Max-Planck-Institut
für Mathematik
in den Naturwissenschaften
Leipzig

Boundedly rational strategic interaction and the interplay between complexity and simplification: Implications for strategy science

by

Jürgen Jost and Timo Ehrig

Preprint no.: 97 2019

Boundedly rational strategic interaction and the interplay between complexity and simplification: Implications for strategy science

Jürgen Jost and Timo Ehrig

Abstract

It is a basic question of strategy science how we can cope with an uncertain and complex world, taking into account limitations of information access or computing power. Using new insights from complexity science and exploring analogies with current machine learning which in many respects is faced with the same problem, we can not only postulate that one should simplify to cope with complexity and uncertainty, but we can also analyze how one should best simplify. In particular, we argue that strategists need to apply metaheuristics to infer and create structure that organizes interactions, both among detailed choices that firms make, and among strategic agents such as competitors, suppliers and customers. Our framework contributes to the literature on dynamic capabilities, as we develop a systematic and new understanding of heuristics to cope with different types of change. Moreover, it contributes to our theoretical understanding of search, opportunity capture and competition, as we qualify and extend the notion of a competitive landscape.

1 Introduction

Boundedly rational actors dealing with complex problems need to construct simplified models (Simon, 1957). They form simplified models of their available actions and their potential consequences. To anticipate potential consequences, they also form simplified models of their interaction partners, such as customers, suppliers and competitors. Thus, boundedly rational strategy interaction involves *modelling modellers*. Moreover, strategy and organization science models such modelling modellers to understand the aggregate dynamics of search and competition. But doesn't that get paradoxical? We started with simplification, but it seems that such nested modelling becomes ever more complex. Is that inevitable? Our answer here is *No*, the more complex the interactions that characterise a situation, the simpler the models should be to cope with it. Only when the situation becomes better understood, the models should gradually become more detailed. Thus, we shall propose that superior means to simplify when complex interaction environments are modelled are important managerial capabilities and, ultimately, sources of competitive advantages.

The idea that strategy is at its heart a simplification is not new. It has been articulated at least 30 years ago (Mintzberg, 1990). In the past, simplification has primarily been regarded as negative (Miller, 1993). However, according to (Leiblein et al., 2018) strategic decisions are characterized by interactions among detailed choices (as in the notion of activity systems; see (Porter & Siggelkow, 2008)), interactions among strategic agents (as in game theory, see Brandenburger and Stuart, 1996) and interactions of choices made on different time scales. Clearly, the different types of interactions that render a decision strategic interrelate, and interrelated interdependencies create complexity. To cope with it, strategic decision-making *necessarily* requires a process of simplification (Bettis, 2017). Therefore, we suggest to consider complexity reduction as a characteristic feature of *strategic* decisions. Given the necessity to simplify, superior simplification skills may provide individual or firm-level advantages. There exist other fields that are dealing with analogous problems, like machine learning and data science, robotics and theoretical biology, and more generally, complexity science, and it is therefore natural to seek some guidance from them. That is what we shall start in this paper.

We shall analyze simplification as a *process* that can also be iterated (Schilling, 2019). Since different types of situations may require different simplification strategies, one may first need a simple method to decide which such strategy is adequate. Drawing on newer results in the sciences of complexity, we can define, conceptualize and analyze different simplification tasks of strategists. In fact, such an interaction between strategy science and the sciences of complexity is not new. The roots of organization and strategy science are intimately tied to complexity science. The analogies between organization science on one hand and artificial intelligence and computer science on the other hand can not only be seen in the work of Simon who contributed to both fields, but also, for instance, in the role that heuristics continue to play in these fields, compare for instance (Gigerenzer and Selten, 2001, Gigerenzer and Gaissmaier, 2011) and (Pearl, 1984, Pearl, 1988/1997). As we shall discuss, we define heuristics as simple rules, or in other words, as the outcomes of simplification processes. However, heuristics also need to be applied to *select* appropriate heuristics to deal with a given situation. Such heuristics are *metaheuristics*, as the help to select appropriate heuristic responses.

As heuristics are outcomes of simplifications, applying them will not always yield optimal outcomes. However, individuals and organizations operate in environments that are too complex, and too unknown, to be dealt with using global optimization schemes. While economics by and large still tries to find such optimization schemes, organization and strategy science took a radical departure from this view of rationality and decision making. Bounded rationality, as proposed by Herbert Simon, is deeply connected with complexity reduction. Individuals and organizations need to cope with environments that are vastely more complex than internal representations held by managers or distributed in organizations could capture (Bettis and Prahalad, 1995). Yet some individuals and organizations are successful to operate in, and shape such environments. This raises fundamental questions that are only partially answered: What exactly is complexity? How is it related to uncertainty, that is, a condition under which action possibilities and their consequences are unknown, or even unknowable? How does strategic interaction create complexity, and in turn, how does complexity influence strategic interaction, and in particular, competition? How can it theoretically be reduced, and how do individual agents and organizations reduce it in practice?

In the last 20 years, progress has been made about answering these questions. Complexity has been characterized as the degree to which detailed choices of a business organization interact and create dependencies (Rivkin & Siggelkow, 2007). In this context, a bulk of studies investigated when and how mental representations successfully help firms to find high performing strategies, here understood as sets of interdependent choices (Gavetti and Levinthal, 2000, Martignoni, Menon & Siggelkow, 2016).

Less is understood about the complexity arising from the interaction of boundedly rational agents in a fundamentally uncertain world. It has become clear that economic evolution is not just driven by creative insights about novel resource combinations that are of potential value. The competitive landscape is cognitively constructed, at least to some degree. It therefore changes not only due to external factors or technological and other innovation, but also because of the insights, the conceptions and the misconceptions of the agents. These agents have to cope with the fact that they are ignorant about most opportunities to shape the future, and they can be surprised when unrealized possibilities are discovered by other agents. In sum, neither all relevant agents, their action possibilities, nor the consequences of actions can be fully known. That does not imply, however, that some agents cannot be better in anticipating interaction outcomes than others. For instance, Kodak was successful in building digital imaging capabilities, but failed to anticipate the resulting new logic of competition (Grant, 2016). By contrast, IBM successfully anticipated the new logic of competition in the post mainframe age of computing. The thesis of this article is that success and failure in the anticipation of complex interaction can be theoretically explained and grounded in insights from the sciences of complexity.

Our guiding concept will be *strategic bounded rationality* (henceforth abbreviated as **SBR**). SBR is related to the concept of ecological rationality (Gigerenzer and Selten, 2001), but extends beyond it. SBR describes interaction situations in which all agents are limited in access to information, computation time, and computational abilities, and moreover, the environment evolves: The set of interaction rules and choice sets of agents is not fixed, but to a certain degree shaped by the agents' actions. While efforts have been made to model such situations (Gavetti, Helfat and Marengo, 2017), we provide a theoretical concept of

SBR that is grounded in the sciences of complexity. This grounding is productive, for two reasons. First, it will enable us to work towards a theory of heuristics in strategic and evolving settings. While Gigerenzer and colleagues started to discover theoretical principles that underlie the positive function of heuristics, in particular, the bias-variance trade-off, we go further and explore a range of other principles of simplification known from the sciences of complexity to explain the mechanisms behind the success of heuristics in strategic settings. As strategic decisions are defined by their complexity (as noted above, (Leiblein et al., 2018)'s definition of strategic decision implies complexity) insights from the sciences of complexity naturally apply to strategic decision-making.

Second, our argument will provide an additional rationale of why (meta-)heuristics are dynamic capabilities. Metaheuristics are dynamic capabilities as they allow to systematically and reliably detect problem structures (Baer, Dirks, and Nickerson, 2013). Metaheuristics are useful to infer the type of change a strategist is confronted with and to select appropriate cognitive and organizational tools to adapt to the change. Thus, metaheuristics help to reconfigure organizational resources when firms face change and thus meet the defining property of dynamic capabilities. Our theoretical discussion informs the discourse on dynamic capabilities, as some of the principles that help to detect and deal with change, known from the sciences of complexity, were simply not introduced into the strategy discourse until now. Teece's (2007) concept of sensing is of course related, but Teece does not address how the problem structure is inferred, and we develop some systematic insights.

To approach this, we revisit and extend the notion of a competitive landscape. While the ultimate underlying reality may neither be known to scientists nor to modelling strategists in firms, strategists can use glimpses of feedback to infer the actual structure of reality, that is, to get a sense how the competitive landscape in a novel setting may look. In other words, we argue that a key simplification skill is to infer the structure of a possible competitive landscape on which agents adapt, and thereby construct such a landscape. To illustrate this, in our view, the landscape on which Tesla's cars are positioned today did not readily exist before Elon Musk shaped it by creating a "proof of concept" (the Tesla S model). The S model created expectations that made major players (VW etc) move and those moves then "span" a new competitive landscape that is a result of demand patterns, technological expectations, etc. Musk correctly anticipated a possible structure of the transition from the combustion engine to electric cars: That electric cars can first be luxury goods that do everything better than combustion engine cars but can do things that combustion engine cars can't (eg., extraordinary acceleration). Thus, Musk inferred a possible demand pattern and a possible logic of competition (technological race in the premium segment) correctly. We can derive insights about how agents should make such inferences drawing on concepts such as "correlation length" from the sciences of complexity. Heuristics to infer the structure of a transition are metaheuristics, as they then help to deploy appropriate heuristics responses. For instance, if the transition to electric cars is about a technology race in the premium segment, a heuristic response is to first create a powerful prototype to create positive expectations about the transition. By contrast, this response would have resulted in very negative outcomes for the case of Kodak and the transition to digital images. The structure of the competitive landscape in digital imaging is entirely different: Digital images are a by-product, the new bottlenecks are with digital devices such as smartphones. If a metaheuristic had allowed Kodak to infer this correctly, a suitable heuristic response would be to not focus on developing a high quality prototype of a digital imaging process, but a prototype of a digital lifestyle device that creates a new customer experience.

The approach in this article is different from empirical studies on heuristics. We will discuss theoretical explanations of well performing heuristics under SBR from a complexity science point of view. In Section 4.2, we shall see how statistical learning theory tells us how to manage the conflicting demands of model accuracy on a given data set and its predictive power for future data. This will lead to a deeper understanding of the bias-variance trade-off (Gigerenzer and Brighton, 2009, Ehrig and Schmidt, 2019) and extend beyond it. We will add explanations from other fields of machine learning to provide some kind of a meta-perspective about selecting simplification tools that are appropriate in a given context. When these tools are heuristics, they are metaheuristics for selecting a heuristic that is adapted to the situation at hand.

We have already mentioned *complexity* several times. But what complexity is and how it can be handled may depend on the setting. To clarify that, we shall recall two basic and well established distinctions. First,

Table 1: **Types of interactions under SBR**

1 well-structured + choice interaction	2 well-structured + strategic interactions
3 ill-structured + choice interactions	4 ill-structured + strategic interactions

The boxes are labeled as 1,2,3,4 so that we can subsequently refer to them simply by their label.

task environments can be well structured or ill-structured (Simon, 1973). In the former case, the structure of the problem is clear, and what counts as a solution is not controversial, and one could in principle compute the solution if one had sufficient resources. But in practice, the required resources may be so demanding that an effective solution is computationally, organizationally or cognitively intractable[1]. In the latter case, the environment is not provided in the form of a well-defined symbolic description. Rather, the situation is opaque. It is not clear who the key factors and players are, who knows what and what the options and the possible pitfalls are.

Second, as argued above, complexity can arise from interactions of detailed choices that a strategist makes, but it can also arise from strategic interaction. Thus, there are two basic distinctions: well-structured versus ill-structured and choice interactions versus strategic interaction. The two basic distinctions combine. Their combination yields four basic types of complexity, leading to the table (1).

We now describe the details of table (1). Box 1, that is, the strategic consequences of choice interactions in well-defined task environments are relatively well understood. An approach for Box 2, that is, strategic interaction in well-structured task environments, analyzes the complexity of games in the sense of game theory (Shubik, 1997). In fact, even for well-structured tasks, there are issues of complexity, analogous to those of computational complexity in computer science. The latter refers to problems like chess or the traveling salesman problem with explicit rules that are exactly solvable in principle, but not in practice, because that requires much more computational resources than realistically available. Therefore, an important line of research in this field was devoted to finding search strategies that are simple and computationally inexpensive, but still are efficient in the sense that they find a reasonable (but not necessarily optimal) solution quickly. Such search strategies are called heuristics (Pearl, 1984), and the fact that heuristics also play an important role in strategy science in the context of Simon's bounded rationality is not accidental. In fact, Herbert Simon was not only a pioneer of modern organization science, but also influential in the early years of computer science and artificial intelligence, see (Newell and Simon, 1972, Newell and Simon, 1976). Importantly, whether a problem is well- or ill-structured depends on the perspective. The environments that are modelled by strategy scientists are well-structured almost by definition. For instance, the landscape models (see Section 2.3) present a model environment where all available options and their consequences are explicit. One can then let myopic agents (March, 1991, Levinthal and March, 1993) search in them for local or global optima. From the perspective of such an agent who cannot oversee the landscape, but at best only his local vicinity, it may then appear ill-structured. We may ask whether modeling myopic agents in a complex, but well-defined environment captures the challenges of actual strategists who need to construct representations of their task environments that are simply not symbolically described and therefore, by definition, ill-structured.

