta}_2 x_2$, $\hat{y} = \hat{\beta}_3 x_3$, $\hat{y} = \hat{\beta}_{1,2} x_1 x_2$, $\hat{y} = \hat{\beta}_{1,3} x_1 x_3$, or $\hat{y} = \hat{\beta}_{2,3} x_2 x_3$.

3.3 Performance

Firm performance is a measure of the firm's ability to screen projects. We assume that the firm only approves projects that are predicted to be profitable (i.e., approve only if $\hat{y}(x_1, \ldots, x_M) > 0$). Hence, the expected performance of the firm is defined as

$$\text{Performance} = \mathbb{E}[y \mathbb{1}[\hat{y} > 0]], \tag{1}$$

which is a function of the representation's complexity K', the environment's complexity K and uncertainty U, and the decision maker's experience E and informedness I. To facilitate interpreting our results, the performance reported in the analyses is scaled to fall in the 0–1 range. Here, 1 represents the best possible performance (i.e., approving all projects with $y > 0$ and rejecting the rest) while 0 represents the natural low-performance benchmark of simply approving all projects (i.e., a "lazy" screener who approves everything and thus performs no screening at all).

Performance is computed via simulation. Each screening decision is simulated as follows: (i) a random environment of complexity K and uncertainty U is generated; (ii) E random projects are drawn from that environment (for each project, the x's are independently drawn from a Normal$(0, 1)$ distribution); (iii) a representation (of a given complexity K' and informedness I) is estimated by running an OLS regression on the E projects observed; (iv) a new random project is drawn from the same environment; (v) the estimated representation is used to predict the new project's value ($\hat{y}$); and (vi) if the predicted value is positive ($\hat{y} > 0$) then the project's true value y is recorded (otherwise, 0 is recorded). The average of the recorded numbers (for each combination of parameter values) converges to the definition of performance (Equation 1) as the number of screening decisions simulated increases. To ensure that our results are reliable and not a function of any particular random draw, we run 10 million simulations for each combination of parameter values. The simulation code is available in the online appendix.

4 Results

Here we use the model developed in the previous section to study how the effect of representational complexity (K') on performance is contingent on the environment's complexity (K) and uncertainty (U) as well as on the decision maker's experience (E) and informedness (I).

To convey the results in an intuitive yet precise manner, our presentation is organized around a series of plots that are representative of the model's behavior (see Figure 4). The parameter values used to generate these plots were chosen according to the following logic. The number of main effects in the environment is set constant to $M = 3$.[12] Given this value of M, the environmental complexity K can vary from 1 to 7 ($= 1 + 3 + \frac{3(3-1)}{2}$). To understand the impact of K, it is enough to look at its effect at low, medium, and high values; therefore, the figure plots performance at $K = 1$, 4, and 7.

Representational complexity K' is the main independent variable in our analysis and is the parameter producing the most elaborate effects on performance. For that reason, all plots vary this parameter along its entire range (i.e., K' varies from 1 to 7 along the x-axes in Figure 4). To understand the effect of experience E and informedness I, it is enough to look at the effect of low and high values on each of these parameters. In Figure 4, these values appear as the "inexperienced" and "experienced" conditions ($E = 7$ and $E = 25$, as columns) and as the "uninformed" and "informed" conditions ($I = 0$ and $I = 1$, as rows).[13] Because varying uncertainty U affects performance in a straightforward way (increasing U dampens all performance curves similarly), most analyses in this section keep the value of U fixed at 0.5.

4.1 Results overview

An overall observation from Figure 4 is that the optimal representational complexity (marked with a "$\circ$" at the top of each curve) depends nontrivially on the model's contingencies—there is no ideal, one-size-fits-all representational complexity. The optimal level of representational complexity is a function of the circumstances: (i) in panels (a) and (b) it is best to use complex representations (i.e., curves peak around $K' = 6$–7); (ii) in panel (c) it is best to use simple representations (curves peak

[12] We tried other values of M and obtained results that were qualitatively equivalent to those presented here. For more details, see the Robustness Checks subsection.

[13] Experienced managers are represented by $E = 25$ because, in 99% of the simulations, performance has plateaued at this level and no further improvements are achieved by increasing experience. Inexperienced managers are represented by $E = 7$—the minimum number of observations needed to estimate a representation (i.e., solving for $K' = 7$ unknowns requires $E = 7$ equations). Values of experience in between these two extremes affect performance in a gradual and a predictable way, so it is enough for Figure 4 to plot just these two levels.

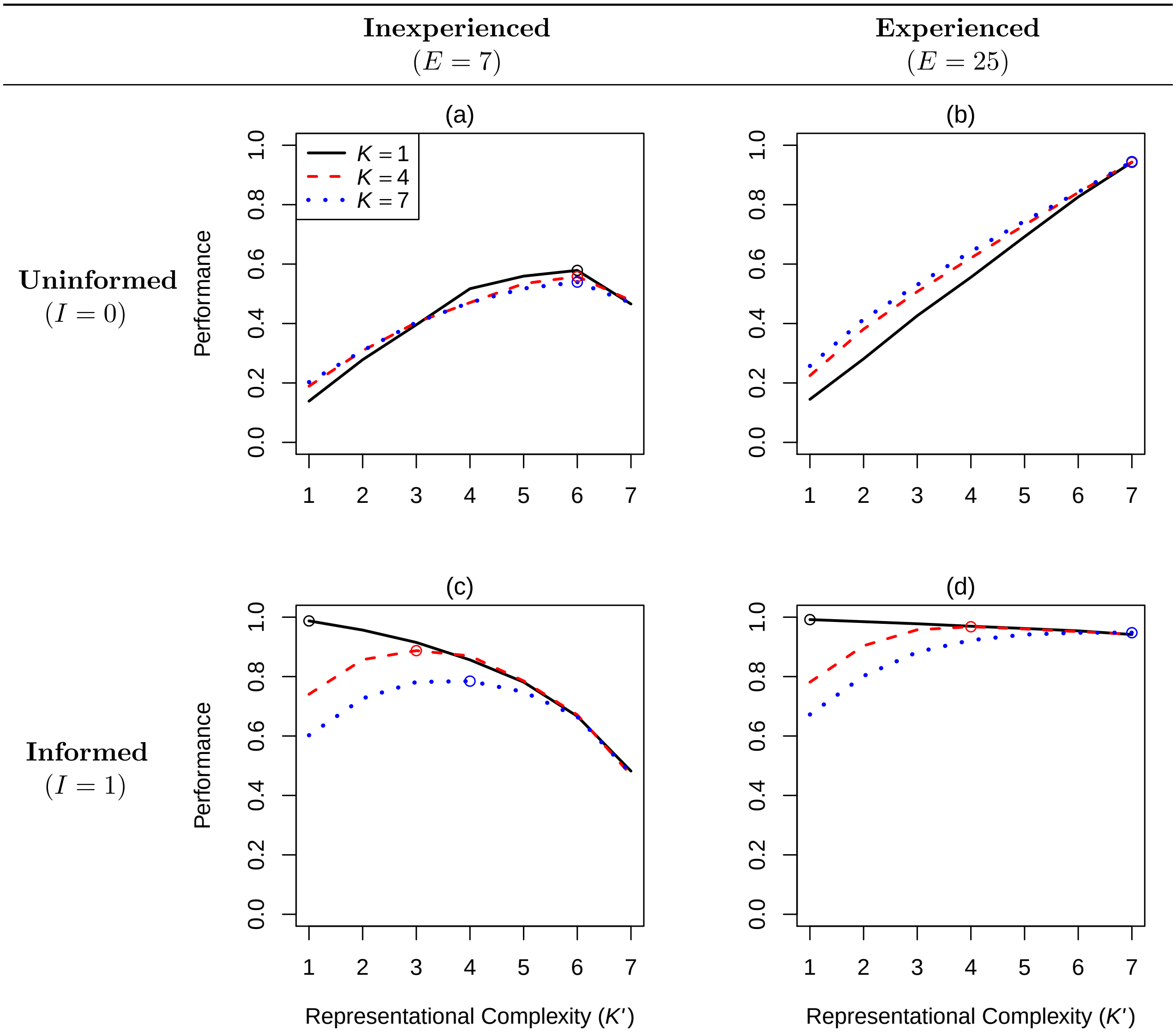

Figure 4: Performance as a function of representational complexity (K', on the x-axes), environmental complexity (K, as curves), experience (E, as columns), and informedness (I, as rows). The $\circ$'s mark the optimal representational complexity for each situation.

around $K' = 1$–4); and (iii) in panel (d) it is best to use representations that are as complex as the environment (i.e., each curve peaks exactly when $K' = K$). We remark that each of these three cases lends support to each of the three different theories of representational complexity discussed in the Introduction; namely, case (i) supports using complex representations (Weick 1979:261), case (ii) supports using simple representations (Gigerenzer and Goldstein 1996), and case (iii) supports using representations that match the environment's complexity (Ashby 1956).

