



Modelling Bounded Rationality in Organizations: Progress and Prospects

Journal:	<i>Academy of Management Annals</i>
Manuscript ID:	ANNALS-2013-0071.R4
Document Type:	Article
Keywords:	DECISION MAKING, LEARNING, RESEARCH METHODS, ORGANIZATION

SCHOLARONE™
Manuscripts

Modelling bounded rationality in organizations: Progress and Prospects

Phanish Puranam, Nils Stieglitz, Magda Osman, Madan M. Pillutla

Abstract:

Much of the formal modelling work in the organizational sciences relies on Herbert Simon's conception of bounded rationality, and it stakes a claim to drawing on behaviorally plausible assumptions about human behavior and action in organizations. The objectives of our review are three-fold. First, we summarize the formal literature by "model families" - classes of models sharing the same analytical structure- to highlight sharply the behavioral assumptions being made. Specifically, we discuss model families involving a) adaptation through search and learning by individual agents, b) mutual adjustment of interacting agents, and c) information aggregation in organizational decision-making. Second, we examine to what the extent these models of bounded rationality in organizations are in fact consistent with the behavioral evidence in psychology and other related fields. Finally, we discuss opportunities for further research that strengthens the links between formal modelling in organizations research, and its behavioral foundations. In particular we highlight the promise of experimental methods that translate organizational models to multiple-subject experiments in the behavioral laboratory.

Note:

An early draft of this article was prepared as a background note for a workshop on "The psychological foundations of organizational models" held at London Business School, May 11-12, 2012. We also thank Felipe Csaszar, Jerker Denrell, Dan Levinthal, and Thorbjørn Knudsen for helpful comments on earlier drafts of the paper. All remaining errors are ours.

1. Introduction

The work of Herbert Simon (1916-2001), his colleagues and others inspired by his vision of constructing a science of organization has always had a strong component of formal (mathematical and computational) modelling (e.g. Simon, 1955; Cyert and March, 1963). Formal models allow for a precise statement of assumptions in a manner that makes them transparent, and allows for an examination of the possibly surprising ways in which a system characterized by these assumptions behaves over time. Verbal theorizing does not offer either this precision or insight into dynamics, unless practiced by an extraordinarily gifted theorist (Lave and March, 1975; Davis et al., 2007; Harrison et al. 2007; Adner et al., 2009).

The tradition of modelling behavior in and of organizations inspired by Simon and his colleagues stands in contrast to other forms of modelling in the social sciences primarily because of the centrality of the notions of bounded rationality.

Simon's (1955) notion of bounded rationality assumed that 1) the decision making process begins with the search for alternatives 2) the decision maker has egregiously incomplete and inaccurate knowledge about the consequences of actions, and 3) she chooses actions that are expected to be satisficing (attain targets while satisfying constraints). Later work by Cyert and March (1963) added an important elaboration - problemistic search- which indicates that the search process described above is itself triggered by a failure to meet aspiration levels (see also recent overviews by Gavetti et al (2007; 2012) of the tradition of research inspired by Simon). We will refer to this form of bounded rationality, to paraphrase Simon, as a form of "adaptive rationality" as it emphasizes the process of searching for better alternatives by reacting to feedback (e.g. Simon, 1996).¹

¹ Simon's conception of global rationality was as follows (Simon, 1997:25) :“This is what I call full or global rationality: people are making their decisions to maximize utility in a world which they either understand exactly or in terms of a known probability distribution (i.e. they are maximizing subjective expected utility).” Since bounded rationality is defined by its departure from global rationality, it follows that there are many possible ways in which rationality can be bounded. There is, of course, a significant body of modelling work with some form of bounded rationality in economics; see for instance Rubinstein (1998) or Spiegler (2011), as well as an extensive empirical tradition (e.g.