Landscape models have been criticized in (Felin et al., 2013), because in contrast to the finite setting assumed in such models, in reality, there are infinitely many unexplored combinations and potential innovations. From such a perspective, reality is always ill-structured. *But, and this is our crucial point, human agents can cope with ill-structured environments, and differences in this coping ability can lead to competitive advantages.* In general terms, learning means detecting structure, that is, making the situation better structured for the agent involved. And in modern computer science, a similar shift of perspective has taken place. That is seen in the field of machine learning. There, one wants to find and exploit structure in (typically very large) data sets that at the onset may lack any apparent structure. Again, the approach

[1]The analogy between algorithms, organizations and humans extends that of (Bettis, 2017) and is intended.

then needs something that one might want to call heuristics, that is, simple, but efficient tools for seeing and exploiting some structure. Many such heuristics have been developed, and important theoretical foundations have been provided. In the same manner that Simon exploited the analogy between human and computer problem solving (Newell and Simon, 1972, Newell and Simon, 1976, Simon, 1969/96), we now want to explore the analogy between modern machine learning and human techniques for finding structure and thereby rendering a situation less ill-structured. Because an ill-structured environment first needs to be rendered well-structured before responses can be formulated, and that applies to both quadrants 3 and 4, the study in this paper is different from studies that couple agents that jointly adapt on NK landscapes using markets (eg. (Lenox et al., 2006, Knudsen, Levinthal, Winter, 2014)). In these studies, the set of possible actions, and the interaction possibilities, although complex, are given in principle. We are interested in situations in which structure needs to be inferred before action and interaction possibilities are readily available and a competitive landscape can be explored.

Our theoretical discussion deeply resonates with fundamental questions in strategy. For instance, (Mintzberg, 1990) criticizes the design school of strategy, as presented in (Christensen et al., 1987) (although Mintzberg identifies earlier sources than those authors), for assuming that agents handle well-structured environments. In essence, he criticizes the assumption that the CEO can gather all information relevant to making a strategic decision and that he or she can compute the optimal course of action for the organization given this information. We argue that this assumption is not just problematic because of the bounds to the information processing abilities of CEOs. It is also problematic as environments may be ill-structured. As in ill-structured environments, by their definition, information cannot be readily processed, the actual cognitive task of strategists is fundamentally different. Strategists need to translate an uncertain and opaque environment into a representation and thereby translate an ill-structured environment into a well-structured simplification (Levinthal, 2011). We will discuss various insights from the sciences of complexity when and why such translation is successful.

More generally, the positioning view of strategy implicitly or explicitly assumes a well-structured environment. Our approach is different: In our view, it is a key aspect of dynamic capabilities to allow the translation of ill-structured problems into well-structured problems, because structure is needed to redeploy resources in a systematic way. Thus, dynamic capabilities may at their heart be tools to detect latent, and/or create new structure (see also (Baer, Dirks, and Nickerson, 2013)). Such structure may include interaction rules. For instance, the Digital Hub concept of Steve Jobs allowed a transition from a "quadrant 3" problem to a "quadrant 1 problem" (see table (1)): myriads of possibilities to think about digital lifestyle were rendered into a well structured set of possibilities by templates how portable devices and media content interact with personal computers (eg., the iTunes/ iPod concept.)

But the example also shows that metaheuristics to infer structure directly relate to taming the complexity arising from *strategic* interaction, too: Once structure is created, it can align actions of agents (including competitors), by organizing the interaction rules. For instance, the music publishers initially insisted on copy protection of .mp3 files when they agreed on iTunes (Renner, 2004), and Steve Jobs agreed and implemented this. However, the iTunes/iPod structure was so successful that copy protection of songs was abandoned in 2009. The logic of competition changed: iTunes is an attractive marketing device, but major revenue streams in the music industry today come from performances and streaming. So protecting songs from being copied is no longer effective to secure profits. The change of the logic of competition, however, was a result of the digital lifestyle transition that Apple co-created. Thus, the ability to detect structure in how digital lifestyle can be organized provides also means to create new interaction rules in the music industry - value chains transformed by a new idea how music can be embedded into everyday life devices. In the context of this article, the key message of this example is that metaheuristics to render a decision problem in quadrant 3 into a decision problem in quadrant 1 in Table (1) and often also means to render decision problems in quadrant 4 into decision problems in quadrant 2. In yet other words, given ill-structured problems, coping with choice complexity and coping with strategic complexity goes hand in hand.

Table (2) provides a summary of our arguments that are inspired from machine learning and their implications.

Table 2: **Principles from machine learning and their implications for search and creation**

Principle	**Implication**
Metaheuristic to identify type of structure	Find landscape type to select appropriate heuristic
Separate data from noise by penalizing complexity	Compromise between accuracy and simplicity of models
Adapt complexity of model to number and quality of observations	Start with simple models and gradually add more detail
Information bottleneck	Optimal compression increases efficiency

All principles in table (2) have the potential to provide theoretical reasons when and why some metaheuristics work successfully in the strategy context. Thus, these principles are candidates of theoretical explanations of effective dynamic capabilities. They may explain why some metaheuristics detect a change and allow to react appropriately, while other metaheuristics may fail. In the discussion section, we will return to table (2) and discuss in how far its entries can explain dynamic capabilities that are already discussed in the literature.

In the next section, we will discuss the basic concepts of bounded rationality, competitive landscapes, and complexity, to set the stage for our argument, followed by a discussion of simplification. We will then present newer insights from the sciences of complexity on how ill-structured situations can be rendered into well-structured situations, and apply these insights to strategy. This will be followed by a discussion of what these ideas imply for settings with strategic interaction. Finally, we conclude.

2 Rationality and complexity

2.1 Models

Modeling is an important part of modern strategy science. In particular, models are developed to understand learning and search processes of economic agents. Models are simplifications that aim at identifying and reproducing key qualitative features of a more complex world. From the perspective of the modeler, the model is a closed world in which all solutions and all available options are listed and can be evaluated and compared and which he can then study. This world is usually given in terms of a symbolic representation, like that of an NK-model (Levinthal, 1997), where the modeler sets N and K and the Boolean functions and therefore knows the resulting landscape. For a modeler, such a model world is well-structured, since explicitly symbolically represented. He can thus model or simulate how an agent acts in this model world, for instance an NK-landscape, that is known to the modeler, but only partially known to the agent. In particular, he can investigate how the agent's limitations lead to non-optimal behavior and how the agent can cope with that situation and also, how several agents might interact. In particular, he can study how agents construct their own models in such a model world and search and act according to them. Thereby, he can identify sources of error and give strategic advice. In the basic version, the model is independent of the actions of the agents. In a further step, he may also model how the actions of the agent affect the model parameters.
As we shall argue in Section 2.2, following Herbert Simon, economic agents also develop models. The two types of models, that of the modeler and that of the agent, differ in important ways. For such an

agent, the model is a simplification of the actual situation she is confronting.[2] While a modeler's purpose might be to better understand and explain certain features of the real world, the strategist might be most interested in the predictive power of her model. The mental model of an agent need not possess a symbolic representations. It may simply be based on certain statistical regularities and observed correlations. For the formation of the model, she may use certain implicit or explicit heuristics. While her model may have a certain stability, ideally it should also have enough flexibility to cope with changing circumstances. And this then becomes important in interactive situations where everybody involved develops models and behaves according to their models. Therefore, it becomes important to understand the models of others, their scope and their limitations. That is, an agent should develop a model of the models of others. In the context of SBR, this again requires simplification, and so, an agent will construct a simplified model of the simplified models of others. In practice, of course, this need not be explicit. Rather, agents use heuristics, for good reasons as we have argued, and so, in an interactive situation, an agent then might develop heuristics to guess the heuristics of others. And instead of simply trying to infer the models of others, an agent might rather try more generally to assess their modeling capabilities and limitations, and perhaps try to compare them with her own. That is, instead of "What is his model?", the question might be "Is he a better modeler than me, or is he likely to ignore or overlook relevant aspects?".

Again, whatever is guessed, modeled or inferred, importantly, this will almost inevitably involve some simplification. In particular, all the participants may have to cope with uncertainty, and it may then become uncertain how others handle uncertainty. This then creates additional, higher-level uncertainty.

In the terminology alluded to in the introduction and explained in more detail in Section 2.2, the above modeler's model is well-structured, because of its explicit symbolic representation. The only reason why the agent may not be able to find an optimal solution are her cognitive and computational limitations, that is, her bounded rationality, or more precisely, as we shall see below, one aspect of her bounded rationality. If equipped with more powerful resources, she could in principle become unboundedly rational. Real agents, however, cannot be unboundedly rational, not even in principle. The environment in which they operate is not given in form of a symbolic description (Wulff, Mergenthaler-Canseco, Hertwig, 2018). Rather, agents need to experience the environment, and to the extent that they try to create a symbolic description or representation of the environment, this is a part of their rational attempt to cope with it. But as argued already by Knight (Knight, 1921), an agent may simply operate with statistical regularities instead of symbolic representations. As representations, be they symbolic or not, are necessarily selective and incomplete, the rationality of agents can only be bounded - there are always contingencies that are overlooked when optimization schemes are applied.

In this contribution, we wish to explore the consequences of this, in particular in interactive contexts, and outline implications for strategy.

2.2 Simon's concept of *Bounded Rationality*

As (Gavetti and Levinthal, 2000) put it, the notion of bounded rationality (Simon, 1955) has been a cornerstone of organizational research (March and Simon, 1958/93, Cyert and March, 1963). In (Simon, 1969/96), p.166,[3] bounded rationality is defined as "The meaning of rationality in situations where the complexity of the environment is immensely greater than the computational power of the adaptive system."

The adjective "bounded" refers to limitations in terms of abilities, values, and knowledge (Simon, 1947/97), p.46. It carries a slightly negative connotation of being not fully rational. In fact, as (Simon, 1947/97), p.118, put it, "The central concern of administrative theory is with the boundary between the rational and the nonrational aspects of human behavior". (Simon, 1947/97), p.128, distinguishes between well- and ill-structured problems; problems are well-structured when the goals and the schemes for finding potential solutions are clear, and ill-structured otherwise. The concept of bounded rationality applies in either case. Even if the problem is well-structured, an individual or an organization may not possess the

[2]We will discuss below whether *awareness* of agents that their views are based on only imperfect models is a managerial capability that may lead to competitive advantages.

[3]Quotations and page numbers are always taken from the last edition listed in the bibliography.

ability to find an optimal solution (Bettis, 2017). In such situations, one may be content with satisficing, that is, finding a good enough solution, and one may use heuristic search strategies to arrive at such a solution. We shall have more to say about that below. In general, (Simon, 1957), p.198, an actor will simplify the problem, "The first consequence of the principle of bounded rationality is that the intended rationality of an actor requires him to construct a simplified model of the real situation in order to deal with it. He behaves rationally with respect to this model, [...] To predict his behavior we must understand the way in which this simplified model is constructed. " As a consequence, one might try to break up the problem into two steps. First, the construction of the simplified model, second, its solution. For the second step, standard methods of game theory might suffice, whereas for the first step, we would need other tools, in particular mechanisms of complexity reduction. As we shall see below, however, this is too simple. In situations of reciprocal bounded rationality in particular, the two steps are necessarily intertwined. Also, even if the problem posed by the simplified model may be well-structured, this second step may be computationally intractable and may therefore require the agents to find shortcuts, that is, to further simplify. This means that the concept of bounded rationality applies to both well- and ill-structured problems. Both the problem itself and the search for a solution can be simplified.

Whether a problem is, or better, appears well- or ill-structured depends on the perspective. The perspective of an agent may be different from that of the modeler. A problem that looks perfectly well-structured to a modeler may appear opaque and unclear to an agent who has to solve it because she lacks some relevant knowledge or understanding of the situation. One may also argue that any interactive real-world situation is ill-structured, because social reality is created by the framings by the agents, and interaction is only possible when some basic agreement about the norms and rules exists. Playing chess in a tournament may be well-structured, as long as the rules are not questioned, but playing chess with an infant might require some negotiations about those rules, and perhaps ways to cope with their violation. And while in soccer, it may be relatively clear what counts as a goal, it is often less clear what should count as a foul. That is, the agents need to agree on certain rules, and often also on mechanisms for their enforcement. These rules then serve to evaluate the performance of an agent and to provide criteria for sanctioning behavior that does not conform to them.

In the context of business strategy, agents need to agree on what value is before a performance landscape is definable. Before positions can be searched for, the framing of the transactional space needs to stabilise (Garud, Kumaraswamy, & Karne, 2010). And strategy needs a market for evaluating performance. Again, a market depends on agreements about values and rules, and those might be negotiable to a certain degree.

If we accept that real environments are typically ill-structured, at least from the perspective of the economic agents, we need to develop theoretical arguments that apply to such environments.

2.3 Landscapes

The notion of complexity arising from interactions of detailed choices is well-established in the strategy literature (Porter & Siggelkow, 2008, Gavetti and Levinthal, 2000). The basic question of these studies is how firms can establish advantageous positions in a landscape of possible strategies by superior search strategies. Two of the basic premises of this line of research are that the "landscape" of possible strategies is given, and that strategists' bounded rationality makes global, exhaustive search for the optimal strategy impossible. Rather, strategists need boundedly rational search strategies. And in the context of such a given landscape, agent behavior can be simulated, and the implications for understanding competitive advantages can be explored (see Baumann, Schmidt and Stieglitz, 2019 for a review).