Another observation from Figure 4 is that the optimal representational complexity depends more on characteristics of the firm (experience and informedness) than of the environment (complexity). In other words, the optimal representational complexity depends more on the panel than on the curve one is in. As a matter of fact, in panels (a) and (b) the optimal representational complexity is unaffected by environmental complexity; and in panels (c) and (d), although environmental complexity is consequential, choosing a medium representational complexity yields performance that is close to optimal. From a theoretical standpoint, this observation is relevant because the literature on "fit" (Donaldson 2001, Burton and Obel 2004) has emphasized how critical is the fit between environment and firm; in contrast, the critical fit relationship in our model hinges not on the environment but instead solely on characteristics of the firm (i.e., K' must primarily fit with I and E, not with K). This observation is also relevant because the literature on simple rules posits environmental unpredictability as its primary contingency (see, e.g., Davis et al. 2009); our research suggests that cognitive contingencies (i.e., experience and informedness) could have an even more important role in determining the appropriateness of simple rules.

To understand the mechanisms behind these and other results, we now examine in more detail each of the situations depicted in Figure 4.

4.2 Uninformed and inexperienced managers

Panel (a) in Figure 4 describes managers who are inexperienced ($E = 7$) and uninformed ($I = 0$). In comparison to the other panels in Figure 4, the managers in panel (a) know the least about the environment: they have neither experience nor any information about what characteristics of the environment might matter most. This panel is representative of radically novel situations, with regard to which no manager could have significant experience or knowledge (e.g., in 1998 neither Google nor its competitors had much experience with or knowledge about how to develop and commercialize an

Internet search engine).

A first observation from panel (a) is that performance follows an inverted-U shape with respect to representational complexity K'. This panel is consistent with the general intuition of the bias/variance trade-off mentioned above: neither too simple nor too complex representations are optimal. But as mentioned above, the trade-off does not explain much. For instance, it does not explain why the optimal is so high in this panel—a surprising observation given that from the literatures on heuristics and simple rules one would expect that managers who know little (i.e., have low experience and low informedness) would be better off by using simple representations. It also does not explain why in the other panels there is so much variation regarding the optimal representational complexity (in panel (b) the best is to set $K' = K$; and in the other panels there are cases where the best is to set $K' = 1$). Moreover, it does not explain why, to a large extent, the inverted-U shape only occurs in panel (a). To understand what drives the results in this and the other panels, it is useful to define the concepts of representational relevance and representational accuracy.

We define *representational accuracy* as the proximity of the representation's coefficients to the environment's corresponding coefficients. Representational accuracy can be expressed formally as $-\frac{1}{K'}\sum_{i\in\{i|\hat{\beta}_i\neq 0\}} |\beta_i - \hat{\beta}_i|$.[14] In turn, we define *representational relevance* as the extent to which a representation includes relevant terms (i.e., main effects or interactions that affect the outcome). We measure relevance by counting how many of the K coefficients that are relevant in the environment are part of the representation (i.e., the cases in which both β_i and $\hat{\beta}_i$ differ from 0). The mathematical formulation of representational relevance is $\frac{1}{K}\sum_{i\in\{i|\beta_i\neq 0\}} \mathbb{1}[\hat{\beta}_i \neq 0]$.

Representational accuracy and relevance are closely related to the concepts of bias and variance. Increasing relevance, decreases bias (as the representation captures more dimensions of the environment and, hence, will exhibit fewer consistent estimation errors). In turn, increasing accuracy, decreases variance (as the representation will produce estimates that are closer to reality). But discussing the results in terms of accuracy/relevance rather than bias/variance greatly simplifies the presentation of the results. This is so for three reasons. First, bias/variance decompositions add up to a total error; but in the organizations literature—and in the simple-versus-complex debate here studied—the dependent variable is typically performance, not errors. Second, even if there was a way to restate performance in terms of total error (e.g., by measuring a distance to optimal performance), there is

[14] In this expression, the set $\{i \mid \hat{\beta}_i \neq 0\}$ designates all the coefficients that are part of the representation; the initial negative sign makes the measure increase in the right direction.

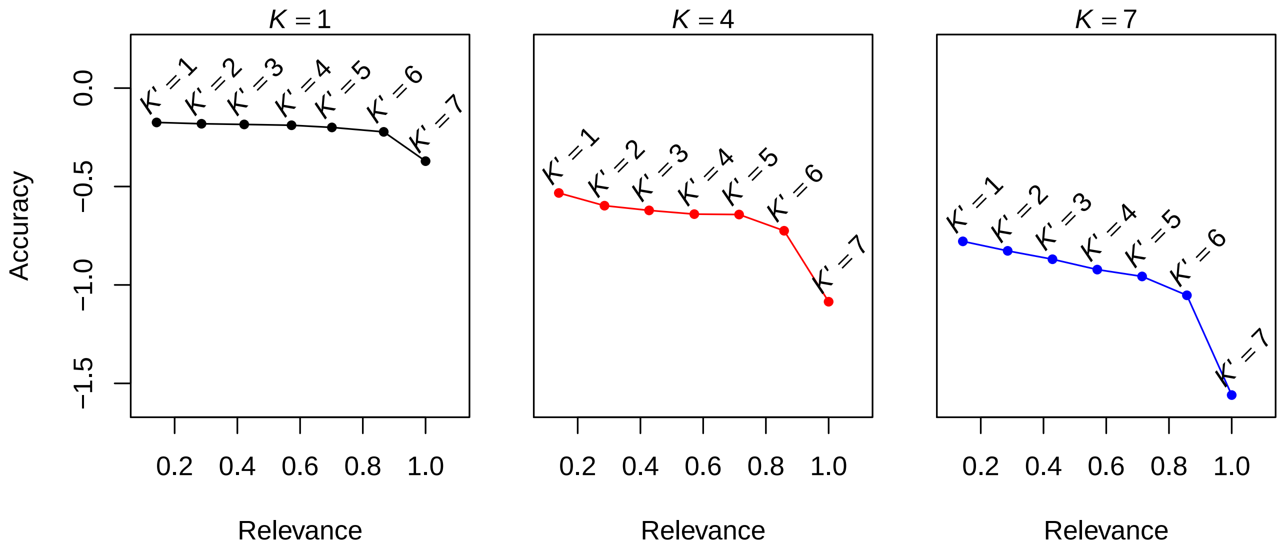

Figure 5: Median relevance and accuracy as a function of representational complexity (K') under uninformed and inexperienced managers. Medians (rather than means) are plotted in order to illustrate the behavior of a typical simulation.

no preestablished bias/variance decomposition for the type of problem studied here (as Equation 1 combines elements of both classification and prediction); hence, even if the bias/variance language was used, any numbers reported would not effectively "decompose" performance into two components. Finally, the accuracy and relevance measures here proposed allow for a more direct discussion of the mechanisms driving the results (as the effect of varying model parameters, such as K' or I, is easier to understand on accuracy/relevance than on bias/variance).

Figure 5 plots relevance against accuracy for the same conditions used in Figure 4(a). The downward slopes in Figure 5 demonstrate that, regardless of the environment's complexity, there is a trade-off between relevance and accuracy: by changing representational complexity K' (i.e., moving along each curve), it is impossible to increase one of these measures without decreasing the other.