1
2
3
4
5 Simon's intention was to theorize about and to model organizations in a manner that took
6
7 account of the nature of the human agents who constituted them. Specifically, if human
8
9 agents are best characterized as being bounded, rather than globally rational when confronted
10
11 with typical problems in organizational settings, then we should model organizations with
12
13 these 'correct' behavioral assumptions since organizations are aggregations of these human
14
15 agents. It is not the purpose of this essay to review the evidence for the proposition that
16
17 individuals are boundedly rather than unlimitedly rational, which we take as granted.² Instead,
18
19 our focus is on investigating whether models of organizations that assume boundedly rational
20
21 agents make psychologically 'correct' assumptions about humans-in other words to assess
22
23 their behavioral plausibility.
24
25
26

27 Since Cyert and March (1963), about a 100 papers invoking models of bounded, adaptively
28
29 rational behavior in and by organizations have been published in the leading outlets for
30
31 research about organization and management theory. A common feature of these papers is
32
33 their explicit declaration of building on behaviorally plausible models of human action- i.e. on
34
35 realistic assumptions about human behavior in organizations. A few have achieved canonical
36
37 status such as the "garbage can" model of Cohen, March and Olsen (1972) and March's
38
39 model of exploration and exploitation in organizational learning (1991). These models and
40
41 their insights have percolated widely beyond the community of modellers in the organization
42
43 sciences to an extent where it is no longer necessary to make them the focus of review articles
44
45
46
47

48
49 Kahneman and Tversky, 1979; Camerer, 2003). In this literature, bounded rationality is conceptualized
50
51 variously in terms of the existence of communication and search costs, malleable and unstable
52
53 preferences, or outright processing errors (e.g. Kahneman and Tversky, 1979; Gilboa and Schmeidler,
54
55 2001), while the modelling typically retains the notion of optimization subject to these features.

56
57 ² The statement by Schoemaker (1982) in his review of the evidence on expected utility (EU)
58
59 maximization remains a good summary: "*For well-structured repetitive tasks, with important stakes,
60
and well-trained decision makers, EU maximization may well describe the actual decision process.*"
In other words unbounded rational choice may be a useful theory when there is a small or well-known
space of choices, (expected) utilities associated with the choices have been discovered and are now

1
2
3 such as this one (there is an excellent one by Gavetti, Greve, Levinthal, & Ocasio (2012) in
4 this very journal for those interested in the research inspired by these classic models from the
5 Carnegie tradition). Rather, our objective in this essay is to do three things not undertaken so
6 far:
7
8
9

10
11
12
13 **First**, as noted, we wish to examine the growing body of modelling work based on adaptive
14 rationality in order to examine to what extent these models of organizations are in fact
15 consistent in their assumptions about behavior with the available scientific evidence. If this
16 consistency is a central claim for the distinctiveness of modelling by organization and
17 management theorists, then it would be good for the field to explicitly evaluate the strength of
18 these claims. Drawing on assumptions about human behavior that are better grounded in
19 psychology than traditional rational choice models is an appealing feature of these models for
20 management and organizations researchers that sometimes remains obscured behind the
21 technical details of the models. By exposing this, we aim to increase the appreciation for
22 formal models among organizations researchers, which as Adner et al. (2009) point out is still
23 fairly limited.
24
25
26
27
28
29
30
31
32
33
34
35
36

37 **Second**, we will systematically summarize this body of work by “model families” - classes of
38 models that share an underlying analytical structure, with a distinctive set of assumptions and
39 mechanisms, and resulting insights. Specifically, we discuss model families which examine a)
40 individual agent adaptation through search and learning, b) mutual adjustment of an
41 organization’s constituent agents and c) information aggregation. Such an exercise we believe
42 has enormous instructional value for modellers as well as organizations researchers in general,
43 by making transparent underlying assumptions, as well as the kinds of problems to which the
44 models may bring insight, and finally by suggesting avenues for empirical research.
45
46
47
48
49
50
51
52
53
54

55
56 well known, and there is sufficient difference between the outcomes of choices to make maximization
57 easy. Such situations, Simon argued, are rare within organizations.
58
59
60

1
2
3
4
5 **Third**, we discuss opportunities for further research that strengthen the links between formal
6 modelling in organizations research, and its behavioral foundations. Surprisingly, direct tests
7 of the empirical predictions of organizational models of adaptive rationality have been rare,
8 with these models more often being treated as metaphors that shed light on how previously
9 known mechanisms might interact in surprising ways. While this is valuable, what makes
10 models most useful is their ability to predict. Beyond the usual empirical approaches
11 management researchers favour (e.g. large sample archival data analysis, surveys, case
12 studies), in our view, there is also an opportunity to use experimental methods to test model
13 predictions. These experiments could translate agent-based models to multiple-subject
14 experiments in the behavioral lab as a methodology to test the predictions. Pragmatically,
15 since both modeling and experimental work deal with interaction structures among a small
16 number of agents, the possibilities for theory building, calibration and theory testing are
17 numerous at their interface. Empiricists may benefit from sharper statements of hypotheses
18 that model based predictions can generate, while modelers may benefit from the data that
19 empiricists gather that can help falsify current models and force them to improve in their
20 veridicality and usefulness.
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38