The notion of such a landscape, however, did not originate in strategy science, but rather in theoretical biology. In fact, ever since their introduction by Sewall Wright in 1932 (Wright, 1931, Wright, 1932), fitness landscapes have been a fundamental metaphor in evolutionary biology (Gavrilets(2004)). Their importance and suggestive power is perhaps only surpassed by Darwin's phylogenetic trees. In neodarwinian evolutionary theory (one of whose architects Wright was), the fitness of an organism or the mean fitness of a population is determined bythe interaction of a number of discrete factors (the genes of evolutionary theory). Wright's visual metaphor replaces that typically quite large number of discrete factors by one or two continuous factors for which the graph of the fitness function can be readily visualized. That graph is the fitness landscape, and its local maxima are called peaks. The landscape could be flat without any peaks, smooth and single-peaked, or irregular and rugged with many peaks. Wright, in fact, viewed a typical fitness landscape as rugged with many local peaks of different heights that are separated by valleys of different depths.

According to this metaphor, biological organisms or populations try to find peaks by moving in the landscape through local adaptations (genetic mutations). Typically following the uphill direction of increasing fitness, they get stuck at local maxima, because even if a nearby peak is higher than that currently occupied, getting there might require to move through a deep valley, corresponding to low fitness values, and organisms of low fitness have little survival probability. This metaphor has been criticized for being too static. In fact, the actions and interactions of many organisms will affect the structure of the landscape, and therefore, one should consider landscapes undergoing dynamical changes.

Stuart Kauffman (Kauffman(1989)) then carried the metaphor further by identifying a class of landscapes that are very rugged, and where therefore local search is bound to lead to results that may be far from optimal. These are the NK landscapes. Here, at each time step, N factors, or in another interpretation, N agents act simultaneously, and the outcome of the action of an agent depends not only on that action itself, but also on those of $K-1$ other agents, according to some random Boolean function (the possible factor or agent values are assumed to be binary for simplicity). The iterated application of this rule can lead to dynamics that appear rather chaotic and unpredictable. And after Daniel Levinthal (Levinthal, 1997) had introduced NK-landscapes into strategy science, they became a leading metaphor there as well, for understanding the possibly rather chaotic and largely unpredictable consequences of many simultaneous actions or policy choices. Again, the contribution of a policy choice to the overall performance of a firm (the analogue of fitness) depends also on $K-1$ other choices. (Levinthal, 1997) used NK-landscapes to model the dynamics of populations of organizations that are all trying to find fitness peaks, and also how changes in the landscapes affect the resulting dynamics. (Gavetti and Levinthal, 2000) then modelled search processes of individual agents in NK-landscapes and systematically investigated the influence of the structure of such a landscape on cognitive and experiential search processes. Again, the visual metaphor replaces a high-dimensional discrete situation by the graph of a continuous function over the two-dimensional plane.[4]

For low values of K, the landscape is rather smooth, while for larger K, it becomes rather rugged, and there is little correlation between the fitness values even at nearby locations. In fact, even for low values of K, higher order correlations can build up through the iterated application of the dynamical rule (Ay et al.(to appear)). (Gavetti, 2012) proposes that strategists can overcome their limited local views by changing their cognitive representations, and he models that as jumps on the landscape of possible strategies that enable the agent to escape from the basins of attraction of local peaks.

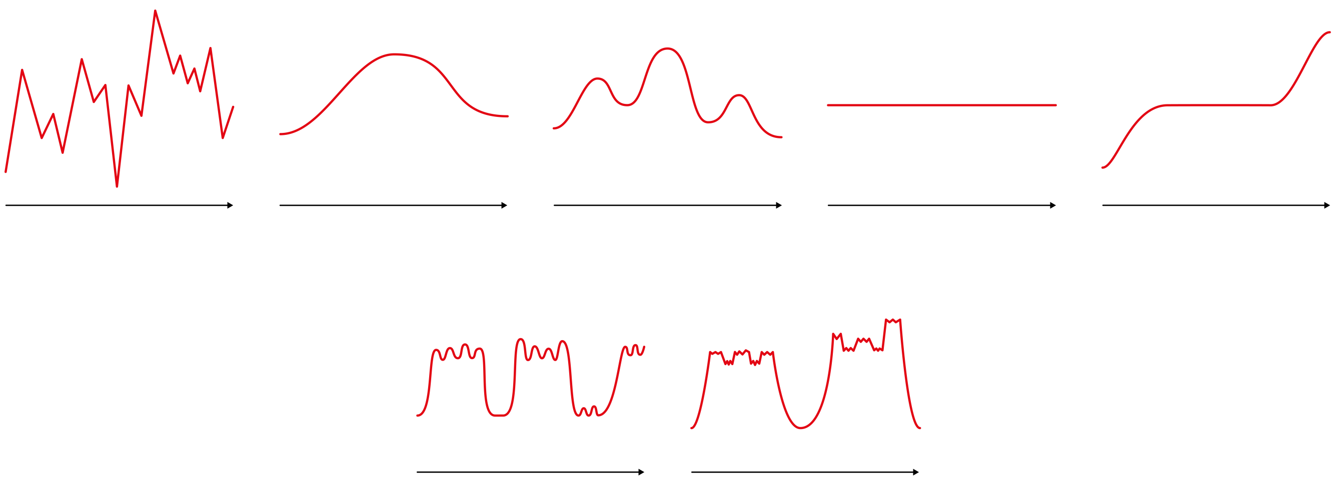
While keeping in mind its shortcomings, i.e., replacing a discrete high-dimensional structure by a continuous low-dimensional picture, the landscape metaphor is useful for illustrating a couple of further points. These are of a more general nature than those highlighted by the NK model.

In a general landscape, different regions might be quite dissimilar (see Figure (1)). Some parts may be rugged, others smooth with one or at most a few dominant peaks, still others may be flat without any peak, possess a large flat plateau with sharp edges (this would be a neutral region in the terminology of (Fontana et al., 1993), see (Jain and Kogut, 2014)), or exhibit a hierarchical structure.

Clearly, all these types require different strategies. In a rugged landscape, random actions may by

[4]We should point, however, that in the strategy literature, .e.g. (Gavetti and Levinthal, 2000), the term "landscape" typically refers to the discrete high-dimensional situation, that is, formally, a function defined on the vertices of an N-dimensional cube. The visual metaphor replaces this by the graph of a continuous function over the two-dimensional plane for purposes of illustration or in (Gavetti and Levinthal, 2000) also as a model of an agent's model of the landscape. The visual metaphor draws on the suggestive power of visual images in three-dimensional space. Being aware of its shortcomings, we nevertheless find it helpful to illustrate some important concepts. Our thinking is geometric, that is, adapted to three-dimensional space, and we find it difficult to visualize high-dimensional structures.

Figure 1: **Different types of landscapes**

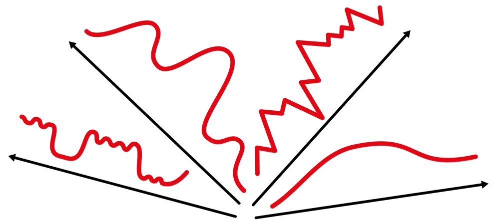


chance reach a good peak. In a single-peaked landscape, a simple hill climbing strategy works well. In a flat landscape, whatever one does has little consequences, and on a flat plateau, one can likewise freely experiment until one comes to an edge where one further change has dramatic consequences. In a hierarchical landscape, one should first adopt a coarse strategy to land in a good region and then explore that region by a more fine-grained strategy to find a local peak (see also (Csaszar and Levinthal, 2016) in this regard). In situations where several agents with different information search together, but along different dimensions of the landscape, (Knudsen and Srikanth, 2014) find that this can lead to mutual confusion or joint myopia.

There is an implication for strategy science here. Before employing a search heuristic in a landscape, like climb the next hill, make random jumps, cautiously explore your vicinity, or stay where you are, one first needs a *metaheuristic*, that is, a scheme for identifying the general structure of the landscape. Such a metaheuristic then tells the agent what search heuristics might be appropriate.
In the general theory of landscapes, there is the concept of the *correlation length* which quantifies up to what distance the structure of the landscape is correlated with that at the current position (see for instance (Stadler and Stephens, 2003))). Moving a distance larger than the correlation length brings an agent into a region that has little similarity with that where she came from, and the strategies that in her experience have worked well for searching the landscape will no longer be useful. Of course, when the landscape is not uniform, its correlation length may also be different in different regions. In any case, wherever an agent is, she should try to gauge the correlation length, that is, estimate how far her current strategic logic will remain applicable and when a radical change is required. And even if she is transported to a distant region in the landscape, she might still succeed with her strategic logic if that region of the landscape is structurally similar enough to the region where she comes from. The concept of correlation length may help us to better understand dynamic capabilities (Teece et al., 1997, Eisenhardt and Martin, 2000, Winter, 2003). Clearly, a strategist who can apply a metaheuristic to infer the correlation length in her task environment can also infer whether past wisdom should be applied in current decisions or not. While the strategy field is aware of the merits of forgetting since decades (Bettis and Prahalad, 1995), forgetting is a passive way of dealing with change. Being aware of the correlation length implies an active way of dealing with change in environments. Also, the landscape need not be isotropic, meaning that it may look different in different directions (see figure (2)).

Changing one factor may have a benign and smooth effect, changing another one may have very little effect, whereas changing a third one may lead to dramatic and highly fluctuating changes of the fitness. While of course the NK metaphor teaches us that the consequences of varying individual factors may not be independent of each other, in a general landscape one may still want to determine which factors can cause the largest or the least predictable changes. In general, this will not be a single factor, but a collection

Figure 2: **Different directions of a landscapes**



of dominant ones. In many situations, these may be rather few, and varying the others may have little consequences. Also, a multidimensional landscape with qualitatively different directions is not compatible with the positioning view of strategy (Porter, 1980) that views a strategy as a position in a one-dimensional price-quantity continuum.

There are two types of change that we need to distinguish here (Ehrig and Schmidt, 2019). One is the variability of a given landscape. In our landscape metaphor, the agent walks aroundand searches in a static landscape, but the landscape *may appear* changing as its local structure is different at different points in space. We have already discussed the correlation length as the scale on which correlations decay. And recalling Figure 1, the agent might come from smooth to a rugged portion of the landscape, or conversely. But there is also the possibility that the landscape itself actually changes. Then an agent may stay at his position while the landscape around him becomes different. That is, we should assign the dynamics to the landscape rather than to the agent. But again, the question is whether the new landscape is sufficiently correlated with the old one so that the agent's strategic logic is still adequate.

Thus, it might be more realistic to consider a situation where the agent stays, but the landscape changes its structure. From a formal perspective, however, it does not matter who or what is staying or changing. What is relevant is that the agent becomes exposed to a different landscape structure and needs to cope with that. Transporting the agent to a different region of the same landscape simply allows us to illustrate the qualitative effects of a change within the visual metaphor of a fixed landscape.

The preceding discussion can be rephrased as the difference between variations within a given landscape and the change of the landscape itself. When an agent is confronted with an unusual constellation, she may ask herself whether this simply represents extremal parameter values within her model class or whether this indicates a systematic change of that model class itself. This may have important consequences for her selection of an appropriate strategy.

Furthermore, there are the constructive aspects. Landscapes could be social or mental constructs, but agents might also shape landscapes economically. These issues have, of course, been discussed at length in the strategy literature (for instance, strategy as positioning vs. strategy as shaping), but this is not our topic.

The landscape metaphor is meaningful in the context of our boxes 1 and 3. More precisely, it is relevant for the distinction between those two boxes, that is, between well- and ill-structured problems. To the extent that the local landscape is known to an agent, the problem may be considered as well-structured, but search in unknown parts should appear as ill-structured. For the boxes 2 and 4, that is, strategic interactions, its static nature makes it ill-suited, and it needs to be extended. As noted, a similar criticism had been raised in evolutionary biology where a key issue is the dynamics resulting from interacting agents and species.

As already argued, we should make the landscape dynamic. Within that metaphor, the dynamics of landscapes is typically caused intrinsically by active organisms or by the process of evolution in the biological contexts, and by the activities of the strategic actors in the economic setting, rather than by external shocks.

(Felin et al., 2013) identified another shortcoming of the landscape metaphor, that it cannot address the frame problem. This means that in reality, the possibilities for innovation are limitless and exceed whatever is possible in some predefined landscape. According to their argument, innovation requires new conceptualizations beyond prestatable options. While invoking analogies with recent discussions in evolutionary biology, they do not provide guidance how to achieve that, however. Our perspective is somewhat different and, or so we claim, more constructive. We also want to move away from predesigned landscapes where agent behavior can be readily simulated by imposing some simple behavorial rules. From the perspective of the agent, the question rather is how to detect structure in complex settings that may appear as intransparent, opaque, and lacking apparent structure. But such a move, from a well-structured setting where complexity refers to computational limitations and therefore heuristics serve as search rules, to one where nobody knows what the structure, if any, is, has also occured in computer science. The most active current direction lies in machine learning, with tools like deep neural networks. A fundamental question is how to find structure in data. In that context, heuristics also play a new and different role. They are no longer search rules in well-defined environments, but methods for guessing appropriate structural priors with the help of which structure in data can be revealed. As we shall argue, this is also a way to move forward for strategy science.

2.4 Complexity – at last

The preceding has prepared the ground for a discussion of complexity in the context of this paper. We shall not attempt a formal and quantifiable definition of complexity here. Such a definition may be useful in computer science or machine learning, but when we draw analogies to problems in strategy science, the possibility to quantify things will usually get lost. In any case, the following is fundamental (Jost, 2004). On one hand, an agent is situated in an overwhelmingly complex environment, and he may want to capture and utilize as much as possible of that external complexity. On the other hand, his internal representation of that complexity should be as simple as possible, in order to use his limited resources in the most efficient manner.
In our context, we have to deal with rather different types of complexity.