The relevance–accuracy trade-off faced by managers who are both uninformed and inexperienced is explained by the following logic. On the one hand, relevance *increases* with K' because a larger K' increases the likelihood of the uninformed manager including relevant terms in his representation.[15] On the other hand, accuracy *decreases* with K' because as K' increases, the managers' limited experience becomes increasingly insufficient to estimate the greater number of coefficients. For instance, a $K' = 1$

[15] Suppose that an environment is described by $K = 7$ terms. Then a $K' = 1$ representation would include 1/7 of the relevant terms. The same representation in a $K = 1$ environment would have a 1/7 chance of including the relevant term. In both examples, an increase in K' would generate a proportional increase in the expected representational relevance.

representation can be estimated with relative accuracy by an inexperienced manager (since his $E = 7$ experience could be brought to bear on estimating a single $\hat{\beta}$); yet if asked to estimate a $K' = 7$ representation, the same manager would perform poorly (because his $E = 7$ experience would be spread too thin when estimating seven different $\hat{\beta}$'s).

In sum, Figure 4(a) exhibits an inverted-U shape because, when managers are uninformed and inexperienced, both simple and complex representations are problematic: simple representations are likely to be irrelevant, and complex representations are likely to be inaccurate. Hence, the optimal representational complexity is somewhere *between* the simple and the complex. This explains why, although managers in this panel know little about the environment (i.e., they are uninformed and inexperienced), they are better-off by using more complex representations than the more "modest" representations that their relative lack of knowledge might suggest.

A striking implication of panel (a) is that it contradicts two commonly espoused principles: Occam's razor and Ashby's law of requisite variety. Figure 4(a) runs counter to Occam's razor because the optimal representation is not parsimonious: whereas a scientist would not consider proposing a theory based on six covariates while using only seven data points as supporting evidence, panel (a) indicates that the manager should do exactly that in such case. It runs counter to Ashby's law because the optimal K' does not increase with K (all curves peak at $K' = 6$). We discuss the rationale for the Occam's razor result here and delay discussing Ashby's requisite variety until a few more results have been developed.

The key to understanding why the optimal representational complexity is lower for scientists and managers is to acknowledge that these professions pursue different goals; in particular, scientists seek to prove causal relationships while managers seek to make good predictions (Shmueli 2010). Hence, the goal of scientists is not to minimize the total prediction error (i.e., see Figure 3), but to demonstrate how a given factor affects the outcome. And when data is limited, as in Figure 4(a), demonstrating causal relationships calls for estimating few parameters, as this will narrow the confidence intervals and allow for more precise tests of hypotheses. In this panel, hence, Occam's razor is good advice for scientists but not for managers.

The role of uncertainty. The previous analyses have maintained a fixed uncertainty of $U = 0.5$. We now study the effect of varying uncertainty. Figure 6 illustrates what happens to the $K = 4$ curve in Figure 4(a) as uncertainty varies from an extremely low value ($U = 0.125$) to an extremely high one

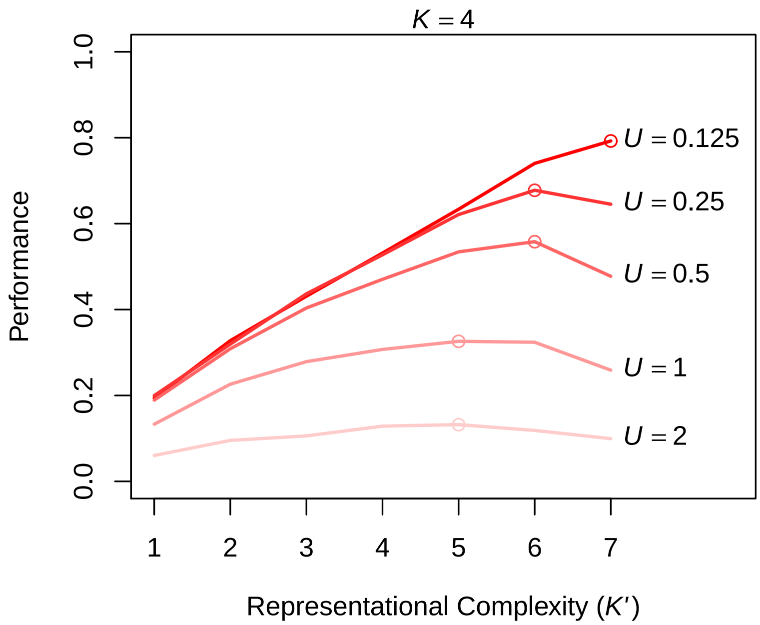

Figure 6: Effect of uncertainty (U) under uninformed and inexperienced managers.

($U = 2$).

A first observation from Figure 6 is that increasing uncertainty dampens performance. This observation (expected from our review of the heuristics literature) follows because increasing uncertainty decreases representational accuracy and therefore reduces the accuracy of predictions made using that representation.

Another observation from Figure 6 is that increasing uncertainty lowers the optimal representational complexity (as U increases, each $\circ$ in Figure 6 moves leftward). This occurs because increasing uncertainty decreases representational accuracy but *without* affecting representational relevance. So as U increases, accuracy costs more in terms of relevance (i.e., the downward slopes in Figure 5 become steeper with increasing U) and hence it becomes optimal to reduce K'. It is thus reasonable to expect that firms in uncertain environments will use simpler representations than firms in more certain environments.

This result corroborates the view that simple rules are particularly useful in uncertain environments (Eisenhardt and Sull 2001, Artinger et al. 2015:37). In turn, this result seems to run against Galbraith's (1973:4) well-known proposition that "the greater the uncertainty of the task, the greater the amount of information that has to be processed between decision makers." Yet, a closer inspection of the context in which Galbraith's states his proposition shows that he interprets uncertainty as "[not] understanding the task prior to performing it" (p.4). This interpretation corresponds not to our measure of uncertainty but to either uninformedness or inexperience. And, indeed, Galbraith's proposition interpreted in this way is consistent with our model's predictions: in Figure 6, the optimal representational complexity increases as one moves from the bottom to the upper row of panels (i.e., as uninformedness increases) and as one moves from the right to the left column of panels (i.e., as inexperience increases). These observations point to a benefit of the modeling approach we use: it forces one to define all constructs used and, hence, increases theoretical precision. In particular, we show that below the commonly-used term "uncertainty" underlie at least three possible meanings: unpredictability of the environment (high U), not knowing what matters (low I), and having little experience (low E). We suggest that future research on uncertainty specifies which definition they follow.

A final observation from Figure 6 is that, when uncertainty takes extremely low or high values, the inverted-U shape described so far degenerates into different shapes (see the $U = 0.125$ and $U = 2$ curves in Figure 6). If uncertainty is extremely low (e.g., $U = 0.125$), then estimation is so effective that increasing K' boosts relevance without reducing accuracy; therefore, performance increases monotonically with K' (in the extreme, if $U = 0$ then one can perfectly estimate all coefficients provided $E \geq K'$).

In contrast, if uncertainty is extremely high (see the $U = 2$ curve in Figure 6) then—although maximal performance is nominally achieved at some intermediate value of K'—choosing any K' produces almost equally low performance. This irrelevance of K' under high U, when combined with the greater cost of complex versus simpler representations (e.g., the former may take more resources to learn and use), suggests that firms in extremely uncertain environments are more likely to use simple than complex representations.

4.3 Uninformed and experienced managers

Panel (b) in Figure 4 differs from panel (a) in that, despite remaining uninformed ($I = 0$), managers are now experienced ($E = 25$). Panel (b) is representative of long-standing industries (and thus ones rife with experienced managers) in which there is *no* well-established knowledge about what drives performance (and so managers do not know a priori what are the environment's most relevant terms). For instance, the semiconductor industry has experience going back to the 1970s, but firms in this industry are still not certain about what characteristics make for high-quality manufacturing facilities (which is why, as mentioned previously, Intel replicates *all* the characteristics of its successful plants; Winter and Szulanski 2001). This panel is also representative of managers who, not knowing the rules driving a situation, feed copious amount of data ("big data") to a machine-learning algorithm in the hopes of finding useful relationships. In short, the uninformed and experienced condition is representative of situations where managers know many cases but not principles.