39 Thus, at a broader level, our goal in this article is to bridge the chasm between empirical
40 researchers (including experimentalists) and theoretical researchers in order to help stimulate
41 the production of scientific (i.e. behaviorally grounded) models in organization and
42 management theory. Our efforts are guided by the words of the Nobel Laureate Vernon Smith
43 (1982: 924) who argued that ‘theory should be ever more demanding of our empirical
44 resources. Simultaneously, data should be ever more demanding of the empirical relevance of
45 theory and of the theorist’s expertise in working imaginatively on problems of the world,
46 rather than on stylized problems of the imagination’. The modeling work on adaptive
47 rationality fulfills Vernon Smith’s criterion of theory that captures problems of the world as
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 its avowed purpose is to model the important problem of ‘organizing.’ It is therefore
4
5 deserving of empirical testing and we hope that our paper stimulates that.
6
7

8
9 The rest of this this article is organized as follows: in section 2, we describe the common
10
11 building blocks that are present in all models of bounded adaptive rationality. Section 3 is the
12
13 heart of the paper, and examines each of three important model families that appear often in
14
15 organizations research- models of individual adaptation, mutual adjustment, and information
16
17 aggregation. For each model family, we describe the model’s structure, discuss the extent to
18
19 which the assumptions of the model have sound foundations in the evidence about how
20
21 individuals and groups behave, and then state the key insights emerging from the model. We
22
23 highlight the core results that emerge for each model family rather than review the specific
24
25 results that are derived in each paper within this literature; this is because the results in
26
27 particular papers point to “moderators” that weaken or strengthen the effects underlying core
28
29 results. This aggregation into model families also implies that we are unable to do justice to
30
31 every influential model (e.g. Carley, 1992; Prietula, Carley, Glasser, 1998; Davis, Eisenhardt,
32
33 Bingham, 2009), as we pick those which have underlying platforms that are shared across
34
35 multiple papers and authors. Finally, section 4 concludes with a discussion of future research
36
37 opportunities.
38
39

40
41 To preview the fruits of this review, (which are discussed more fully in the last section), we a)
42
43 identify the four key behavioral assumptions that underlie the models we have reviewed,
44
45 namely exploratory choice, learning through reinforcement, search through local hill
46
47 climbing, and problemistic search, and verify whether these assumptions have an empirical
48
49 basis b) discuss testable implications from these models that have not so far been examined
50
51 by empirical researchers, and provide a useful mapping between model parameters and
52
53 empirically observable data in Table 1, and c) discuss the possibilities of using behavioral
54
55 laboratory experiments to test models of adaptive rationality.
56
57
58
59
60

2. Basic elements of a model of bounded adaptive rationality

Before moving to a discussion of different model families, we briefly describe the foundational elements of models of boundedly adaptive rational behavior.

1. *Task Environment*

This describes an objective reality which can be characterized with respect to an agent in terms of a set of possible *Actions* $\{A_1, A_2, \dots, A_m\}$ which map into *Performance outcomes* $\{\Pi_1, \Pi_2, \dots, \Pi_m\}$ for the agent. The task environment need not be stable; it could contain other agents, whose behavior may itself contribute to the instability of the task environment. Implicitly, given a task environment, the objective of the agent in an adaptive rationality model is to take actions that yield a level of performance deemed acceptable (often termed the aspiration level). Task environments thus capture a mapping between the possible courses of action available to an agent, and the likelihood of attaining the desired goal of the agent for each course of action. Task environments are tied to particular goals; a change in goals will imply a change in task environments.

2. A *Representation* of the Task Environment

An agent at any point in time may have beliefs (not necessarily knowledge) of the *Task Environment*, in terms of two things:

- i. a set of feasible actions $\{a_{1,t}, a_{2,t}, \dots, a_{m,t}\}$,
- ii. and the associated set of performance outcomes $\{\pi_{1,t}, \pi_{2,t}, \dots, \pi_{m,t}\}$.