1. In Box 1, the structure of the problem is clear, and what counts as a solution is not controversial, and one could in principle compute the solution if one had sufficient resources. But in practice, the required resources may be so demanding that an effective solution is computationally, organizationally or cognitively intractable.

2. In Box 4, the situation is opaque, and it is not clear who the key factors and players are, who knows what and what the options and the possible pitfalls are. But everybody else also is in the same situation, even though there may be variations in available information, resources and ability.

Boxes 2 and 3 are between those extremes, but as our boxing scheme indicates, complexity spans at least two different dimensions. In Box 2, as in Box 1, complexity stems from the difficulty or inability of solving a well-defined problem. In Box 3, complexity arises from a lack of understanding the structure of the problem. In Boxes 2 and 4, also the others agents are a source of complexity. We might therefore distinguish between *structural* and *strategic* complexity. Nevertheless, we shall try to find some general principles.

To handle complexity requires simplification. And different types of complexity require different types of simplification. Heuristics have been suggested as efficient simplification techniques both in strategy science (see the discussion in Section 3.1 and in artificial intelligence (see for instance (Pearl, 1984)). Our distinction between different types of complexity will also have consequences for the analysis of heuristics.

Importantly, simplification should not just be considered as a second-best option when the best option, a full solution of the problem, is not possible. In organizational research, also positive consequences of the fact that bounded rationality requires simplification have been explored. (Gigerenzer and Selten, 2001) pointed out that solutions found by simplified strategies like heuristics that typically ignore much of the available information often turn out to be better than those constructed by more elaborate schemes. As a possible reason, they suggest a bias-variance trade-off (Gigerenzer and Brighton, 2009). We propose the bottleneck principle (Tishby et al., 1999) as a deeper reason why searching under constraints like those of bounded rationality can lead to better solutions. In the strategy literature, *bottleneck* refers to a position in the supply chain where an agent is in a position to capture value (Gans & Ryall, 2017). In the machine learning context, however, the *bottleneck principle* tries to find an optimal balance between compression and prediction, that is, to achieve a maximally compressed mapping of the input variable that keeps as much information as possible about the output variable. The latter then forces the compression to be as efficient as possible, that is, to identify the most informative aspects and to suppress the rest. In strategy, this principle may apply in particular to research and development. More generally, computational, organizational, or cognitive limitations require an efficient utilization of resources. This may be the key to finding a good solution. This is a positive aspect that has been somewhat neglected in the prior discussion.

More generally, we shall argue that in opaque and uncertain situations, dealing with ill-structured problems in the sense of (Simon, 1947/97), simplification is often the only approach that could reasonably be called rational. In situations of strategic complexity, the key to success may no longer be the ability to outsmart others, for instance by ascending to a higher level of reasoning, but rather the ability to find the most efficient simplification. And in reciprocal situations that then implies that it is rational to assume tha all actors involved simplify. How such simplifications are carried out, however, may vary between actors. Therefore, more efficient simplifications can be a source of strategic advantage. Anticipating how others might simplify can be another source of advantage. Anticipating the collective effects of simplifications by all actors can be still another such source.

The important question then is how to best simplify. And we shall also encounter the metaquestion of finding a simple way to choose between different simplification strategies.

In contrast to rational analysis in clearly defined scenarios as in game theory, simplification is fallible. Therefore, it is a key strategic imperative to seek weaknesses in one's simplification schemes. These weaknesses can be of different types. A scheme could be too simple, or it might not be simple enough. We hope that our subsequent discussion of simplification strategies in other fields like machine learning by analogy can also provide insight into this issue in the context of strategy.

Sometimes, one can not only simplify one's model of the underlying structure, but also that structure itself. While we shall be mainly concerned with the former, we shall also see some instances of the latter, like platforms.

3 Simplification

3.1 Intractable problems and heuristics

As explained, Simon's concept of *bounded rationality* essentially means that an agent may lack the cognitive or computational resources to arrive at an optimal solution in practice, even if such a solution exists in theory. In fact, Simon saw the analogy with problems in artificial intelligence where also many problems cannot be optimally solved in practice, like playing chess, even if the rules are clear and explicit. In the practice of computer science, this applies to many problems like the famous traveling salesman problem where an agent or a computer program has to find the shortest roundtrip connecting a collection of cities. The information provided, the table of distances between all pairs of cities, is in principle sufficient to solve this problem, but in practice already for moderate numbers of cities, no computer is powerful enough to compute the optimal solution. Such problems are *computationally intractable*, and Bettis (Bettis, 2017) suggested the analogy of *organizationally intractable* problems in organization science. Again, a problem may be solvable in principle, but an organization may lack the resources to determine an optimal solution. Similarly, in psychological contexts, one may speak of *cognitively intractable* problems that humans with their limited cognitive resources cannot solve. This issue has profound implications. Importantly, the limitations that are captured by the concept of bounded rationality also have positive, constructive consequences. When faced with computational, organizational or cognitive limitations, agents are forced to make use of their limited resources in the best possible way. They therefore need to carefully distinguish the most relevant and salient aspects of the situation at hand and base their deliberations on those, while ignoring the rest. That is, they are forced to *reduce the complexity* of the situations they find themselves in, see (Jost, 2004). For concreteness, let us consider again the traveling salesman problem. If one had unlimited computational power, one could simply evaluate all possible routes through the cities. After one has evaluated all routes, one knows which one is the shortest. Obviously, this is a rather stupid, slow and inefficient way to solve the problem. But with unlimited resources, there would be no need to become efficient and quickly discard those routes that clearly will be very long. With limited computational power, however, one should search for strategies that make the search most efficient and that will quickly find short routes, even though the very shortest one might not be among them. For instance, one could use a greedy algorithm that at each step goes to the closest city not yet visited before. This may not find the best solution, but at least it should end up with a route that is not too long.

In dynamical contexts, it is also an issue that computation takes time, but during the time the computation is carried out, the situation may already change. Economic theory is mainly concerned with equilibria, and the resulting collapse of time makes it essentially blind to that aspect. But in reality, agents who extract the relevant features of a setting and reach a good solution quickly may have a crucial advantage over agents who attempt to plan for all contingencies and then search for the optimal solution.

The greedy algorithm just discussed for the traveling salesman problem is an example of a heuristic in the sense of (Pearl, 1984). Such a heuristic could be used not only by a computer program, but also by a human that has to solve a similar problem, like a real traveling salesman. (Bettis, 2017) suggested to view heuristics as strategies of complexity reduction by boundedly rational agents, and we shall explore some of the consequences of this and also try to take the next steps.

Heuristics try to identify and exploit the few most salient bits of information in a given situation, while ignoring further details, to reach a fast and frugal decision (Hafenbrädl et al., 2016). One of the reasons why heuristics often work remarkably well seems to be that they can detect and exploit subtle weak cues in the environment that are informative about some relevant issue. Many of these cues seem to be the result of collective processes, like market forces or public opinion, that for various reasons achieve a better information aggregation than individual forecasts.

It turns out that heuristics often lead to better decisions than more elaborate schemes that utilize more information or process the available information in more depth. It has also been argued that particular heuristics, special rules of thumb about, for instance, which business opportunities to explore, constitute some of the main immaterial assets of successful firms (Eisenhardt and Martin, 2000). While the term *heuristic* has somewhat different meanings and connotations in those research lines, the important aspect for our concern is that they constitute means to reduce the complexity of situations that the agents cannot fully grasp and to find ways to operate quickly and without having to devote too much effort to the process of reaching a decision or finding a solution. This is needed in situations where the agents cannot grasp all details or are overwhelmed by the amount of available information. In contexts of RBR, other agents will use such heuristics as well. This has two implications. First, one could try to infer what kind of heuristics other are utilizing and predict what their heuristics would yield, both to understand their way of thinking and for extracting forecasts without having to look at the situation oneself. Secondly, the collective use of heuristics may have collective effects, and that might be a source of both risk and opportunity. For an example in investment banking, see (Ehrig, Jost, Katsikopoulos and Gigerenzer, 2019). Heuristics can also miserably fail and produce inferior results, when not properly adapted to the context.

Whether a heuristic is successful, or at least stands a good chance of being successful, depends on the context and the circumstances. Therefore, (Gigerenzer and Selten, 2001) propose a theory of ecological rationality that delineates the conditions wherein particular heuristics (or other decision strategies) are better suited than others. Here, *metaheuristics* help to search for a good heuristic adapted to a specific context, in the same way that a heuristic searches for a good solution. Metaheuristics are heuristic rules to learn, or to select among, decision heuristics (Ehrig and Schmidt, 2019). This issue will be taken up below when we turn to the analogy with machine learning.

In machine learning, one is also often confronted with situations where the probabilities are not known, that is, situations of uncertainty instead of risk in Knight's terminology, but instead of accepting that as fate, one tries to infer the probabilities as well as one can by taking repeated samples. As will be explained in Section 4.2.1, theory (Vapnik, 2013) then tells us that the estimates should prefer simpler over more complex models, with technical precisions that are not relevant for our discussion. The balance between accuracy and simplicity may be delicate, but the point is that there exist compelling theoretical reasons for avoiding overly complex models, even if they provide a better fit to the data at hand. Thus, in that sense it is *rational* to build simple models. By analogy, in reciprocal situations, it may also be rational to assume that the others likewise prefer simpler models. More abstractly, in intransparent situations, one should adopt schemes of complexity reduction, and reciprocally assume that others do that as well.It is important for understanding strategy how much complexity can be reduced without arriving at models that are so simple that they grossly misrepresent the data is a question that cannot be answered in general terms, but depends on more precise circumstances. Therefore, there is still room for competitive advantages, in contrast to the rationality paradigm of game theory that fixes an equilibrium.

4 Learning and the boundary between ill- and well-structured problems

4.1 Simplification as a process

In contrast to the positioning view, the resource-based view of strategy focuses on the physical, human or organizational assets that a firm can employ for value-creating strategies. Such resources are not only valuable in static situations, in particular, when others do not possess and cannot copy them, but they also include abilities to adapt to changing circumstances. These are called dynamic capabilities, defined in (Eisenhardt and Martin, 2000) as "the organizational and strategic routines by which firms achieve new resource configurations" in changing markets. While the details of dynamic capabilities may be idiosyncratic, (Eisenhardt and Martin, 2000) also identify features that are common across firms. (Eisenhardt and Martin, 2000) also find that the behavorial rules representing dynamical capabilities are simpler in high-velocity than in moderately changing markets. They appeal to insights from behavorial learning theory (Argote, 1999,Haleblian and Finkelstein, 1999).
From our perspective, which is based on machine rather than behavorial learning theory, we can offer some conceptual clarification. The distinction between variability and change, which is well established in other fields (as an example, see (Rieke et al.(1997))) and which we have discussed in Section 2.3, is important here. In Section 4.2.1, based on statistical learning theory, we shall develop the general thesis that in more complex or opaque environments, models should be simpler than in environments that are better understood.
In (Eisenhardt and Martin, 2000), dynamical capabilities appear as routines. In slowly changing contexts, they are detailed and utilize existing knowledge, whereas in rapidly changing markets, they are simpler and rely on newly created knowledge. As routines, they are results of simplification processes. Here, we are interested in that simplification process itself, how it should be structured, what kind of cues and tricks it can possibly use, and in particular, how the learning patterns should evolve along with an increase in understanding.
Also, there is the question which dynamical capability is adequate under which circumstances. As we have argued, agents need metaheuristics to select among the available capabilities.

The distinction between well- and ill-structured environments depends not only on the perspective, as argued in Section 2.1, but is also fluid. In fact, learning is a process whereby agents discover and understand some of the structure of their environment, and as a result, the boundary between what is ill- and what is well-structured is shifted. At least from the perspective of the agent. As a consequence, also the dynamical capabilities may have to be adjusted as a result of learning that lets the environment appear better structured.

The traditional view of heuristics in artificial intelligence, as for instance in (Pearl, 1984), is to conceive of heuristics as search strategies for problems that are well-structured, but which for the agent are computationally intractable, like playing chess or solving the traveling salesman problem. The view of heuristics in strategy science seems similar. This is different in modern machine learning. There, one is often confronted with data sets that are ill-structured in the sense that nobody knows what structure they might contain. One then needs to make an educated guess at the type of structure. And if one has guessed the type of structure correctly and then applies the appropriate, often very powerful algorithms, one may uncover the details of that structure in the data. When the guess is wrong and therefore an inappropriate algorithm is applied, no detailed structure will be revealed. We shall discuss concrete examples below. The point we want to make here is that in the field of computer science (as the science encompassing both artificial intelligence and machine learning), the problems have changed, from search in well-structured, but computationally intractable model problems to the analysis of data sets where structure needs to be found. We wish to argue that an analogous shift will be insightful for strategy science, as it should bring it closer to real world problems. And since machine learning has developed some powerful tools to deal with such problems, there is hope that by analogy, strategy science can succeed as well.

4.2 Learning

Learning in and from ill-structured data sets is a basic problem addressed in the field of statistical and machine learning theory. Researchers have come up with a number of principles. Since there is some analogy between such learning tasks in data and in strategy science, we wish to explore to what extent those principles can also be useful in strategy science. Here, we discuss three such abstract principles for learning models from data.