The main observation from Figure 4(b) is that performance now increases with representational complexity. In panel (a), performance follows an inverted-U shape because of the relevance–accuracy trade-off: past a given level of representational complexity, the representation's increased relevance carries too high of a price in terms of accuracy. This trade-off ceases to be critical in panel (b) because, even if the representation includes the same number of coefficients as the environment, experience is vast enough to enable the accurate estimation of *all* coefficients accurately. In contrast, setting K' below its maximal value impairs performance because it then becomes more likely that the representation will omit relevant terms.

The fact that in panel (b) performance increases with representational complexity, makes this panel consistent with Weick's (1979) advice to "complicate yourself!" and also with Intel's decision to behave as if *any* factor could affect the performance of its manufacturing facilities.

4.4 Informed and experienced managers

We now turn our attention to panel (d) of Figure 4; thus we study managers who are not only experienced but also *informed*—they can identify the environment's most relevant aspects. These managers are representative of situations in which there are well-known principles regarding what drives performance (e.g., graduates of an MBA real estate course know the main determinants of property values). Recall that informedness is modeled as estimating the K' coefficients that are most relevant in the environment.

An overall regularity that stems from comparing the informed and uninformed panels of Figure 4 is that the optimal K' is lower for informed managers (i.e., each curve in the bottom row of panels peaks before the corresponding curve in the top row). The reason is that the marginal benefit of increasing K' is high for uninformed managers (since the expected relevance of their representations increases with K') whereas such marginal benefit is lower for informed managers (since each additional term they add is less relevant than the previous terms). So according to our model, a manager who transitioned from being uninformed to being informed would likely benefit from reducing her K'. This outcome is consistent with—and offers a complementary explanation of—the "simplification cycling" phenomenon described by Bingham and Eisenhardt (2011:1454), who note that managers simplify their decision-making processes as they become more aware of the deeper structure of their respective environments.

Panel (d) in Figure 4, which applies only to managers who are both informed and experienced, is representative of situations where managers have vast experience as well as extensive training on what are the most relevant aspects of the environment. Such could be the case of managers in industries that are old and stable. For instance, the fields of retail banking and real estate are replete with managers who are both experienced (e.g., have worked for decades in the industry) and informed (e.g., have undergone extensive training).

The main observation from panel (d) is that performance is maximized when representational complexity exactly matches the environment's complexity. That is: for each curve in panel (d), the maximum occurs when $K' = K$. Hence this panel is consistent with Ashby's (1956) law of requisite variety because the representation is just as complex as the environment it represents.

To understand why managers who are informed and experienced are better-off choosing a representation with the same complexity as the environment, it is useful to consider the alternative cases: having such managers choose a representation that—in comparison with the environment—is either simpler ($K' < K$) or more complex ($K' > K$). Choosing $K' < K$ decreases representational relevance (i.e., omits relevant terms from the representation) without increasing representational accuracy by much (since managers are experienced). Choosing $K' > K$ is also damaging because it complicates the representation with spurious terms that will not increase relevance but will reduce accuracy (even if experience is vast, the coefficients estimated for the spurious terms will differ from their true value of 0). Because managers in panel (d) are experienced ($E \gg K'$), they pay only a small penalty for increasing representational complexity beyond environmental complexity; that is why the curves in panel (d) all have only a slightly downward slope after reaching the point where $K' = K$.

4.5 Informed and inexperienced managers

We now analyze panel (c) of Figure 4, where managers are informed but inexperienced. This panel is representative of situations in which there are well-understood principles but no experienced managers.

This could be the case for a new industry that is structurally similar to an old one but in which no experienced managers are employed (perhaps because experienced managers prefer to work for established firms). For example, managers in the smartphone apps industry and the PC software industry use similar principles for developing software (these principles are well established and broadly available from handbooks such as McConnell 2004); therefore, managers in these industries are similarly *informed*. However, early app development companies typically had less *experienced* managers than did PC software companies.

Another way of understanding the case described in panel (c) is by contrasting it to panel (a). In both panels, managers are inexperienced; yet panel (a) describes a radically novel situation (such as Google pioneering the search business) whereas panel (c) describes a less innovative situation.

The main observation from panel (c) is that, under these circumstances, firms are best served by using relatively simple representations. In fact, it is only in this panel that the performance of each curve is maximized under the least representational complexity. Thus, for example, the $K = 7$ curve in this panel achieves its maximum at $K' = 4$—in contrast to the other panels, where the maximum of the $K = 7$ curve occurs at $K' = 6$ or 7.

To understand why it is best to use simple representations in panel (c), it is useful to compare it with panel (d). The curves in both panels have similar shapes, but in panel (c) they exhibit a steeper decline after achieving their peaks. This is so because, in that panel, the penalty for using overly complex representations is more pronounced: experience is so limited in panel (c) that using it to estimate irrelevant coefficients is highly detrimental to accuracy.

In sum, our model shows that, under the conditions of panel (c) (i.e., informed but inexperienced managers), it is better to use simple representations. Hence this panel is consistent with the research advocating for simple representations (e.g., Gigerenzer and Goldstein 1996). One can argue that the situations studied by the heuristics literature match this panel. Indeed, the heuristics literature assumes little experience (Gigerenzer and Goldstein 1996:652) and, although it has not been shown generally, informedness appears to be a critical assumption. For example, for take-the-best (perhaps the most studied heuristic) to work well, it is crucial that high-validity cues are taken into account first (Katsikopoulos et al. 2010). In other words, take-the-best requires high informedness.

4.6 Robustness checks

We performed a number of additional analyses to check the robustness of our results. Overall, these analyses confirmed the qualitative robustness of the results described so far.

A first robustness check was to increase the number of main effects M. Increasing M changes the scale of some of the parameters (as the upper limit of K and K' depends on the value of M) without changing the results in relative terms. For example, in situations corresponding to Figure 4(a), the optimal K' is roughly 85% of the maximal K' regardless of M's value.

A second robustness check was to change the family of functions producing the environment. Here we tried a model with main effects only and no two-way interactions (i.e., $y = x_1 + \cdots + x_M$; in such a model, K and K' can vary only from 1 to M). The results here were qualitatively similar to those of the main model, since adding a main effect or a two-way interaction increases the environment's difficulty similarly.[16]

Finally, we tried a model under which K affects only the interactions in the environment (which is akin to how complexity is defined in NK models; Levinthal 1997). In this robustness check, all environments had M main effects yet K affected only the number of two-way interactions. Because the environment's difficulty is equally affected by adding main effects or interactions, this robustness check is effectively equivalent to restricting the analysis of the main model to values of K that are greater than M.

5 Discussion

This study introduces a model of how representational complexity affects firm performance in a way that is contingent not only on the environment's uncertainty and complexity but also on the decision maker's experience and informedness. The model provides a behaviorally plausible and organizationally relevant depiction of this phenomenon by combining ideas from the cognitive science literature and extending them to take into account contingencies and problematics that are germane to the organizations

[16] This is because the two-way interactions are uncorrelated with the main effects (i.e., $\text{corr}(x_i, x_i x_j) = 0$ if $x_i, x_j \sim \text{Normal}(0, \sigma)$) and, hence, estimating a main effect or an interaction is equally costly in terms of accuracy and relevance. The no-correlation assumption can be understood as the x's having been "orthogonalized" first (e.g., being the outcome of a principal component analysis). One could also explicitly model correlation, but doing so makes the model analysis unwieldy without providing much new insight, as its effect is not hard to predict: as correlation increases, the interactions have a larger effect on the outcome, hence correctly estimating the $\hat{\beta}$'s of the interactions becomes more important. Therefore, the effect of increasing correlation is equivalent to the effect of decreasing K (as fewer terms will be effectively driving y). This effect resembles the one described by Clemen and Winkler (1985:430) in the context of combining predictions from correlated experts.