3. A *Choice* process *given* the Representation

This is a procedure by which the agent chooses an action within his *Representation*.

1
2
3 These three elements of a model of boundedly rational behavior correspond to the three
4
5 hierarchically linked components that have been used to characterize behavior in
6
7 classical organization theory (e.g. Simon, 1946; Cyert and March, 1963): Goals
8
9 (operationalized through the task environment), Representations and Choice processes.
10
11 Interestingly, these elements are present in rational choice models as well. The difference
12
13 is that the task environment in a rational choice model is typically assumed to be known
14
15 perfectly to the agent so that representation and task environment are identical (or at least the
16
17 former is a known and unbiased probability distribution over the latter), and the choice
18
19 process is the maximization of subjective expected utility. To make the modelling of the
20
21 choice process mathematically tractable, it is also common in many such models to assume
22
23 certain mathematical properties of the task environment that make it easy to apply standard
24
25 techniques to find the best action.
26
27
28
29

30 Bounded adaptive rationality models are different. The first (and in our view both necessary
31
32 and sufficient) point of distinctiveness for such models is that the representation is incomplete
33
34 and/or incorrect in terms of the set of possible actions as well as their associated outcomes.
35
36 The agent 'has egregiously incomplete and inaccurate knowledge about the consequences of
37
38 actions', to paraphrase Simon. Thus at any point in time, the agents Representation of actions
39
40 $\{a_{1,t}, a_{2,t}, \dots, a_{k,t}\}$ may contain only a small subset of actions from $\{A_1, A_2, \dots, A_m\}$ (i.e they
41
42 have a "coarse" representation), and the beliefs about associated outcomes $\{\pi_{1,t}, \pi_{2,t}, \dots, \pi_{k,t}\}$
43
44 may be incomplete, inaccurate or both with respect to the actual values of outcomes in the
45
46 Task Environment $\{\Pi_1, \Pi_2, \dots, \Pi_m\}$.
47
48
49
50

51
52 A second point of departure (which, in our view is sufficient but not necessary to make it a
53
54 model of bounded rationality) is that the Choice process may not involve maximization even
55
56 *within* the representation. More specifically, alternatives to maximization include the notion
57
58
59
60

1
2
3 of “non-greedy” or exploratory choice (Sutton and Barto, 1998), whereby an agent may with
4 some probability select an action other than the one that is associated in the Representation
5 with the best outcome. A related concept is “satisficing” which involves making a choice
6 relative to an aspirational level (Cyert and March, 1963; Greve, 2003), rather than pursuing a
7 course of action aimed at the best possible outcome.
8
9

10
11
12
13
14
15 4. Critical to a model of bounded but adaptive rationality is a **Transformation** procedure that
16 modifies Representations and possibly Choice processes over time. The transformation
17 process may be a consequence of communication with other agents, a process of learning- a
18 change in beliefs as a consequence of experience- or indeed random fluctuations in beliefs.
19 The process of Transformation may itself be triggered by the consequences of the Choice
20 process, in the form of feedback from the Task Environment- the realized outcome. It may
21 affect not only Representations but also Choice processes by changing aspiration levels.
22 There is usually at least a tendency for the Transformation process in models of adaptive
23 rationality to move Representations towards alignment with Task Environments, but this can
24 take time, and is seldom either complete or comprehensive.³
25
26
27
28
29
30
31
32
33
34
35
36

37 Thus the typical structure of an adaptive rationality model involves some or all of the
38 following (possibly) recursive sequence:
39

- 40 1. Agent has (initial) **Representation** of the **Task Environment**.
- 41 2. Agent selects an action through a **Choice** process.
- 42 3. Feedback is generated from the consequences of the agents action in the **Task**
43 **Environment**.