1. Control the number of degrees of freedom of your model class
2. Identify appropriate structural assumptions
3. Penalize complexity

We will discuss them now. In order to illustrate those principles in more concrete terms, let us consider the following problem. A strategist observes some market movements, like prices, sales, regional differences etc. and in order to understand this better and make reasonable predictions, he wants to build a model with some factors and dominant actors causing those movements.

4.2.1 Number of degrees of freedom, or bias vs. variance

Every model is taken from some class. In our example, the parameters of the model class could be the number of factors and actors and their abilities. According to learning theory (see for instance (Vapnik, 2013)), the expected error of any prediction has two parts, the approximation error (bias) and the sample error (variance). The approximation error comes from having too few parameters so that the model cannot match the observations well. The sample error comes from allowing for too many parameters, resulting in overfitting the observed data. The statistical learning theory of (Vapnik, 2013) can estimate these two errors. The essential result is that in order to keep the prediction error, that is, the sum of these two components, small, the model class should adapt to the number of independent observations. The number of parameters should grow with the number of observations, to decrease the approximation error, but more slowly than that number, to keep the sample error under control.

Let us provide some more details for a central example. Our discussion will first be technical. We will present an important result from statistical learning theory and then apply it to strategic settings.

We want to estimate an unknown relation between an input x and an output y. x might come from any space X, but for simplicity, we assume that the output y is a real number. We draw samples (x_i, y_i), and on the basis of those samples, we want to construct a function $f(x) = y$. The samples are drawn from an unknown distribution p. The crucial object is the error or risk functional

$$R(f) = \int_{X\times\mathbb{R}} (f(x) - y)^2 p(x, y) dx dy. \tag{1}$$

Of course, the more samples we draw, the more information we have about p and the true relation between the variables x and y, and the more accurate our model f could become. From the samples, we obtain the empirical risk functional

$$R_{\text{emp}}(f) := \frac{1}{m}\sum_{i=1}^{m}(f(x_i) - y_i)^2. \tag{2}$$

Naively, we could thus choose an f that makes $R_{\text{emp}}(f) = 0$, that is, satisfies $f(x_i) = y_i$ for all $i = 1, \dots, m$. But this is not a good idea, as this leads to overfitting, as should become clear in a moment. After all, we want to make R small, and not R_{emp}. That is, the estimated f should generalize well to future samples. The solution requires to constrain f to avoid overfitting. That is, one specifies some model class from which one chooses f, and the model class should be large enough to account reasonably well for the observed samples, but also not too large so as not contain too complicated functions f. This can be made precise. The key result is that the number of degrees of freedom in the set of functions from which f is chosen must asymptotically be smaller than the number m of observations. The number of degrees of freedom is measured by the Vapnik-Chervonenkis (VC) dimension, a technical concept that, roughly speaking, quantifies how many distinctions the model class allows for. The details are not essential for our qualitative discussion, and we refer to (Vapnik, 2013).

Of course, since we are in a probabilistic setting, the risk R in (1) can be bounded only with a certain probability, and the bound gets worse if one demands higher probability. We want to have the bound satisfied with probability at least

$$1 - \eta \,,$$

for some small η, that is, we want to have the bound with high confidence. We put

$$\mathcal{E} := \frac{4h}{m}\left(\log\frac{2m}{h} + 1\right) - \frac{4}{m}\log\frac{\eta}{4}\,, \tag{3}$$

where h is the VC-dimension of the class of functions from which we choose that f that minimizes the empirical risk in (2). $\mathcal{E}$ depends on one hand on η, and it becomes large when η is so small, but because of the term $\frac{4}{m}$, this effect can be mitigated by choosing a large number m of samples. On the other hand, it depends on the ratio $\frac{h}{m}$ between the VC-dimension h, that is, essentially the complexity of the model class from which our estimator f will be drawn, and the number m of samples.

The fundamental estimate of statistical learning theory (Vapnik, 2013) then says that with probability at least $1 - \eta$, we have

$$\begin{aligned} R(f) \quad &\leq \quad R_{\text{emp}}(f) + \frac{B\mathcal{E}}{2}\left(1 + \sqrt{1 + \frac{4R_{\text{emp}}(f)}{B\mathcal{E}}}\right) && (4)\\ &\leq \quad 2R_{\text{emp}}(f) + 2B\mathcal{E}\,. && (5) \end{aligned}$$

Here, B is some general bound for the function class whose precise value is not important for our discussion. The inequality (5) is the simpler one, so let us discuss it first. The first term in (5) controls the empirical risk and depends only on the particular function f that minimizes the empirical risk within the prescribed class whereas the second term includes the confidence, since $\mathcal{E}$ in (3) includes η, and it depends on the entire class and grows if the VC-dimension h grows and is large if $\frac{m}{h}$ is small. Thus, the confidence is small if there are few observations but a large choice of models. Thus, in order to make $R(f)$ small, one has to balance the two terms in this inequality. One has to make the class large enough so that the empirical risk becomes small, but not so large that the confidence gets too small. Looking at the more precise inequality (4), one would like the expected loss, that is, the predictive ability of the estimator f, not much larger than the empirical loss. Therefore, $\mathcal{E}$ needs to be small. In particular, looking at (3), this means that h should be substantially smaller than m.

Thus, the best choice of the VC-dimension h depends on the sample size m. When we have more samples, that is, possess more information, we can make h correspondingly larger, so as to decrease the empirical risk while still keeping $\mathcal{E}$ small. That is, the model can then become more complex. Importantly, we should see this as a process. We collect samples, and by that, we increase our information and can then construct more accurate models with higher confidence.

For instance, when we want to approximate the graph of an unknown function from which we collect samples by a polynomial, we should control the maximal degree k of the polynomial. When we have m data points, then k should be significantly smaller than m (the VC-dimension is of the order of k in this example). On the other hand, when we keep k too small, for instance $k = 1$, that is, when we want to approximate the graph by a straight line, the empirical loss might be quite large. Thus, the degree of the polynomial should grow with the number of samples m, but only moderately.

A special case (a discussion of a linear relation between x and y) of the general insight presented here is discussed in (Ehrig and Schmidt, 2019), and applied to the problem of learning heuristics in firms. Moreover, the here discussed principle explains the success of the take-the-best heuristics discovered by Gigerenzer and colleagues (Gigerenzer and Brighton, 2009).

Let us consider another example that has been discussed in the finance literature (DeMiguel, Garlappi, & Uppal, 2007). The famous Markowicz rule yields an optimal mixture of n items in a portfolio when the relevant expected pay-offs and risks are known. That formula is rather complicated, and Markowicz himself is famously known to utilize for his own investments the simple $\frac{1}{n}$ heuristic, that is, split his investment evenly between the n items. Now, the above discussion tells us that while the Markowicz rule may be too complicated, the $\frac{1}{n}$ rule might be too simple, and a better rule should be between those extremes, that is, use a modest number of parameters that are matched to the available data.

We can now explore some general implications for simplification in the context of strategic decisions.

Thesis 1. *The less you understand, or the more complex the situation appears, the simpler your model should be.*

and

Thesis 2. *Only when you understand the situation better and know more facts, your model should gradually become more detailed.*

In fact, this principle is also supported by a major insight in developmental psychology. When children learn a language (see for instance (Tomasello, 2003)), they also start with very simple models. During a critical phase, they overgeneralize the rules they have identified. Only later, when exposed to more data, they gradually refine the model. Thus, language learning by infants becomes a remarkably efficient process. Returning to strategy, look at the simple example of a market where actors are just buying and selling, and the relevant parameter then is the number of actors. In our example this means that when only few market observations are available, the model should postulate only a small number of key actors or factors, and only when the number of those observations grows, one should gradually allow for more. When we assume that this principle remains valid in interactive situations, we can also apply it to the observation in (Ehrig, Jost, Katsikopoulos and Gigerenzer, 2019) that models used by bankers typically take only a small number of dominant actors into account. This can then be justified not only in terms of cognitive limitations, as it becomes mind boggling to mentally juggle with a large number of agents and their complicated interactions, but also from general principles of learning theory.

4.2.2 Structural assumptions

But let us suppose now that our agent who wants to understand and predict the market behavior has many observations at his disposal, or more abstractly, a large data set. Should he now try to build a model with many actors and factors and complex interactions between them, in order to most accurately account for all the details in the data?

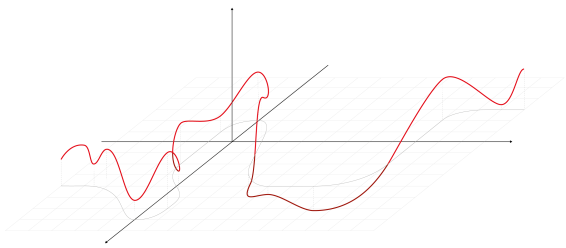
Abundant experience in machine learning tells us that the answer is an emphatic *No*. And the reason for that *No* are not computational limitations, but the insight that in order to find a good model, one should make stringent structural assumptions about the underlying data set. In our example, such an assumption might simply be that are only few dominant actors and factors. The precise number need not be specified, nor does one need to impose an explicit bound on that number. One simply works with a model class that makes the a priori assumption that there are only few sources. In machine learning, this approach is called *compressed* (or *compressive*) *sensing* (Foucard and Rauhut, 2013). It was introduced in (Candès et al., 2006, Donoho, 2006). The motivating example is the following (whose analogy with our example is easily seen). Consider a complex auditory scene, for instance all the speech and the noise recorded in a party room. With traditional methods like Fourier analysis, this is very difficult to analyze, but if one uses a scheme that is based on the assumption that there are only a few auditory sources (speakers), then those can be readily identified and the data make sense. The rest then is simply unexplained noise.

The method of compressed sensing works only under suitable circumstances, like those of our example. In other situations, it is inappropriate and fails. But other schemes might work wonderfully then. The question is which such scheme to apply. In machine learning, many schemes have been developed that work well under appropriate conditions. But when confronted with an intransparent data set, the question then is which scheme to choose. Analogously, in a strategic situation, it is not the question whether heuristics or other simplifications should be used, but rather, how much one should simplify. Let us draw upon some other examples from machine learning to illustrate this point (see (Jost, 2017) for a more general analysis). Machine learning is a field at the intersection of computer science, high dimensional geometry and statistics that is currently rapidly developing. Much of current machine learning is concerned with detecting, extracting and utilizing structure in big data sets. Such big data sets are typically characterized by a couple of Vs, including high volume, velocity (with which they come in), variety (that is, heterogeneity of data types, representation, and semantic interpretation), variability (of their quality and reliability) etc. These properties make a complete and exhaustive analysis impossible. Therefore, one needs mechanisms of simplification.

Of course, there are some well known principles, interpolation (if you have two data points at different times, assume that the data at intermediate times should be averages of those) and extrapolation (assume that a trend continues). Machine learning refines these principles. Many successful machine learning algorithms rely on general structural assumptions about the data set at hand (Jost, 2017). They work well, and often surprisingly well, when those structural assumptions are satisfied, but fail when not. Examples include

- Manifold learning (Belkin and Niyogi, 2003) assumes that the data while coming from a high-dimensional space are intrinsically concentrated on a low-dimensional smooth manifold, that is, without corners, edges etc (see figure (3)).

  Figure 3: **A smooth one-dimensional curve stretching into different directions in 3-space**

  Thus, there are two intertwined structural assumptions. The first one says that there are only few intrinsic degrees of freedom, although the data may appear high-dimensional. It is then the task to identify those important degrees of freedom. One can then ignore variations in other directions. The second assumption is that those degrees of freedom are regular and smooth. In our example, when the dominant actors have been identified, one may want to infer their strategies. Even though the

strategic space might be huge, each actor might only use a small number of parameters to determine her strategy. These may vary according to circumstances, but one might assume that changes are not abrupt and sudden, but rather gradual and smooth. That would be analogous to the situation of manifold learning.
While we want to discuss here manifold learning simply for purposes of analogy, it could also be used concretely for purposes of data analysis in business environments. Let us consider products with many different manufacturing degrees of freedom, like cars or clothes. The important degrees of freedom might vary between different market segments, for instance between luxury cars, sport sedans or family vans. In some segments, engine power is important, whereas in other segments, customers are interested in a large selection of colors, or gasoline consumption might be an issue. But neither of those seems to have a high relevance in the luxury car segment, where customers value a different set of qualities and options. In any segment, probably only a small number of features is relevant, but it will vary between segments which those are. Thus, the consumer choice data may be concentrated on an intrinsically low dimensional manifold, and a manifold learning scheme should be able to identify that manifold in the high dimensional space of possible factor combinations.
More abstractly, the essential point is the following. When the manifold on which the data are concentrated is low dimensional, that means that on each portion of it, only a small number of features is relevant, and one may project on the corresponding dimensions. But which these dimensions are may vary along the manifold. Therefore, methods that seek to identify the same dimensions on the entire space, like PCA or ICA, may not capture that variability, and may fail to give an accurate picture. Similarly, when agents project a landscape onto a lower dimensional model, as in (Gavetti and Levinthal, 2000), they may miss this important aspect. Everywhere in the landscape, only some dimensions may be relevant, but which these are may vary across the landscape.

- We have already discussed compressed sensing (Candès et al., 2006,Donoho, 2006,Foucard and Rauhut, 2013) which similarly assumes that there are only few sources that generate the data. This assumption leads to algorithms with a performance that is superior to others that may fail to identify any source because they are postulating too many. As discussed, in a strategic situation, one might analogously assume that there are only a few actors that have a significant influence on the market. One should try to identify and model those and relegate the rest to some kind of noise term.