	Simple Representations (Low K')	Complexity-Matching Representations ($K' = K$)	Complex Representations (High K')
Main proponents	• Gigerenzer and Goldstein (1996)	• Ashby (1956)	• Weick (1979)
Situation when most applicable	• Uncertainty is high (e.g., firm facing an unpredictable market, such as in the midst of political turmoil) • Informed but inexperienced managers (e.g., “low-tech” startup managed by newly minted MBAs)	• Informed and experienced managers (e.g., firm in a long-standing industry, such as banking or real estate, where well-established principles exist and where managers have observed many past projects)	• Uninformed managers (e.g., firm in a radically new industry, such as Google in 1998)

Table 1: Summary of the results derived from the model. The table shows the situations under which representations of different complexity are most applicable.

literature. The analysis of the model leads to several predictions regarding the optimal representational complexity for different situations.

Armed with insights derived from the model, in what follows we propose how previous conflicting empirical findings can fit together, present broader implications of our research for organization design and managerial cognition, and propose avenues for further research.

5.1 Resolving the simple-versus-complex debate

We started this paper by pointing out to the conflicting recommendations regarding the optimal complexity of representations. Our research shows that the three organizational theories addressing this issue—calling for simple, complex, and complexity-matching representations—can be integrated within a single coherent framework.

Table 1 summarizes the results derived from our model is in terms of the situations that call for different degrees of representational complexity. This table shows the situations that call for simple representations (left column), complex representations (right column), and representations whose complexity matches that of the environment (middle column). The three columns can be interpreted as situations under which the respective theories of Gigerenzer and Goldstein (1996), Weick (1979), and Ashby (1956) are most applicable.

From Table 1 we can see that, first, the calls for simple representations are most applicable when uncertainty is high (cf. Figure 6) or when managers are informed but inexperienced (cf. Figure 4(c)). Second, Weick’s call for complex representations is most applicable when managers are uninformed (cf.

panels (a) and (b) of Figure 4). Third, Ashby's call for representations whose complexity matches that of the environment is most applicable if managers are experienced and informed (cf. Figure 4(d)).

Although not explicitly studied by our model, it is worthwhile mentioning another condition that calls for simple representations: the cost of representations. For example, learning a complex representation may require hiring technical experts, and using such a representation may require that many personnel be trained and monitored; in contrast, learning and using a simple representation consumes far fewer resources. In practice, this means that simple representations may be convenient not only in the situations outlined in Table 1's left column but also when the marginal cost of increasing representational complexity exceeds its marginal benefit. Such cost considerations may be especially relevant for startups, which are typically resource constrained and for which increasing representational complexity comes at a high opportunity cost. That dynamic may explain the popularity among entrepreneurs of books advocating for simple rules (Sull and Eisenhardt 2015) and for closely related ideas such as running a "lean startup" and launching "minimum viable products" (see, e.g., Ries 2011, Blank and Dorf 2012). Understanding the costs of different representations is an empirical matter that could be studied by further work.

The conditions we have identified concerning the applicability of each theory fit well with the empirical settings that each theory has used for support. Fredrickson and Mitchell (1984) found support for simple representations in high uncertainty environments, which is consistent with results derived in the context of Figure 6. Similarly, Bingham and Eisenhardt (2011) found support for simple representations among startups executing their internationalization strategy. Arguably, the startups' managers are informed about internationalization (since this topic is taught in marketing and strategy courses) and the internationalization task they face is not extremely complex (e.g., it is less complex than replicating Intel's fabrication process). Hence this task falls within the low or medium K curves in panels (c) or (d), whose performance is maximized by setting a low-to-medium K'—that is, by using simple representations.

In turn, McNamara et al. (2002) found support for complex representations in the banking industry. One could certainly argue that the banking industry has many experienced managers (since this is an established industry) and that its business is not overly simple (as evidenced by the vast operations of many banks). The banking industry should therefore fall within the medium or high K curves in panels (b) or (d), whose performance is maximized by setting a high K'—that is, by using complex representations.

Finally, Ashby (1956) found support for complexity-matching representations in the context of developing an early robotic mechanism (the "homeostat," described in Ashby 1952) that could access a detailed record of the environment and received only informative inputs. Thus Ashby's mechanism is consistent with the experienced and informed curves of panel (d), whose performance is maximized by setting $K' = K$—that is, by using complexity-matching representations.

We conjecture that the recommendation to use simple rules is useful for many firms, as simple representations apply to more situations than do complex or complexity-matching representations. Whereas complex and complexity-matching representations each apply to only one specific situation (see right and middle columns of Table 1), simple representations are advisable in three different cases: high uncertainty, informed and inexperienced managers, or high costs of learning and using complex representations. Indeed, our model suggests that simple representations are suitable for a broad spectrum of firms. These firms range from startups, since they usually face high uncertainty as well as high costs of learning and using complex representations (because of resource constraints); to established firms, since they often employ managers who are informed (due to operating in industries with well-established principles) but inexperienced (owing to, e.g., employee turnover or changing markets).

5.2 Implications for organization design

The ideas here developed have theoretical and practical implications for organization design. From a theoretical standpoint, our work helps develop the role that uncertainty and cognition play in organization design. With respect to uncertainty, we show that the literature has interpreted this concept in different ways. Recall (from our discussion of Figure 6) that the heuristics literature has interpreted uncertainty in terms of unpredictability of outcomes (U in our model), while the classic organization design literature (Galbraith 1973) has interpreted it as lack of experience or lack of informedness (E and I in our model). For the former literature, uncertainty is a property of the environment (the environment is unpredictable); for the latter, it is a property of the organization (the organization does not have enough information to predict the environment). Our work highlights that it is important to distinguish between these different meanings of the word uncertainty, as they carry different implications (e.g., increasing U calls for simpler representations, while increasing E or I calls for more complex representations).

With respect to cognition, our work suggests that cognitive contingencies should play a key role in the contingency theory of organizations. Recall (from our initial discussion of Figure 4) that the optimal representational complexity depends more on the firm's experience and informedness than on the environment's complexity. This suggests that the contingency theory of organizations (whose main contingencies are structural, strategic, and environmental; Burton and Obel 2004:16) could also incorporate cognitive contingencies such as managers' experience and informedness. Doing so seems like a natural extension of the information processing logic that underlies contingency theory and organization design (Burton and Obel 2018).

From a practical standpoint, our work proposes organization design rules that complement prior work on such rules (Burton and Obel 2013). In particular, Table 1 can be read as four rules (one for each of the bullet points in the "situations" row); namely:

- If uncertainty is high, then use simple representations.
- If managers are informed but inexperienced, then use simple representations.
- If managers are uninformed, then use complex representations.
- If managers are informed and experienced, then use complexity-matching representations.

A final implication for organization design is that incorporating the idea of "representations" can extend the scope of organization design. Organization design is usually understood in terms of choosing the macro structure of the firm. For example, choosing whether the firm should use a functional, divisional, or matrix form or whether it should use a flat or a tall hierarchy (see, e.g., Burton and Obel 2004, chap. 2). This paper shows that a number of other structures—not only macro structures—depend on the same logics. That is, because the underlying logic is the same, it should be within the scope of organization design to design structures like rules and frameworks.

5.3 Implications for managerial cognition

Similarly to how thinking in terms of representations extends the purview of organization design, it also extends managerial cognition. Managerial cognition has studied managers' internal representations. But other types of representations can be studied with some of the same tools as internal representations (such as Brunswik's lens model and the bias/variance trade-off that drives our results). For instance, this paper suggests that the complexity of external representations will have similar effects on performance as the complexity of internal representations. Studying managerial cognition by looking at external representations may be a fruitful research path, as vis-à-vis internal representations, external representations are easier to observe and manipulate.

Another way in which our work contributes to the managerial cognition literature is by deepening the understanding of cognitive complexity. Previous research in this literature has shown that managers differ in their cognitive complexity (see, e.g., Tetlock 1983, Kiss and Barr 2015). In fact, measuring cognitive complexity is among the main uses the causal map methodology (see, e.g., Huff 1990, Clarkson and Hodgkinson 2005). The managerial cognition literature has taken a pro-complexity stance by arguing that complex mental representations allow managers to understand issues from multiple perspectives and pick appropriate responses (Bartunek et al. 1983) and that simple representations are associated with a failure to recognize opportunities accurately (Kiesler and Sproull 1982). As summarized in the context of Table 1, we propose that the effect of complexity is more nuanced. In particular, there are many situations (perhaps a majority of situations), where relatively simple representations are preferable.