44
45
46
47
48
49
50
51 ³ A process of **Selection (through competition)** sometimes accompanies organizational models of
52 adaptive rationality (e.g. Levinthal, 1997; Levinthal and Posen, 2007), and may be an alternative (or
53 addition) to the **Transformation** process in terms of generating change over time. Put simply, the
54 individual agents in the model may change their representations through Transformation processes, or
55 the distribution of representations in the population may change through Selecting out those agents
56 who have poor representations, or both. Since selection through competition does not feature in all
57 models of adaptive rationality, we only discuss it when relevant.
58
59
60

- 1
 - 2
 - 3
 - 4
 - 5
 - 6
 - 7
 - 8
 - 9
 - 10
 - 11
 - 12
 - 13
 - 14
 - 15
 - 16
 - 17
 - 18
 - 19
 - 20
 - 21
 - 22
 - 23
 - 24
 - 25
 - 26
 - 27
 - 28
 - 29
 - 30
 - 31
 - 32
 - 33
 - 34
 - 35
 - 36
 - 37
 - 38
 - 39
 - 40
 - 41
 - 42
 - 43
 - 44
 - 45
 - 46
 - 47
 - 48
 - 49
 - 50
 - 51
 - 52
 - 53
 - 54
 - 55
 - 56
 - 57
 - 58
 - 59
 - 60
4. A *Transformation* process of the *Representation* and possibly *Choice* process is triggered by the feedback.
 5. (Step 1 repeats).

As should be clear, our focus is on models that not only feature bounded rationality (i.e., in which agent's representations are imperfect) but are also adaptive (i.e. there is a transformation process that modified representations over time).⁴

Adaptively rational models of organizations vary in whether they model organizations as 'unitary' actors⁵ or as interacting agents whose sometimes conflicting preferences, beliefs and choice procedures cannot be captured conveniently in aggregate form as if they relate to a single agent. This suggests that we must consider two approaches to assessing behavioral plausibility. First, to what extent do unitary actor models of organizations embody assumptions about their adaptive rationality that are consistent with how groups and organizations as aggregations act? Specifically, there is much research in social psychology and organizational behavior about how and when groups and other collectives act (as an aggregate), and indeed act differently compared to individual actors in similar situations. We use that research as the basis to verify the validity of the behavioral assumptions of models of adaptive rationality that employ unitary actor assumptions. The second approach to assessing behavioral plausibility is to see how multi-agent models of organizations embody assumptions about the agents' adaptive rationality that are consistent with how individuals act.

⁴ Models with the first feature but not the second also exist in the organizational literature, and often invoke maximization within an imperfect representation (e.g. Siggelkow, 2002; Kretschmer and Puranam, 2008; Puranam and Vanneste, 2009; Gulati and Puranam, 2009). Since there is considerable variation in the nature of the Task Environments and Representations in these papers, we do not attempt to summarize these here.

⁵ The unitary actor assumption can be justified if there are rules of preference/belief aggregation, and selection into or out of an organization does not destabilize such processes; if an organization is effectively controlled by a powerful individual, or a governing coalition which has stable preferences, beliefs and choice procedures (see for instance Hug (1999) for a formal analysis of the conditions under which unitary actor assumptions may be more or less plausible).

3. Families of organizational models of adaptive rationality

We summarize the modelling literature on adaptive bounded rationality in organizations in terms of three main model families. The first, featuring models of individual agent adaptation, provides some or all of the building blocks for the second (mutual adaptation) and third (information aggregation). The literature is also most extensively developed around the first. Accordingly, the coverage of models of individual agent adaptation – particularly in terms of a discussion of their behavioral plausibility- is more extensive in this review than of the other two model families. It is our hope that the field will develop in a way that allows future reviews, such as this one, to give more coverage to the second and third families of models, which are very relevant to issues of aggregation in organizations.

For each of the models that we present, we begin with some information about the conceptual background of the model and, present its core elements (task environment, representation, choice, transformation of representation) to identify the behavioral assumptions made in the model. We then examine the behavioral plausibility of core assumptions and move on to discussing the key insights from the model.

A. Models of Individual Agent Adaptation

Since the discovery of superior (to status quo) alternatives through an adaptation process is the essence of adaptive rationality, we begin with a look at models of individual agent adaptation. We focus on models in which decision-making by an agent plays an explicit and central role in the model. Two important sub-classes of organizational models are considered within this family. The multi-armed bandit depicts learning and decision-making under uncertainty (and models organizational problems such as resource allocation, including but not restricted to making decisions about different investment opportunities, allocation of resources to different divisions of a company, etc.), while the NK model examines adaptation

1
2
3 in complex task environments that feature interdependencies between choices (capturing
4 organizational problems such as the formulation of new business models, adoption of
5 manufacturing or human resource policies with unknown interactions between them).
6
7
8
9