- Hierarchical tensor decompositions (Hackbusch, 2012) suppose that there is a hierarchical structure underlying the data, and identifying that structure leads to much more compact representation. In strategic situations, one might similarly assume that many local variations are coming from some common dominant influence, and they should therefore be correlated. The task then is to identify that dominant cause. For instance, within each country, region, or market segment, agents might behave similarly. One should then first make a coarse grouping according to those factors, and then perhaps proceed to a more fine-grained analysis of agents within such a group.

- Multiscale methods (Pavliotis and Stuart, 2008) assume an interplay between a coarse global and a fine local structure. In economic situations, we of course have the interplay between prices as global market variables and individual behavior. The idea of the multiscale methods is that one can switch back and forth between the different scales for computational efficiency. For instance, for long stretches of time, it might suffice to observe the dynamics of market prices, and only every now and then a more detailed investigation of changing customer preferences is needed.

- More generally, one may suppose symmetries, invariances and other regularities. This is also what makes the gestalt principles of cognitive psychology work (Breidbach and Jost, 2006).

In order to select such assumptions and the computational schemes based on them, usually some prior knowledge of the domain the data are coming from is needed. It is then an important research question (Jost, 2017) to find a more systematic and formal way to identify such general structures in a data set.

4.2.3 Penalizing complexity

In many practical problems of data analysis, the essential issue is the distinction between noise and structure. When the model is too simple, it does not reflect the data well, and when the model is too complex, it pays too much attention to the noise. Both negatively affect the ability to generalize. A theory that makes more general predictions is a better theory. Of course, given the samples, one does not know what is noise and what is the underlying structure. So, what should one do? In abstract terms, if one does not have good prior insight that restricts the model class, one may try to minimize a combination of the data fit and the model complexity. Thus, one would minimize a combination of a term that measures the data fit and another term that evaluates the complexity of the model. This, in fact, is a widely applied general principle, although in practice the complexity measures is often chosen in a somewhat ad hoc manner.

Let us take an example from image denoising. There the problem is to recover the original image from a version that is corrupted by noise. That version is all that one has. Which of its pixels should be kept, and which should to be changed because they might represent noise? Variational methods, see for instance (Jin et al., 2015), minimize a combination of two terms. One is a fidelity term that quantifies the difference between the image that one has and the image that the scheme produces as an approximation of the unknown original. The other is a regularity term that penalizes contrasts between adjacent pixels. The underlying assumption is that the original should contain many rather homogeneous parts with little fluctuations and that therefore most contrasts should originate from the noise. That is, one assumes that the original is smoother and more regular than the noise corrupted version. If the fidelity term is given too much weight, the noise is not much eliminated. If the regularity term is given too much weight, then all pixels are similar, and the essential structure is lost. The original image was most likely not completely homogeneous and regular, but had some edges and boundaries between different segments, but inside each segment, we should not expect wild fluctuations between adjacent pixels. Thus, one needs to find the right balance. That usually requires experience with such data sets. The regularity term may be considered as a complexity term, because images with too many local fluctuations are difficult to compress and therefore require more computer memory space than more regular ones.

What can we learn from that example? Overly complex models may not represent the underlying structure well. Here, (in contrast to other places in this paper), *complex* simply means having little regularity and exhibiting many local fluctuations. Therefore, one should penalize complex models. But how much they should be penalized, that is, how simple or complex the underlying structure is, cannot be decided by general arguments, but requires experience. The question is to identify the right level of complexity of the model to best reflect the underlying structure.

4.2.4 Correlations

Although strategy science should be concerned with causations rather than correlations, as emphasized for instance in (Bettis, 2017), detecting correlations can be an important source of information. Also, the power of some heuristics or of intuition may depend on the utilization of certain weak correlations. For instance, when selecting that stock for investment that has seen more recent media coverage one may be utilizing some weak correlations between media attention and future performance. The reason for a correlation between the future prospects of a company and its mention in some news medium may be difficult to trace in each individual case, but is quite plausible that it should exist at least on average. Also, the direction of causality, if any, may remain opaque in such cases, but that might not matter for a good prognosis. That is, it does not matter whether media attention positively influences future performance, by whatever mechanisms, or whether the journalists are clever enough to identify potential high performers.

The strength of deep neural networks and other such machine learning algorithms also rests, at least partly, on the implicit detection of certain weak correlations. If this is so, then the reasoning of the preceding sections about the analogy between heuristics in strategy and structural assumptions in machine learning is strengthened further.

5 Strategic Bounded Rationality

In strategic situations, boundedly rational agents encounter, either directly or indirectly, other agents that are also boundedly rational. These other agents could be customers, suppliers or competitors with which one interacts directly. Or the bounded rationality of other agents could be indirectly encountered, for instance through collective effects in markets. And since, as we have argued, in situations of uncertainty, in ill-structured environments or even in well-structured environments that are too complex, bounded rationality requires agents to simplify, this also has the consequence that in strategic contexts, agents need to deal with other simplifying agents.

The concept of rationality, as used for instance in rational choice theory, has a normative component. The concept of bounded rationality seems to abandon that normative component, as it has the consequence that real agents cannot achieve optimal solutions even in well-structured environments when the task is too complex in view of the limitations on their cognitive abilities and computational resources. In ill-structured environments, however, as we have argued, there is no such thing as an optimal solution that could be computed in principle if one had sufficient resources. Therefore, if the concept is to retain a normative component, we need to proceed differently. The best that one can do is to use the theoretically available tools to deal with ill-structured environments. As we have argued, these include heuristics, or more abstractly, appropriate schemes of complexity reduction. In short, agents need to simplify. We have described theoretical insights from statistical learning theory and related fields, and we have also cited some relevant literature from psychology about the appropriate use of heuristics (Gigerenzer and Selten, 2001, Gigerenzer and Brighton, 2009) in opaque situations. Instead of the impossible norm of finding an optimal solution, agents should find simplification schemes that reflect such theoretical insights. Thus, on the basis of their abilities, their available knowledge and information, they should develop models that capture the essential aspects and suppress the others. Since we have conceptualized this as a process the models need to be updated and enriched when information gets richer and knowledge becomes more secure.

This is fundamentally different from the standard approach of game theory and neoclassical economics. Agents that are learning, searching and or making decisions in a context of incomplete and evolving knowledge and limited computational, organizational or cognitive resources should use the best available, applicable and implementable inference rules. Importantly, under those conditions, these rules suggest not to incorporate all available details, but rather to start from simpler models. Leaving out details makes these rules fallible, but the point is that in contexts of uncertainty, this is unavoidable. Therefore, one should take into account the own fallibility as well as that of others. One consequence of this is that agents will typically run into contradictions and surprise.

We suggest to use insights from machine learning that in complex situations, the best strategy is to develop simple models. What does that mean in strategic contexts? As yet, strategic situations don't seem to be covered by machine learning, at least not systematically. Our answer to the question is based on the simple observation that in strategic interactions, all agents involved are boundedly rational and therefore need to reduce complexity and simplify. Strategic bounded rationality needs to build on that. Thus, an agent should not only assume that others build simple models as well, but also needs to decide how to model that in turn. Again, this requires simplification. One solution is to simply relegate other actors to a collective force, like the market. In many contexts, this is appropriate because the market collects information and represents that information by simple variables, the market prices. In contrast, in interactive situations with a few actors, like in oligopolies, one needs some simplified models to model the modelling of others. For instance, one could develop heuristics to gauge the modelling abilities of competitors. While we start here to develop some consequences of this, there is much room for future research.

When real agents that are boundedly rational interact with other agents that are likewise boundedly rational, they find themselves in a situation of reciprocal bounded rationality (**RBR** for short). Reciprocal bounded rationality means the mutual understanding of potential interaction partners that everyone involved in an interaction under uncertainty is limited in access to information, computation time, and computational abilities. This implies a mutual understanding that one's partners or competitors also need to use mechanisms of complexity reduction, like heuristics. RBR is a special case of SBR. RBR is more special, as it implies that agents are aware that all involved agents are aware of their limitations. Note that here limitations are not necessarily negative. Given complexity, no involved agent can deduce an unlimited rational response to others' actions, expect for situations that are well-defined and for which fixed point solutions are available (such as in standard game theory). Given their limitations, the involved agents are modellers (they do not "see" the ultimate reality but only construct a model of it, relative to which their responses are rational; Levinthal, 2011). Advantages then can come about by being a better modeller, so dealing more wisely with one's own limitations.

In Section 2.1, we have introduced a distinction between the model of a modeller and that of an agent. Both models are simplifications, but in importantly different ways. We need to take the perspective of the agents. In interactive situations, all involved agents develop models and behave according to them. Therefore, each agent should develop a model of the models of others, or better, of the modelling processes of others, and this will include their capabilities and limitations. As we have pointed out, this requires simplification, leading to the construction of a simplified model of the simplified models of others. And as argued in Section 2.1, this need not be symbolic and explicit, but could be based on heuristics. Thus, an agent then might develop heuristics to guess the heuristics of others. Since this will inevitably involve some simplification, it will create uncertainty. In particular, it becomes uncertain how others handle uncertainty. This then creates additional, higher-level uncertainty. This cannot be fully resolved.
On one hand, simplifications reduce subjective uncertainty, insofar as a mental model may possess more explicit structure or regularities than the actual world. On the other hand, simplifications create new uncertainty because nobody knows for sure how the others simplify.

What are the consequences of all this?

1. Obviously, boundedly rational agents can make mistakes. One can then try to exploit the mistakes, but at the same time needs to be aware of the possibility of being outsmarted or exploited by others (Menon, 2018). Therefore, agents need to find schemes to gauge their opponents. However, in contrast to (Menon, 2018), our argument implies that more sophistication in modelling others *does not* necessarily yield competitive advantages. For instance, it is *not* always beneficial to reason in more iterations about the reasoning of others (Ehrig, Manjali, Singh, Sunder, 2019).

2. Human agents are often quite good in extracting weak cues from the environment. In contexts of RBR, other agents will use such heuristics as well. This has two implications. First, one could try to infer what kind of heuristics other are utilizing and predict what their heuristics would yield, both to understand their way of thinking and for extracting forecasts without having to look at the situation oneself. Secondly, the collective use of heuristics may have collective effects, and that might be a source of both risk and opportunity. For an example in investment banking, see (Ehrig, Jost, Katsikopoulos and Gigerenzer, 2019).

3. The preceding concerns learning, but for search, something similar applies. Instead of laboriously searching for a solution oneself and developing an appropriate search strategy, one can try to copy the solutions found by others or imitate their search strategies. Yet there are limitations to this (Ryall, 2009), in particular as it may not be perfectly observable how competitors arrived at their strategies or the details of their execution. In any case, joint search by boundedly rational agents may not lead to the optimal outcome (Knudsen and Srikanth, 2014). This has important consequences when outcomes depend on the collective search of several, and perhaps many, potentially competing agents. Factor market prices are a fundamental example. Prices are determined by expectations (Barney, 1986). In the context of RBR, agents need not only be aware of their own limitations, but also need to grapple with the imperfections of other agents. Sometimes, prices rather reflect others' ideas about the future than a fundamental reality (Hong, Scheinkman & Xiong, 2006). Search processes of competing boundedly rational agents have been addressed in prior work (Lenox et al., 2006, Knudsen, Levinthal, Winter, 2014), but the existing studies do not address how agents anticipate each other's expectations

and actions. (Ehrig, Jost, Katsikopoulos and Gigerenzer, 2019) empirically documented how strategists in the banking sector cope with this problem.

In short, besides constructing simplified, perhaps game-like representations, investment bankers gauge their opponents. That is, insteading of modelling an opponent, they rather ask “Is he better than me? Does he have information sources or cognitive abilities that I do not possess?” They utilize weak cues, and they consider market agreement as a source of both risks and opportunities. Furthermore, on longer time scales, central bankers construct narratives that involve putative causal agents like “the market” or “the financial sector”.

As explained in Section 4.2.1 there is theoretical insight (Vapnik, 2013) why simpler, that is, less complex models often have better generalization power. And this not only leads to a rather general heuristic, “choose the simplest possible model that fits the data reasonably well”, but even explains why this is a useful principle that works better than choosing more complex models that fit the data better.
Importantly, that then is a strategy with a rigorous theoretical foundation. Therefore, in reciprocal situations, one should also assume that others apply that strategy. This is a very important point worth emphasizing. Avoiding complex explanations and choosing simple ones instead is what one should rationally do when dealing with intransparent or noisy data sets. And therefore, everybody should rationally do that even in competitive situations. Of course, since this is a somewhat loose principle, it does not completely determine the precise model that is best in a given context, and therefore, there still is room for competitive advantages from selecting a better model.

Applied to strategic interaction, our arguments about simplification also imply that the transition from ill-structured problems to well-structured problems in choice interactions can go hand in hand with the same transition for strategic interactions. Metaheuristics can help to render a situation in quadrant 4 in Table (1) into a situation in quadrant 2. And, even more importantly, some metaheuristics seem to render a set of problems in quadrants 3 and 4 into a set of problems in quadrants 1 and 2. One such metaheuristic seems to be the construction of a strategic representation (Ehrig, Jost, Warglien, 2019). For instance, a very successful strategic representation is the digital hub structure (Jobs, 2011). It changed the possibilities in which media capture and playing devices interact with computers and thereby enabled their embedding into the everyday life of customers. Moreover, the digital hub structure also implied a new logic of competitive interaction. The creator of the platform, Apple, gained bargaining power over suppliers of components of the assembled devices (touchscreens etc) and suppliers of content (music etc). In other words, simplifications organize choice interactions *and* strategic interactions simultaneously.