An intriguing implication of our model in the context of the managerial cognition literature is that, under the right circumstances, some cognitive flaws can be beneficial. For instance, an "ignoramus" manager—one who focuses on a highly restricted set of cues—may be the right type of manager for situations that call for simple representations. Similarly, a "paranoid" manager—one who believes that every possible cue and interaction could matter—may be the right fit for situations that call for complex representations.

5.4 Managerial implications

Three important managerial implications emerge from our study. First, because representational complexity is contingent on the situation, managers must learn how to determine what situation they are in. Otherwise, they risk setting an inappropriate level of representational complexity (e.g., a manager who wrongly believes he is informed would, in effect, use the optimal K' from the lower panels in Figure 4 to set the K' of the upper panels, thereby achieving less-than-optimal results). Hence it is paramount that managers can accurately assess the extent to which they are informed and experienced as well as the degree to which the environment is complex and uncertain.

Second, because performance depends on setting a representational complexity that matches the situation, the firm must be able to influence its representational complexity. Doing so requires the firm to be aware of, and have command over, the managerial levers controlling representational complexity. That is, firms must be able to shape the internal, external, and distributed representations they use. This goal can be achieved in multiple ways. For instance, a firm can change its internal representations by hiring managers who have simpler or more complex representations or by using management education programs to instill representations of the right complexity. Firms can change their external representations by encouraging managers to use frameworks of the appropriate representational complexity (e.g., to choose—as needed—between a simple pros-and-cons analysis and a complex financial model incorporating hundreds of parameters). Finally, firms can change their distributed representations by designing organizational structures that are more or less hierarchical (e.g., relying on a single decision maker or instead on a committee to make decisions).

A final managerial implication concerns the value of pursuing experience and informedness. According to our results, moving from uninformed to informed or from inexperienced to experienced increases performance (in Figure 4, performance increases as one moves from the left to the right column and from the top to the bottom row). So if possible, firms should seek managers who are informed and/or experienced. Note that, in situations that are radically new, such managers may not exist or may be so scarce that it is uneconomical to hire them.

These three managerial implications are illustrated by a historical example. A tenet of many early Internet entrepreneurs was that the Internet had created a "new economy" that called for "new rules" (Kelly 1998). That tenet viewed managers as being uninformed (since they did not know the rules of the new economy) and inexperienced (because they had not managed an Internet business before). In terms of our model, this situation is described by Figure 4(a) and therefore calls for using complex representations. It is crucial to assess the veracity of this tenet because if it was false—and so the Internet did not render previous knowledge obsolete—then it would, in fact, have been useful to hire informed and experienced managers. Note also that hiring such managers would improve performance (since the performance curves are higher in panel (d) than in panel (a)) and would also call for setting representational complexity differently. In this example, then, firm performance depended critically on (i) being able to determine what situation the firm was in, (ii) setting the right representational complexity, and (iii) pursuing, if applicable, managers who were both informed and experienced.

5.5 Conclusions

This study set out to develop a contingency theory of representational complexity. To build such a theory, we imported and augmented ideas from cognitive science and used them to build a parsimonious yet realistic model of the contingent effect of representational complexity on firm performance.

Our study contributes to the organizations literature in three ways. First, making minimal assumptions—essentially, that representations are learned and then used to make predictions—we argue that the optimal representational complexity is contingent on characteristics of the environment and the firm. Second, this research sharpens our understanding of the circumstances under which different theories of representational complexity apply. In particular, we demonstrate that the three organizational theories addressing this issue—which call for simple, complex, and complexity-matching representations—can be integrated within a single coherent framework. Third, our study enriches the organizations literature by underscoring that representational complexity is both pervasive and consequential: pervasive because all decisions are made on the basis of a representation; and consequential because our model shows that firm performance is strongly (yet nontrivially) dependent on representational complexity.

Future research could further study the antecedents and consequences of representational complexity. With regard to *antecedents*, we have assumed that K' can be set but do not delve into how this is accomplished by organizations. Future research could study how different internal, external, and distributed representations affect representational complexity. For example, such research could compare the complexity of representations typically developed by different types of individuals (e.g., MBAs vs. non-MBAs), by different frameworks (e.g., checklist-type vs. more open-ended frameworks), and by different types of decision-making structures (e.g., fluid vs. structured; for a notable first step in this direction, see Davis et al. 2009). Experimental settings such as the ones used by Gary and Wood (2011) and Csaszar and Laureiro-Martinez (2018)—in which the complexity of the environment can be set and the complexity of the representations can be observed—could be used to test some of these questions. Regarding the *consequences* of representational complexity, in this paper we have focused on firms' expected performance at screening good from bad projects. Future research could study how representational complexity affects other performance outcomes, such as risk (i.e., variation in performance) as well as learning and exploration. This could be accomplished by studying how representations interact with well-studied search processes (e.g., Levinthal 1997, Posen and Levinthal

2012). For initial work along these lines, see Csaszar and Levinthal (2016) and Martignoni et al. (2016), which model how varying the accuracy of the representation of a landscape affects the outcomes of search on that landscape.

Viewing a firm's performance in terms of its representations amounts to a new way of conceptualizing the firm. Our paper has explored some of the theoretical and practical insights that flow from this viewpoint. Given that representations—internal, external, and distributed—underlie myriad organizational phenomena, we believe that more insights will emerge from furthering this perspective. The representational view of organizations explored here suggests that designing an organization is, above all, a matter of designing its representations. It follows that one of the central tasks of organizational researchers is to study the trade-offs among different types of representations. Our work on how to determine the optimal representational complexity is one step in that direction.

References

Alpaydin, E. 2014. *Introduction to Machine Learning*. The MIT Press, Cambridge, MA.

Argote, L. 2013. *Organizational Learning: Creating, Retaining, and Transferring Knowledge*. 2nd ed. Springer, New York.

Artinger, F., M. Petersen, G. Gigerenzer, J. Weibler. 2015. Heuristics as adaptive decision strategies in management. *Journal of Organizational Behavior* **36** S33–S52.

Ashby, W. R. 1952. *Design for a Brain*. Wiley, New York.

Ashby, W. R. 1956. *Introduction to Cybernetics*. Wiley, New York.

Bartunek, J. M., J. R. Gordon, R. P. Weathersby. 1983. Developing "complicated" understanding in administrators. *Academy of Management Review* **8**(2) 273–284.

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

Bingham, C. B., K. M. Eisenhardt. 2011. Rational heuristics: The 'simple rules' that strategists learn from process experience. *Strategic Management Journal* **32**(13) 1437–1464.

Blank, S., B. Dorf. 2012. *The Startup Owner's Manual: The Step-by-Step Guide for Building a Great Company*. K&S Ranch, Pescadero, CA.

Brunswik, E. 1952. *The Conceptual Framework of Psychology*. University of Chicago Press, Chicago, IL.

Burton, R. M., B. Obel. 2004. *Strategic Organizational Diagnosis and Design: The Dynamics of Fit*. 3rd ed. Kluwer, Boston, MA.

Burton, R. M., B. Obel. 2013. Design rules for dynamic organization design: The contribution of computational modeling. A. Grandori, ed., *Handbook of Economic Organization: Integrating Economic and Organization Theory*. Edward Elgar, Cheltenham, UK, 223–244.

Burton, R. M., B. Obel. 2018. The science of organizational design: Fit between structure and coordination. *Journal of Organization Design* **7**(1).

Clarkson, G. P., G. P. Hodgkinson. 2005. Introducing Cognizer: A comprehensive computer package for the elicitation and analysis of cause maps. *Organizational Research Methods* **8**(3) 317–341.

Clemen, R. T., R. L. Winkler. 1985. Limits for the precision and value of information from dependent sources. *Operations Research* **33**(2) 427–442.

Craik, K. J. W. 1943. *The Nature of Explanation*. Cambridge University Press, Cambridge, UK.

Csaszar, F. A. 2013. An efficient frontier in organization design: Organizational structure as a determinant of exploration and exploitation. *Organization Science* **24**(4) 1083–1101.

Csaszar, F. A. 2018. What makes a decision strategic? Strategic representations. *Strategy Science* **3**(4) 606–619.