10 11 *a. Reinforcement learning in multi-armed bandit models* 12

13 This model takes its name from the slot machines in gambling establishments (Brand, Wood,
14 & Sakoda, 1956; Brand, Sakoda, & Woods, 1957). Metaphorically, the model describes a
15 gambler who faces a slot machine with multiple alternative arms, each with an unknown
16 payoff distribution. The gambler has a fixed budget and intends to make as much money as
17 possible over multiple trials. She learns about the payoff distributions associated with an arm
18 by selecting it and observing the realized payoff. Obviously, the gambler always wants to
19 select the alternative with the highest expected payoff, but the problem is that she does not
20 initially know the expected payoffs for each arm, and can only learn them by actually trying
21 each of the multiple arms. Thus the gambler's behavior is a process of repeated trials,
22 followed by learning from the outcome of each trial.
23
24
25
26
27
28
29
30
31
32
33

34
35
36 Lave and March (1975) and March (1996) introduced the bandit formulation to organizations
37 researchers; subsequent formal papers include Denrell and March (2001), Posen and
38 Levinthal (2012), Stieglitz, Knudsen, and Becker (2014), Puranam and Swamy (Forthcoming)
39 and Lee and Puranam (Forthcoming).
40
41
42
43
44

45
46 Modellers in the adaptive rationality tradition have made use of the multi-arm bandit's strong
47 structural similarity to resource allocation processes in organizations: the model describes
48 with reasonable fidelity a decision maker who aims to maximize returns from her investments,
49 and is confronted with multiple alternatives with unknown performance consequences. The
50 relative attractiveness of these alternatives can only be gauged by actually trying them (i.e.,
51 making some investments) and observing and interpreting the resulting feedback. The goal is
52
53
54
55
56
57
58
59
60

1
2
3 to discover the alternative with the best returns as quickly as possible. In an organizational
4 context, the uncertain investment alternatives could, for example, represent different
5 entrepreneurial opportunities (Minniti and Bygrave, 2001; Gruber et al., 2008), different
6 divisions in a multi-product firm (Noda and Bower, 1996), or different research projects
7 (Loch and Kavadias, 2002).
8
9
10
11
12

13
14
15 We describe in more detail below the four key elements of the multi-armed bandit model
16 before examining the behavioral plausibility of core assumptions and the key insights from
17 the model.
18
19
20
21

22 23 1. Task Environment 24

25 The task environment in a bandit model consists of a set of ‘m’ discrete alternatives, each
26 representing a possible course of action with distinct payoff consequences. Each alternative is
27 associated with a unique payoff distribution (whose properties are unknown to the agent).
28 This a) creates performance differences among alternatives and b) makes performance
29 feedback noisy and potentially misleading. False negatives may arise – a good alternative
30 looks bad because of an atypical unlucky draw from its payoff distribution, as can false
31 positives – a bad alternative may look good because of an atypical lucky draw from its payoff
32 distribution.⁶
33
34
35
36
37
38
39
40
41
42

43 To illustrate these ideas, consider a bandit with $m=2$, with alternatives A and B. A simple
44 everyday example is finding a good place to eat, with A and B representing two restaurants.
45
46
47

48 One alternative is in an objective sense superior for the decision-maker: alternative A
49
50

51
52 ⁶ The basic bandit setup usually assumes a stable task environment, i.e., the probability distribution of
53 payoffs for each arm is fixed. However, the model can be easily amended to also consider dynamic task
54 environments where the probability distributions fluctuate over time to depict changing markets or
55 technological conditions for the opportunities that each alternative represents (Posen and Levinthal,
56 2012, Stieglitz et al., 2014). Similarly, competence development and lock-in to particular alternatives
57 may be represented by making the probability distribution conditional on the number of times an
58 alternative has been chosen (Denrell and March, 2001).
59
60

1
2
3 dispenses a reward (of 1) with a 0.75 probability, while alternative B only has a success rate
4
5 of 0.5. In the restaurant example, a payoff of 1 would represent good food and a payoff of 0 -
6
7 bad food. A fully informed decision-maker would therefore always pick alternative A. The
8
9 bandit task is trivial as long as agents have full information about the alternatives. However
10
11 the purpose of the model is to describe the process that ensues when the knowledge about
12
13 alternatives – the agent’s representation (i.e., in the above example, they do not know the
14
15 probability of a reward associated with each alternative) – is far from complete initially.
16
17