6 Conclusions

Strategy is more than just establishing positions (Porter, 1980), or accumulating resources that provide sustainable advantages *in a given market*. Processes of adaptation or shaping, in particular the discovery or creation of novel opportunities, may be more important for a firm’s survival than securing positions in a well-structured, given market (Teece et al., 1997). (Eisenhardt and Martin, 2000) started to argue that to understand processes of shaping and adaptation theoretically, we need to draw on newer insights from the sciences of complexity. The contribution of this article is the development of such theoretical foundations. We discussed that simplification involves metaheuristics to detect structure. Such metaheuristics may be considered dynamic capabilities, as they enable the transition from ill-structured problems to well-structured problems for which the appropriate deployment of resources of a firm is known. In the following, we will discuss the implications of our framework for future research in the area of strategic management.

Strategy as the Process of Simplification Often, the strategy of a firm itself can be understood as an ongoing pattern of simplification (Mintzberg, 1978, Miller, 1993). Interdependencies across choices, agents, and different time scales involved in a decision create the need for specific strategic skills (Leiblein et al., 2018). Our argument rests on the thesis that these interdependencies create complexity, and skills to reduce the complexity can be the essence of strategic skills or even lie at the heart of a firm’s strategy. In other words, the need for complexity reduction renders decisions ”strategic”.

This "coreness" of complexity reduction has broader implications for our understanding of strategy. For instance, in how far can a strategy be understood as a position? While a strategy may *result* in an ex post perceived *position*, ex ante a strategy can often not be conceptualized as *taking* a position, because the number of possible strategies is vaster than could be captured by a finite-dimensional landscape (Felin et al., 2013). Before positions can be taken, both in a price-quality continuum (Porter, 1980) or in a space of interdependent choice alternatives (Porter & Siggelkow, 2008), strategists need to simplify. Success in establishing ex post perceived positions in a then perceivable landscape of possible strategists essentially hinges upon skill in the simplification process. Simplifications are schemes that by reducing, channelling and structuring the possibilities, create predictable relations in task environments that are yet unpredictable. Take the example of 'platforms'. A platform such as Apple's digital hub creates a predictable and replicable structure to combine objects (like the camcorder and the PC), when the possible combinations are plentiful and unknown.

As we argued above, simplifications organize choice interactions *and* strategic interactions simultaneously. Organizing interactions here means to provide a structure in which choices and market participants can interact in a given set of rules. The rules are not arbitrary, but a means to cut through the complexity of interactions across decisions, time and agents. For instance, the digital hub concept (Jobs, 2011) specified how software, hardware, and lifestyle interact. Indirectly, it also specified who has control over which parts of the associated value chains. To create concepts such as the digital hub requires metaheuristics (Ehrig, Jost, Warglien, 2019), and of course, not every concept is successful. As (Ehrig, Jost, Warglien, 2019) argue, success again can be a result of handling complexity, in this case in a vastly big *mental* search space, and for this, abstract mathematical tools may help.

Obviously, there is the danger of oversimplification. For instance, (Miller, 1993) gives examples of firms that simplified too much and were then stuck with a wrong simplification that eventually caused failure. A more concrete example in the strategy field of a common, but often harmful simplification is the use of linear regression models in contexts with higher order interactions and various feedback loops (Bettis and Blettner, 2019). There is, however, also the opposite danger of overcomplexification. It is therefore important (not only) for a strategist to find the right balance. Of course, we do not ignore that simplification can be harmful. Rather, we contribute by identifying theoretical principles that inform us which simplifications are useful and why, and which simplifications are harmful.

While the insight that simplification is at the heart of many successful strategies is not new, we take a further step. We do not see a strategy as the fixed outcome of a singular simplification, but as the ongoing process of simplification in environments that continuously or perhaps also discontinuously change. Miller's point was that firms that fixed a simple strategy failed when the environment changed. Our point is that environments change and this change implies ever new, growing complexity. Taming this complexity is an ongoing process. For that reason, processes of simplification are related to dynamic capabilities, which we will discuss next.

Simplification as a Dynamic Capability Much of strategy research is concerned with the change of environments (Eisenhardt and Martin, 2000), and capabilities to cope with such change in a superior way. But what is change, from a complexity science point of view? We discussed in detail that in ill-structured environments, by definition, change is more than a change of interactions of detailed choices in a given landscape.

In ill-structured environments, the first task of a strategist should be to infer what has changed. In particular, not just the means of production, but also the logic of value for the customer and the logic of competition can change. As our discussion of machine learning suggested, important dynamic capabilities of strategists are metaheuristics, to detect the type of change that has occurred and then apply a heuristic that is appropriate for that type of change (Ehrig and Schmidt, 2019). In particular, an important aspect of such capability is to discern whether massive new complexity needs to be tamed on the side of novel products, or on the side of a new logic of competition.

If complexity needs to be reduced to find novel products, strategists may shape the new environment by establishing proofs of concepts (like the Tesla S model). Proofs of concepts create certainty that a new product based on a new value proposition is actually feasible.

Table 3: **Principles from machine learning and their explanations of dynamic capabilities**

Principles in Complexity Science	**Principles of Dynamic Capabilities**
Metaheuristic to identify type of structure	Infer the new logic of competition
Separate data from noise by penalizing complexity	Simple Rules for Strategy Sweet-spot of right amount of structure
Adapt complexity of model to number and quality of observations	Infer environmental characteristics to determine how heuristics should be learned
Information bottleneck	Resource scarcity can be helpful for adaptation

If the logic of competition changes, strategists need to anticipate new bottlenecks, and also, anticipate which existing bottlenecks in markets no longer exist. Take again the example of Kodak . Kodak was very good at building dynamic capabilities to change from analogue to the digital production of images. However, Kodak failed to anticipate that one cannot make money with digital images but that they became rather a by-product of new types of products such as smartphones. Thus, image delivery was no longer a key source of profits.

In sum, our transfer of results from the sciences of complexity suggests that we need more research on how strategists make qualitative inferences about the structure of novel task environments. This question has normative and descriptive dimensions. Like (Leiblein et al., 2018) we are argue that psychological research in the strategic management domain needs to be extended to understand cognitive capabilities for dealing with interdependencies among choices, different time scales of decisions, and agents. The argument in this paper clearly indicates that we need descriptive psychological studies to understand how people simplify.

Psychological research on heuristics so far almost exclusively explains the effectiveness of heuristics by the bias-variance trade-off (eg. (Gigerenzer and Brighton, 2009)). We offer additional theoretical principles that may explain the success or failure of metaheuristics, and, relatedly, dynamic capabilities. Consider table (3). In the first row, we draw the analogy from metaheuristics in machine learning to inferring the new logic of competition in a transforming business environment. We can learn that to infer structure, the decision-maker needs to rely on glimpses of feedback, as discussed above ("weak cues"), and moreover, be aware how much has changed (discussed above under "correlation length"). Moreover, the principle of separating data from noise, mentioned in the second row, offers a general explanation of why there is a sweet spot of structure of a model to learn about an environment. Above, we formally derived this result. The result is general and extends beyond the specific setting studied in (Davis, Eisenhardt, Bingham, 2009). Relatedly, environmental characteristics need to be inferred, also to decide how quickly one should generalize and learn simple rules (third row; (Ehrig and Schmidt, 2019)). Finally, sometimes it is better to have fewer resources, as the information bottleneck principle (last row) suggests. By analogy, (Schilling, 2019) argued that Mariott, Hilton etc would have never come up with the idea of the ice hotel, a highly profitable opportunity, but it was found due to a lack of resources and failure (the original business ideas was an ice sculpture exhibition; it rained; ice sculptures were melting, but guests liked the igloos in which the artists were living).

To conclude, this article presented the argument that simplification is a key capability to cope with uncertainty, complexity and change. At its core, this capability rests on a successful translation of ill-structured into well-structured problems, and this translation organizes both choice interactions and strategic interactions. It is a dynamic capability, as this re-organization enables firms to redeploy their resources.

Of course, we could only sketch the implications of the newer results from the sciences of complexity for strategy science. To do this in detail would extend far beyond a single paper. However, our arguments suggest concrete steps forward: For instance, while we understand how the bias-variance trade-off explains concrete heuristics used by firms (DeMiguel, Garlappi, & Uppal, 2007,Ehrig and Schmidt, 2019), it is an important open question if the other principles we discussed (like compressed sensing) can explain documented heuristics used by firms. Moreover, can the metaheuristics that we discussed be developed into universal management tools that can be taught, for instance, in consulting? Or must metaheuristics be tailored to the specific details of firms, that is, are most implemented metaheuristics idiosyncratic? While we identified possible generic theoretical mechanisms underlying metaheuristics and associated dynamic capabilities, it is an open question whether the details of the implementation of metaheuristics may be specific to firms. Thus, some concretely implemented metaheuristics may provide specific firms with adaptation advantages that cannot easily be copied. Whether this is the case is an open question left for future work.

References

[Argote, 1999] Argote L. (1999) Organizational Learning: Creating, Retaining, and Transferring Knowledge. Kluwer Academic: Boston, MA.

[Ay et al. 2011] Ay N, Olbrich E, Bertschinger N, Jost J (2011) A geometric approach to complexity. *Chaos* 21: 037103.

[Ay et al.(to appear)] N.Ay, M.Christensen, J.Jost, T.Knudsen, Higher-order interactions: Hidden driver of complexity in organizations, submitted to this special issue

[Baer, Dirks, and Nickerson, 2013] Baer, M., Dirks, K.T. and Nickerson, J.A., 2013. Microfoundations of strategic problem formulation. Strategic Management Journal, 34(2), pp.197-214.

[Barney, 1986] Barney, J. 1986. Strategic factor markets: Expectations, luck, and business strategy. Management science, 32(10), pp.1231-1241.

[Belkin and Niyogi, 2003] M.Belkin, P.Niyogi, Laplacian Eigenmaps for Dimensionality Reduction and Data Representation, Neural Comp. 15, 1373-1396, 2003

[Bettis, 2017] R.Bettis, Organizationally Intractable Decision Problems and the Intellectual Virtues of Heuristics, Journal of Management Vol. 43 No. 8, November 2017 26202637 DOI: 10.1177/0149206316679253

[Bettis and Prahalad, 1995] R. A. Bettis and C.K.Prahalad, Strategic Management Journal, Vol. 16, No. 1 (Jan., 1995), pp. 5-14

[Bettis and Blettner, 2019] Bettis R, Blettner D (2019) Strategic reality. Strategic Management Rev.

[Brandenburger and Stuart, 1996] Brandenburger AM, Stuart HW Jr (1996) Value-based business strategy. J. Econom. Management Strategy 5(1):5–24.

[Breidbach and Jost, 2006] O. Breidbach, J. Jost, On the gestalt concept, Theory Bioscienc.125, 2006, 19–36

[Candès et al., 2006] E.Candès, J.Romberg, T.Tao, Stable signal recovery from incomplete and inaccurate measurements, Comm.Pure Appl.Math.59, 1207-1223, 2006

[Christensen et al., 1987] Christensen, C. R., K. R. Andrews, J. L. Bower, R. G. Hamermesh and M. E. Porter. Business Policy: Text and Cases, Irwin, Homewood, IL, sixth edition, 1987

[Csaszar and Levinthal, 2016] F.Csaszar and D.Levinthal, Mental representations and the discovery of new strategies, Strat. Mgmt. J., 37: 2031–2049 (2016)

[Cyert and March, 1963] Cyert, R., and J. G. March (1963), A Behavioral Theory of the Firm. Englewood Cliffs, NJ: Prentice-Hall.

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

[DeMiguel, Garlappi, & Uppal, 2007] DeMiguel, V., Garlappi, L., & Uppal, R. (2007). Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy?. The review of Financial studies, 22(5), 1915-1953.

[Donoho, 2006] D.Donoho, Compressed sensing, IEEE Trans.Inform.Theory 52, 1289-1306, 2006

[Ehrig, Jost, Katsikopoulos and Gigerenzer, 2019] T.Ehrig, J.Jost, K.Katsikopoulos, G.Gigerenzer, 2019 Heuristics for Coping with Strategic Uncertainty in Banking. Working Paper. Max Planck Institute for Mathematics in the Sciences.

[Ehrig and Schmidt, 2019] Ehrig, T. Schmidt, J. 2019. Making biased but better predictions: The trade-offs strategists face when they learn and use heuristics, Strategic Organization 1–22

[Ehrig, Jost, Warglien, 2019] Ehrig, T. Jost, J., Warglien, M. 2019. Changing Strategic Representations – A Formal Language. Working Paper. Max Planck Institute for Mathematics in the Sciences.

[Ehrig, Manjali, Singh, Sunder, 2019] Ehrig, T., Manjali, J., Singh, A., Sunder, S. 2019. Towards a behaviourally plausible theory of expectations about others? behavior: Do emotions regulate strategists? depth of reasoning? Yale Working Paper Series.

[Eisenhardt and Martin, 2000] K.Eisenhardt and J.Martin, Dynamic capabilities: What are they? Strat. Mgmt. J., 21: 1105–1121 (2000)

[Farjoun, 2008] M.Farjoun, Strategy making, novelty and analogical reasoning – Commentary on Gavetti, Levinthal, and Rivkin (2005), Strat. Mgmt. J., 29: 1001–1016 (2008)

[Felin et al., 2013] T.Felin, S.Kauffman, R.Koppl, G.Longo, Economic Opportunity Beyond Bounded Rationality and Phase Space

[Fontana et al., 1993] Fontana W, Stadler PF, Bornberg-Bauer EG, Griesmacher T, Hofacker IL, Tacker M, Tarazona P, Weinberger ED, Schuster P (1993) RNA folding and combinatory landscapes. Physical Rev. E 47:2088–2099.