Csaszar, F. A., J. P. Eggers. 2013. Organizational decision making: An information aggregation view. *Management Science* **59**(10) 2257–2277.

Csaszar, F. A., D. Laureiro-Martinez. 2018. Individual and organizational antecedents of strategic foresight: A representational approach. *Strategy Science* **3**(3) 513–532.

Csaszar, F. A., D. A. Levinthal. 2016. Mental representation and the discovery of new strategies. *Strategic Management Journal* **37**(10) 2031–2049.

Davis, J. P., K. M. Eisenhardt, C. B. Bingham. 2009. Optimal structure, market dynamism, and the strategy of simple rules. *Administrative Science Quarterly* **54**(3) 413–452.

Dawes, R. M. 1979. The robust beauty of improper linear-models in decision-making. *American Psychologist* **34**(7) 571–582.

Dawes, R. M., B. Corrigan. 1974. Linear-models in decision-making. *Psychological Bulletin* **81**(2) 95–106.

DeNisi, A. S., T. P. Cafferty, B. M. Meglino. 1984. A cognitive view of the performance-appraisal process: A model and research propositions. *Organizational Behavior and Human Performance* **33**(3) 360–396.

Denrell, J. 2003. Vicarious learning, undersampling of failure, and the myths of management. *Organization Science* **14**(3) 227–243.

Desmet, D., T. Francis, A. Hu, T. M. Koller, G. A. Riedel. 2000. Valuing dot-coms. *McKinsey Quarterly* (1) 148–157.

Dess, G. G., D. W. Beard. 1984. Dimensions of organizational task environments. *Administrative Science Quarterly* **29**(1) 52–73.

Domingos, P. 2000. A unified bias-variance decomposition for zero-one and squared loss. *Proceedings of the Seventeenth National Conference on Artificial Intelligence*. AAAI Press, 564–569.

Domingos, P. 2012. A few useful things to know about machine learning. *Communications of the ACM* **55**(10)

78–87.
Donaldson, L. 2001. *The Contingency Theory of Organizations*. SAGE Publications, Thousand Oaks, CA.
Dougherty, D. 1992. Interpretive barriers to successful product innovation in large firms. *Organization Science* **3**(2) 179–202.
Eggers, J. P., S. Kaplan. 2009. Cognition and renewal: Comparing CEO and organizational effects on incumbent adaptation to technical change. *Organization Science* **20**(2) 461–477.
Einhorn, H. J., R. M. Hogarth. 1975. Unit weighting schemes for decision making. *Organizational Behavior and Human Performance* **13**(2) 171–192.
Eisenhardt, K. M., D. N. Sull. 2001. Strategy as simple rules. *Harvard Business Review* **79**(1) 106–116.
Fox, C. R., G. Ülkümen. 2011. Distinguishing two dimensions of uncertainty. W. Brun, G. Keren, G. Kirkebøen, H. Montgomery, eds., *Perspectives on Thinking, Judging, and Decision Making*. Universitetsforlaget, Oslo, Norway, 21–35.
Fredrickson, J. W., T. R. Mitchell. 1984. Strategic decision-processes: Comprehensiveness and performance in an industry with an unstable environment. *Academy of Management Journal* **27**(2) 399–423.
Galbraith, J. R. 1973. *Designing Complex Organizations*. Addison-Wesley, Reading, MA.
Gary, M. S., R. E. Wood. 2011. Mental models, decision rules, and performance heterogeneity. *Strategic Management Journal* **32**(6) 569–594.
Geman, S., E. Bienenstock, R. Doursat. 1992. Neural networks and the bias/variance dilemma. *Neural Computation* **4**(1) 1–58.
Gigerenzer, G., W. Gaissmaier. 2011. Heuristic decision making. *Annual Review of Psychology* **62** 451–482.
Gigerenzer, G., D. G. Goldstein. 1996. Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review* **103**(4) 650–669.
Hamm, R. M., H. Yang. 2017. Alternative lens model equations for dichotomous judgments about dichotomous criteria. *Journal of Behavioral Decision Making* **30**(2) 527–532.
Hastie, T., R. Tibshirani, J. H. Friedman. 2009. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2nd ed. Springer, New York.
Hinton, G. E., J. L. McClelland, D. E. Rumelhart. 1986. Distributed representations. D. E. Rumelhart, J. L. McClelland, the PDP Research Group, eds., *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Vol. 1: Foundations*, chap. 3. MIT Press, Cambridge, MA, 77–109.
Hogarth, R. M., N. Karelaia. 2007. Heuristic and linear models of judgment: Matching rules and environments. *Psychological Review* **114**(3) 733–758.
Holland, J. H., K. J. Holyoak, R. E. Nisbett, P. Thagard. 1986. *Induction: Processes of Inference, Learning, and Discovery*. MIT Press, Cambridge, MA.
Huff, A. S. 1990. *Mapping Strategic Thought*. Wiley, Chichester, NY.
Hursch, C. J., K. R. Hammond, J. L. Hursch. 1964. Some methodological considerations in multiple-cue probability studies. *Psychological Review* **71**(1) 42–60.
Hutchins, E. 1995a. *Cognition in the Wild*. MIT Press, Cambridge, MA.
Hutchins, E. 1995b. How a cockpit remembers its speeds. *Cognitive Science* **19**(3) 265–288.
James, G. M. 2003. Variance and bias for general loss functions. *Machine Learning* **51**(2) 115–135.
Kaplan, S. 2011. Strategy and PowerPoint: An inquiry into the epistemic culture and machinery of strategy making. *Organization Science* **22**(2) 320–346.
Karelaia, N., R. M. Hogarth. 2008. Determinants of linear judgment: A meta-analysis of lens model studies. *Psychological Bulletin* **134**(3) 404–426.
Katsikopoulos, K. V., L. Martignon. 2006. Naive heuristics for paired comparisons: Some results on their relative accuracy. *Journal of Mathematical Psychology* **50**(5) 488–494.
Katsikopoulos, K. V., L. J. Schooler, R. Hertwig. 2010. The robust beauty of ordinary information. *Psychological Review* **117**(4) 1259–1266.
Kelly, K. 1998. *New Rules for the New Economy*. Viking, New York.
Kiesler, S., L. Sproull. 1982. Managerial response to changing environments: Perspectives on problem sensing from social cognition. *Administrative Science Quarterly* **27**(4) 548–570.
Kirsh, D. 2010. Thinking with external representations. *AI & Society* **25**(4) 441–454.
Kiss, A. N., P. S. Barr. 2015. New venture strategic adaptation: The interplay of belief structures and industry context. *Strategic Management Journal* **36**(8) 1245–1263.
Krackhardt, D. 1990. Assessing the political landscape: Structure, cognition, and power in organizations. *Administrative Science Quarterly* **35**(2) 342–369.