18 19 2. Representation:

20
21 Commonly, in bandit models, the set of alternative actions in the agent’s representation is
22
23 complete and correct with respect to the Task Environment, but the associated performance
24
25 outcomes are not. This corresponds to situations where the space of alternatives is known, but
26
27 not their value. Thus the CEO of a multi-business firm may know the number of divisions in
28
29 the firm that need capital in the budgeting process, but not the true expected rates of return for
30
31 investing in them. Similarly, the number and location of restaurants may be known, but not
32
33 their quality. Instead, for each action in the agent’s representation, the agent may have at any
34
35 point in time (possibly egregiously inaccurate) beliefs (Yechiam and Busemeyer, 2005; Posen
36
37 and Levinthal, 2012). In the technical literature these are referred to as “action value
38
39 estimates” (Sutton and Barto, 1998). For example, the action value estimates for alternatives
40
41 A and B could be 0.5 and 0.66 respectively in the example above. Assumptions about these
42
43 prior beliefs represent the current state of knowledge of the agent, e.g., she could have heard
44
45 from friends that restaurant B is much better than A.
46
47

48 49 3. Choice process:

50
51 Given the agent’s representation, there are broadly two kinds of choice rules: maximizing and
52
53 probabilistic choice rules (Yechiam and Busemeyer, 2005). Maximizing choice rules assume
54
55 that agent always picks the alternative currently associated with the highest expected payoff
56
57 in the agent’s representation. This is also sometimes called greedy action selection (Sutton
58
59
60

1
2
3 and Barto, 1998). For example, an agent following a maximizing choice rule would select
4
5 alternative B based on her priors in our example, which is the incorrect choice. In the
6
7 literature, such a choice would be termed an exploitative one, as it picks the alternative
8
9 considered best within the current representation.

10
11
12 Probabilistic choice rules, in contrast, assume that choices are probabilistically determined in
13
14 proportion to the current relative strengths of the representations. A common probabilistic
15
16 choice rule in organizational models is the so called “softmax” action selection rule (Luce,
17
18 1959; Sutton and Barto, 1998; Posen and Levinthal, 2012). In this rule the probability of a
19
20 choice k at time t depends on its relative action value estimate V in the agent’s representation
21
22 and an exploration parameter τ .

$$23$$

$$24$$

$$25$$

$$26 \quad p(k)_t = \frac{e^{V_{k,t}/\tau}}{\sum_{b=1}^m e^{V_{b,t}/\tau}}$$

$$27$$

$$28$$

$$29$$

$$30$$

31 The parameter τ tunes the degree of exploration – the probability of choosing an alternative
32
33 currently seen as inferior within the representation – by discounting the current
34
35 representations. A higher τ discounts current representations more strongly and thereby
36
37 increases the probability of selecting a (currently believed to be) inferior action. For example,
38
39 given the (wrong) representations for A (0.5) and B (0.66), the probability of choosing B with
40
41 a very low degree of exploration $\tau = 0.1$ is 0.75. In contrast, with $\tau = 1$, the probability of
42
43 choosing B is just 0.52. Note that there is a clear benefit to exploration here, because the
44
45 agent currently holds a flawed representation, believing wrongly that alternative B is superior.
46
47
48

49 For a given task environment, Gittins and Jones (1974) formally proved the existence of an
50
51 optimal level of exploration in the bandit problem. However, as is generally acknowledged,
52
53 the information and computational power required to implement this optimal exploration level
54
55 makes this intractable except under very particular assumptions (Gittins, 1979; Arthur, 1991;
56
57 Posen and Levinthal, 2012). Organizations research inspired by these models has tended to
58
59
60

1
2
3 use either constant levels of exploration, comparing across levels (in order to study the effects
4 of different levels of exploration on system performance), or simple behaviorally plausible
5 rules for deciding the level of exploration dynamically (in order to focus on other parameters
6 of interest in the model). For instance, empirical work in the behavioral theory of the firm
7 tradition argues that organizations increase exploration as a response to underperformance
8 relative to an aspiration level (e.g. Greve, 2003; Hu et al., 2011). Applied to the softmax, this
9 would mean that parameter τ is contingent on performance relative to aspiration, with a
10 higher τ if the organization underperforms relative to its aspiration.
11
12
13
14
15
16
17
18
19
20