[Foucard and Rauhut, 2013] S.Foucard, H.Rauhut, A mathematical introduction to compressive sensing, Birkhäuser, 2013

[Fudenberg and Levine, 1998] D.Fudenberg and D.Levine, The theory of learning in games, MIT Press, 1998

[Gans & Ryall, 2017] Gans, J., & Ryall, M. D. (2017). Value capture theory: A strategic management review. Strategic Management Journal, 38(1), 17-41.

[Garud, Kumaraswamy, & Karne, 2010] Garud, R., Kumaraswamy, A., & Karne, P. (2010). Path dependence or path creation?. Journal of Management Studies, 47(4), 760-774.

[Gavetti, 2012] Gavetti G. 2012. Toward a Behavioral Theory of Strategy. Organization Science 23, 267–285

[Gavetti, Helfat and Marengo, 2017] Gavetti, G., Helfat, C. E., & Marengo, L. 2017. Searching, shaping, and the quest for superior performance. Strategy Science, 2(3), 194-209.

[Gavetti and Levinthal, 2000] Gavetti G, Levinthal DA. 2000. Looking forward and looking backward: cognitive and experiential search. Administrative Science Quarterly 45(1): 113–137.

[Gavrilets(2004)] S.Gavrilets, Fitness landscapes and the origin of species, Princeton Univ.Press, 2004

[Ghemawat & Levinthal, 2008] Ghemawat, P., & Levinthal, D. (2008). Choice interactions and business strategy. Management Science, 54(9), 1638-1651.

[Gigerenzer and Brighton, 2009] Gigerenzer G, Brighton H. 2009. Homo Heuristicus: why biased minds make better inferences. Topics in Cognitive Science 1: 107–143.

[Gigerenzer and Gaissmaier, 2011] Gigerenzer G, Gaissmaier W. 2011. Heuristic decision making. Annual Review of Psychology 62: 451–482.

[Gigerenzer and Selten, 2001] Gigerenzer G, Selten R (eds). 2001. Bounded Rationality: The Adaptive Toolbox. MIT Press: Cambridge, MA.

[Granovetter, 1985] Granovetter, M., 1985. Economic action and social structure: The problem of embeddedness. American journal of sociology, 91(3), pp.481-510.

[Grant, 2016] Grant, R. 201. Contemporary Strategy Analysis. Wiley

[Hackbusch, 2012] W. Hackbusch, Tensor spaces and numerical tensor calculus. - Heidelberg : Springer, 2012. - xxiv, 500 p. (Springer series in computational mathematics ; 42) ISBN 978-3-642-28026-9.

[Hafenbrädl et al., 2016] Hafenbrädl S, Waeger D, Marewski J, Gigerenzer G. 2016. Applied decision making with fast-and-frugal heuristics. Journal of Applied Research in Memory and Cognition 5(2): 215–231

[Haleblian and Finkelstein, 1999] Haleblian J, Finkelstein S. 1999. The influence of organizational acquisition experience on acquisition performance: a behavioral learning perspective. Administrative Science Quarterly 44(1): 2956.

[Hong, Scheinkman & Xiong, 2006] Hong, H., Scheinkman, J., & Xiong, W. 2006. Asset float and speculative bubbles. The journal of finance, 61(3), 1073-1117.

[Jain and Kogut, 2014] A.Jain and B.Kogut (2014): Memory and Organizational Evolvability in a Neutral Landscape. Organization Science 25(2), 479–493

[Jin et al., 2015] Y. Jin, J. Jost, G.F. Wang, A new nonlocal variational setting for image processing, Inverse Problems and Imaging, 9(2), 415–430, 2015

[Jobs, 2011] Jobs, S. 2011. The Digital Hub. Presentation. https://www.youtube.com/watch?v=lmvmtmqqbeI

[Jost, 2004] J. Jost, External and internal complexity of complex adaptive systems, Theory Biosci.123, 69–88, 2004

[Jost, 2017] J.Jost, Object Oriented Models vs. Data Analysis Is This the Right Alternative?, in: J.Lenhard, M.Carrier (eds.), Mathematics as a tool, Boston Studies in the Philosophy and History of Science 327, pp. 253– 286, Springer, 2017

[Jost, 2019a] J.Jost, Leibniz und die moderne Naturwissenschaft, Monograph, Series *Wissenschaft und Philosophie, Science and Philosophy, Sciences et Philosophie*, Springer, in press, 2019

[Jost, 2019b] J.Jost, Biologie und Mathematik, Springer, in press, 2019

[Jost, 2019c] J.Jost, Mathematical Principles of Topological and Geometric Data Analysis, Lecture Notes, 2019

[Jost, 2019d] J.Jost, Biology, Geometry and Information, submitted, 2019

[Kauffman(1989)] Kauffman, S. 1989, Adaptation on rugged fitness landscapes. In D. Stein (ed.), Lectures in the Sciences of Complexity: 517–618. Reading, MA: Addison-Wesley.

[Knight, 1921] F.Knight, Risk, uncertainty, and profit, 1921

[Knudsen and Srikanth, 2014] T.Knudsen and K.Srikanth (2014), Coordinated Exploration: Organizing Joint Search by Multiple Specialists to Overcome Mutual Confusion and Joint Myopia, Administrative Science Quarterly 59: 409–441

[Knudsen, Levinthal, Winter, 2014] Knudsen, T., Levinthal, D.A. and Winter, S.G., 2014. Hidden but in plain sight: The role of scale adjustment in industry dynamics. Strategic Management Journal, 35(11), pp.1569-1584.

[Laland et al., 2000] K.Laland, J. Odlin-Smee, M.Feldman, Niche construction, biological evolution, and cultural change. Behavorial and Brain Sciences (2000) 23, 131–175

[Laubichler and Renn, 2015] M.D. Laubichler and J. Renn, Extended evolution: A conceptual framework for integrating regulatory networks and niche construction. Journal of Experimental Zoology Part B: Molecular and Developmental Evolution 324(7):565–577, 2015

[Leiblein et al., 2018] M. Leiblein, J. Reuer, T. Zenger (2018) What Makes a Decision Strategic?. Strategy Science 3(4):558–573.

[Lenox et al., 2006] Lenox, M.J., Rockart, S.F. and Lewin, A.Y., 2006. Interdependency, competition, and the distribution of firm and industry profits. Management Science, 52(5), pp.757-772.

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

[Levinthal and March, 1993] Levinthal, D. A., J. G. March. 1993. The myopia of learning. Strategic Management J. 14(S2) 95112.

[March, 1991] March, J. G. 1991. Exploration and exploitation in organizational learning. Organ. Sci. 2(1) 7187.

[March and Simon, 1958/93] J.March and H.Simon, Organizations, Wiley, 1958; 2nd ed., Blackwell, 1993

[Martignoni, Menon & Siggelkow, 2016] Martignoni, D., Menon, A., & Siggelkow, N. (2016). Consequences of misspecified mental models: Contrasting effects and the role of cognitive fit. Strategic Management Journal, 37(13), 2545-2568.

[Menon, 2018] A.Menon, Bringing cognition into strategic interactions: Strategic mental models and open questions, Strat. Mgmt. J. 39:168–192, 2018

[Miller, 1993] Miller, D. 1993. The architecture of simplicity. Academy of Management review, 18(1), 116-138.

[Mintzberg, 1978] Mintzberg, H., 1978. Patterns in strategy formation. Management science, 24(9), pp.934-948.

[Mintzberg, 1990] Mintzberg, H. (1990). The design school: reconsidering the basic premises of strategic management. Strategic management journal, 11(3), 171-195.

[Newell et al., 1958/1963] A.Newell, J.C.Shaw, H.A. Simon, Chess-playing programs and the problem of complexity, IBM J.Res.Develop. 2, 320–355, 1958; reprinted (with corrections) in: Computers and Thought, E.Feigenbaum and J.Feldman (eds.), New York, McGraw-Hill, 109–133, 1963

[Newell and Simon, 1972] A.Newell, H.A. Simon, Human problem solving, Englewood Cliffs, NJ, Prentice Hall, 1972

[Newell and Simon, 1976] A.Newell, H.A. Simon. Computer science as empirical inquiry: Symbols and search. Communications of the ACM, 19(3): 113–126, 1976.

[Oerlemans and Meeus, 2002] L. Oerlemans and M. Meeus, 2002. Spatial embeddedness and firm performance: an empirical exploration of the effects of proximity on innovative and economic performance, ERSA conference papers ersa02p054, European Regional Science Association.

[Pavliotis and Stuart, 2008] G.Pavliotis, A. Stuart, Multiscale methods. Averaging and homogenization, Springer, 2008

[Pearl, 1984] J. Pearl, Heuristics: Intelligent search strategies for computer problem solving. Reading, MA: Addison-Wesley, 1984.

[Pearl, 1988/1997] J. Pearl, Probabilistic reasoning in intelligent systems: Networks of plausible inference. Morgan Kaufmann Publishers, San Francisco, 1988; 4th printing, 1997

[Porter, 1980] M.Porter 1980. Competitive Strategy: Techniques for Analyzing Industries and Competitors. Free Press: New York

[Porter & Siggelkow, 2008] Porter, M., & Siggelkow, N. (2008). Contextuality within activity systems and sustainability of competitive advantage. Academy of Management Perspectives, 22(2), 34-56.

[Prahalad and Bettis, 1986] Prahalad, C. K. and R. A. Bettis (1986). 'The dominant logic: A new linkage between diversity and performance', Strategic Management Journal, 7(6), pp. 485–501.

[Renner, 2004] Renner, T. 2004. Kinder, der Tod ist gar nicht so schlimm! Berlin: Campus.

[Rieke et al.(1997)] F.Rieke, D.Warland, R.de Ruyter van Steveninck, W.Bialek, Spikes. Exploring the neural code, MIT Press, 1997

[Rivkin & Siggelkow, 2007] Rivkin, J. W., & Siggelkow, N. (2007). Patterned interactions in complex systems: Implications for exploration. Management science, 53(7), 1068-1085.

[Ryall, 2009] Ryall, M. D. (2009). Causal ambiguity, complexity, and capability-based advantage. Management Science, 55(3), 389-403.

[Schoener, 2009] T.Schoener, Ecological niche, in: S.Levin et al. (eds.), The Princetion Guide to Ecology, Princeton Univ.Press, pp.3ff., 2009

[Shubik, 1997] Shubik, M. 1997. Game theory, complexity and simplicity part 1: A tutorial. Complexity, 3(2), 39-46.

[Schilling, 2019] Schilling, M., Exploring the Microfoundations of Forward-Looking Strategy, Symposion at the Academy of Management Conference, Boston: 2019.

[Simon, 1947/97] H.Simon, Administrative behavior, New York, 1947; 4th ed., 1997

[Simon, 1955] H.Simon, A behavioral model of rational choice. Quarterly Journal of Economics, 69: 99–118, 1955

[Simon, 1957] H.Simon, Models of Man. John Wiley, 1957

[Simon, 1969/96] H.Simon, The sciences of the artificial, MIT Press, 1969, 3rd ed., 1996

[Simon, 1973] Simon, H. A. 1973. The structure of ill structured problems. Artificial intelligence, 4(3-4), 181-201.

[Stadler and Stephens, 2003)] P. Stadler and C. Stephens. Landscapes and effective fitness. Comments on Theoretical Biology, 8:389–431, 2003.

[Teece, 2007] . Explicating dynamic capabilities: the nature and microfoundations of (sustainable) enterprise performance. Strategic management journal, 28(13), 1319-1350.

[Teece et al., 1997] D.Teece, G. Pisano, A.Shuen A. (1997). Dynamic capabilities and strategic management. Strategic Management Journal 18(7): 509–533.

[Tishby et al., 1999] N. Tishby, F. Pereira, W. Bialek (1999). The Information Bottleneck Method (PDF). The 37th annual Allerton Conference on Communication, Control, and Computing. pp. 368–377

[Tomasello, 2003] M.Tomasello, Constructing a language, Harvard Univ.Press, 2003

[Tripsas and Gavetti, 2000] M.Tripsas, G. Gavetti. Capabilities, Cognition, and Inertia: Evidence from Digital Imaging. Strategic Management Journal, Vol. 21, No. 10/11, Special Issue: The Evolution of Firm Capabilities (Oct. - Nov., 2000), pp. 1147–1161

[Vapnik, 2013] Vapnik, Vladimir. The nature of statistical learning theory. Springer, 2013.

[Winter, 2003] S.Winter, Understanding dynamic capabilities, Strat. Mgmt. J., 24: 991–995 (2003)

[Wright, 1931] Wright, S., Evolution in Mendelian populations. Genetics, 16: 97–159, 1931

[Wright, 1932] Wright, S., The role of mutation, inbreeding, cross-breeding and selection in evolution. Proceedings Xl International Congress of Genetics, 1: 356–366, 1932

[Wulff, Mergenthaler-Canseco, Hertwig, 2018] Wulff, D.U., Mergenthaler-Canseco, M. and Hertwig, R., 2018. A meta-analytic review of two modes of learning and the description-experience gap. Psychological bulletin, 144(2), p.140.