Levinthal, D. A. 1997. Adaptation on rugged landscapes. *Management Science* **43**(7) 934–950.
Lindsay, P. H., D. A. Norman. 1977. *Human Information Processing: An Introduction to Psychology*. Academic Press, New York.
Martignon, L., U. Hoffrage. 2002. Fast, frugal, and fit: Simple heuristics for paired comparison. *Theory and Decision* **52**(1) 29–71.
Martignoni, D., A. Menon, N. Siggelkow. 2016. Consequences of misspecified mental models: Contrasting effects and the role of cognitive fit. *Strategic Management Journal* **37**(13) 2545–2568.
McConnell, S. 2004. *Code Complete*. 2nd ed. Microsoft Press, Redmond, WA.
McNamara, G. M., R. A. Luce, G. H. Tompson. 2002. Examining the effect of complexity in strategic group knowledge structures on firm performance. *Strategic Management Journal* **23**(2) 153–170.
Milliken, F. J. 1990. Perceiving and interpreting environmental change: An examination of college administrators' interpretation of changing demographics. *Academy of Management Journal* **33**(1) 42–63.
Norman, D. A. 1991. Cognitive artifacts. J. M. Carroll, ed., *Designing Interaction: Psychology at the Human-Computer Interface*, chap. 2. Cambridge University Press, Cambridge, UK, 17–38.
Payne, J. W., J. R. Bettman, E. J. Johnson. 1993. *The Adaptive Decision Maker*. Cambridge University Press, Cambridge, UK.
Porac, J. F., H. Thomas, C. Baden-Fuller. 1989. Competitive groups as cognitive communities: The case of Scottish knitwear manufacturers. *Journal of Management Studies* **26**(4) 397–416.
Posen, H. E., J. S. Chen. 2013. An advantage of newness: Vicarious learning despite limited absorptive capacity. *Organization Science* **24**(6) 1701–1716.
Posen, H. E., D. A. Levinthal. 2012. Chasing a moving target: Exploitation and exploration in dynamic environments. *Management Science* **58**(3) 587–601.
Priem, R. L., L. G. Love, M. A. Shaffer. 2002. Executives' perceptions of uncertainty sources: A numerical taxonomy and underlying dimensions. *Journal of Management* **28**(6) 725–746.
Ries, E. 2011. *The Lean Startup*. Crown Business, New York.
Robbins, P., M. Aydede. 2009. A short primer on situated cognition. P. Robbins, M. Aydede, eds., *The Cambridge Handbook of Situated Cognition*, chap. 1. Cambridge University Press, Cambridge, UK, 3–10.
Sah, R. K., J. E. Stiglitz. 1986. The architecture of economic systems: Hierarchies and polyarchies. *American Economic Review* **76**(4) 716–727.
Shmueli, G. 2010. To explain or to predict? *Statistical Science* **25**(3) 289–310.
Siggelkow, N., J. W. Rivkin. 2005. Speed and search: Designing organizations for turbulence and complexity. *Organization Science* **16**(2) 101–122.
Simon, H. A. 1957. *Models of Man: Social and Rational*. Wiley, New York.
Simon, H. A. 1962. The architecture of complexity. *Proceedings of the American Philosophical Society* **106**(6) 467–482.
Simon, H. A. 1990. Invariants of human behavior. *Annual Review of Psychology* **41** 1–19.
Spender, J. C. 2013. Herbert Alexander Simon: Philosopher of the organizational life-world. M. Witzel, M. Warner, eds., *The Oxford Handbook of Management Theorists*, chap. 15. Oxford University Press, Oxford, UK, 297–357.
Sull, D. N., K. M. Eisenhardt. 2015. *Simple Rules: How to Thrive in a Complex World*. Houghton Mifflin Harcourt, Boston, MA.
Tetlock, P. E. 1983. Accountability and complexity of thought. *Journal of Personality and Social Psychology* **45**(1) 74–83.
Thagard, P. 2005. *Mind: Introduction to Cognitive Science*. 2nd ed. MIT Press, Cambridge, MA.
Vuori, N., T. Vuori. 2014. Comment on "Heuristics in the strategy context" by Bingham and Eisenhardt (2011). *Strategic Management Journal* **35**(11) 1689–1697.
Walsh, J. P. 1995. Managerial and organizational cognition: Notes from a trip down memory lane. *Organization Science* **6**(3) 280–321.
Wegner, D. M. 1986. Transactive memory: A contemporary analysis of the group mind. B. Mullen, G. R. Goethals, eds., *Theories of Group Behavior*. Springer-Verlag, New York, 185–208.
Weick, K. E. 1979. *The Social Psychology of Organizing*. 2nd ed. Addison-Wesley Publishing Company, Reading, MA.
Winter, S. G., G. Szulanski. 2001. Replication as strategy. *Organization Science* **12**(6) 730–743.
Wirth, N. 1976. *Algorithms + Data Structures = Programs*. Prentice-Hall, Englewood Cliffs, NJ.
Zhang, J. J. 1997. The nature of external representations in problem solving. *Cognitive Science* **21**(2) 179–217.

Online Appendix: Code

The following code (in R) can be used to re-create any result from the model.

```
create_env <- function(M, K, U, E, test_rows=2) {
  main_effects <- paste0('x',1:M)
  interactions <- apply(combn(main_effects, 2), 2, paste, collapse='*')
  all_terms    <- c(1, main_effects, interactions)
  terms        <- all_terms[sort(sample(1:length(all_terms), K))]
  coeffs       <- rnorm(K)
  fmla         <- paste(apply(rbind(coeffs, terms), 2, paste, collapse='*'),
                        collapse=' + ')
  nrows        <- E + test_rows
  Xs           <- as.data.frame(matrix(rnorm(nrows*M), nrow=nrows, ncol=M,
                                       dimnames=list(1:nrows, main_effects)))
  Ys_no_noise  <- eval(parse(text=fmla), envir=as.list(Xs))
  Ys           <- Ys_no_noise * rnorm(nrows, mean=1, sd=U)
  list(all_terms=all_terms, terms=terms, coeffs=coeffs, fmla=fmla, U=U, Xs=Xs[1:E,],
       Ys=Ys[1:E], test_Xs=Xs[(E+1):(E+test_rows),], test_Ys=Ys[(E+1):(E+test_rows)])
}
# Example usage:
# env <- create_env(M=3, K=2, U=0.5, E=7, test_rows=2); env

OLS <- function(X, Y) tryCatch(qr.solve(X, Y), error=function(e)
{ message('Ill-conditioned matrix'); dput(cbind(X,Y)); rep(NA, ncol(X)) })

estimate_repr <- function(env, Kprime, I) {
  if (I==0) {
    terms <- env$all_terms[sort(sample(1:length(env$all_terms), Kprime))]
  } else {
    terms <- c(env$terms[order(-abs(env$coeffs))],
               sample(setdiff(env$all_terms, env$terms)))[1:Kprime]
  }
  Xs_obs  <- eval(parse(text=paste0('cbind(',paste(terms, collapse=','),')')),
                  envir=as.list(env$Xs))
  colnames(Xs_obs) <- terms
  if (length(terms)==1 && (terms=='1')) Xs_obs <- matrix(Xs_obs, length(env$Ys))
  coeffs  <- OLS(Xs_obs, env$Ys)
  fmla    <- paste(apply(rbind(coeffs, terms), 2, paste, collapse='*'), collapse=' + ')
  pred_Ys <- eval(parse(text=fmla), envir=as.list(env$test_Xs))
  if (length(terms)==1 && (terms=='1')) pred_Ys <- matrix(pred_Ys, nrow(env$test_Xs))
  list(terms=terms, Xs_obs=Xs_obs, coeffs=coeffs, fmla=fmla, pred_Ys=pred_Ys)
}
# Example usage:
# repr <- estimate_repr(env, Kprime=1, I=1); repr

sim <- function(M, K, U, Kprimes, E, I, nsims, ntests, rawdata=FALSE) {
  onesim <- function(x) {
    repeat { # redraw if matrix is ill-conditioned (1-in-10 million chance)
      env  <- create_env(M, K, U, E, test_rows=ntests)
      repr <- estimate_repr(env, Kp, I)
      if (is.numeric(repr$coeffs)) break()
    }
    data.frame(test_Y=env$test_Ys, pred_Y=repr$pred_Ys)
  }
  res <- list()
  for (Kp in Kprimes) {
    D <- do.call(rbind, lapply(1:nsims, onesim))
    D <- cbind(expand.grid(Kprime=Kp,ntest=1:ntests,nsim=1:nsims), D)
```

```
    D$decision     <- 1*(D$pred_Y>0)
    D$actual_perf  <- 1*(D$pred_Y>0) * D$test_Y
    D$optimal_perf <- 1*(D$test_Y>0) * D$test_Y
    perf <- (mean(D$actual_perf)-mean(D$test_Y))/(mean(D$optimal_perf)-mean(D$test_Y))
    res[[Kp]] <- c(perf=perf)
  }
  res <- do.call(rbind, res); rownames(res) <- Kprimes
  if (rawdata) { D } else { res }
}

# Replicate any line in Figure 2
K=4; U=0.5; E=7; I=1
D <- sim(M=3, K=K, U=U, Kprimes=1:7, E=E, I=I, nsims=1000, ntests=100)
plot(D, ylim=c(0,1), type='l', xlab="Representational Complexity (K')", ylab='Performance',
     cex.main=1, font.main=1, main=sprintf('K=%s   U=%s   E=%s   I=%s',K,U,E,I))
```