21 4. Transformation process:

22
23 Since most bandit models assume that the agent's representation is complete in terms of the
24 action space, but is incorrect ex ante in terms of the associated payoffs, the transformation
25 processes of interest in such models involve how the representation of the task environment
26 changes over time. Agents typically engage in reinforcement learning- whose variants are
27 known under the labels of trial-and-error learning, experiential learning, operant or
28 instrumental conditioning, "win-stay-lose-shift" rules in the relevant literatures in psychology,
29 computer science, organization theory and evolutionary biology (Thorndike, 1911; March,
30 1991; Domjan, 2010; Nowak and Sigmund, 1993; Sutton and Barto, 1998). All instantiate
31 Thorndike's Law of Effect: "*Responses that produce a satisfying effect in a particular*
32 *situation become more likely to occur again in that situation and responses that produce a*
33 *discomforting effect become less likely to occur again in that situation.*" Thorndike's law of
34 effect implies that favourable feedback tend to positively reinforce the belief about an
35 alternative and thereby make its choice more likely.
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50

51 The most basic way to model the reinforcement learning process is a simple averaging rule,
52 where the representation simply reflects the average of all payoff signals received in the past.
53 For example, assume that an agent visited restaurant A four times and received the following
54
55
56
57
58
59
60

1
2
3 payoffs [1,0,1,0]. Taking the average leads to a belief of 0.5. Restaurant B was selected six
4
5 times, with payoffs [1,1,1,0,0,1], so the belief in the agent's representation is 0.66.
6
7

8
9 A more general approach to agent's reinforcement learning is based on statistical learning
10
11 models (Bush and Mosteller, 1955; Denrell and March, 2001) that explicitly consider a
12
13 learning rate. For example, the exponential recency weighted average rule discounts rewards
14
15 that occurred long ago:
16

$$V_{m,t} = V_{m,t-1} + \phi[r_{m,t} - V_{m,t-1}]$$

17
18 Where $\phi \in [0,1]$ is called the step size parameter (Sutton and Barto, 1998) and can be
19
20 interpreted as a learning rate. The higher the parameter ϕ , the more recent rewards matter in
21
22 the representation, and the more rapid is the updating process. For instance, assume that the
23
24 current representation of B is 0.66 and the agent selects B and receives no payoff. With $\phi =$
25
26 0.9, the new representation is 0.066, a massive re-evaluation of the alternative. In contrast,
27
28 with $\phi = 0.1$ the new representation is 0.594, a much more modest change in representation.
29
30
31
32
33

34 The bandit model can also be extended by considering competition based selection in a
35
36 population of agents, in which the relative cumulative payoffs influence survival probability
37
38 (Denrell and March, 2001; Stieglitz et al., 2014). Some recent organizational models have
39
40 begun to look into a decision-making context in which actions are not followed by immediate
41
42 feedback (Denrell, Fang & Levinthal, 2004; Rahmandad, 2008; Fang and Levinthal, 2009).
43
44 Rather, feedback may only be available after a sequence of actions has been performed. The
45
46 delayed feedback conditions relate to the notion of credit assignment, i.e., the problem of
47
48 assigning payoffs to alternative actions (Fang, 2012). The modelling of learning in this
49
50 context is also based on reinforcement learning principles in which action value estimates are
51
52 assigned to intermediate actions.
53
54
55
56

57 ***Behavioral plausibility of bandit models***
58
59
60

1
2
3 There are two key behavioral assumptions about the agent in a bandit model: **reinforcement**
4
5 **learning** and **exploration in choice**.
6
7

8
9 There is an extensive body of experimental evidence documenting that reinforcement learning
10
11 is a good model for the behavior of individual adaptation in situations where the initial
12
13 understanding of a complex task environment is limited, (e.g. Erev and Roth, 1998; Camerer,
14
15 2003; Domjan, 2010). The fundamental validity of the Law of Effect in such situations is
16
17 beyond dispute (Thorndike, 1911; Domjan, 2010; Nowak and Sigmund, 1993). Indeed, there
18
19 is evidence suggesting that as the computational load of a task that an agent is performing
20
21 increases, they are more likely to rely on reinforcement learning rather than alternative
22
23 heuristic type strategies in choice tasks (Otto et al, 2010).
24
25

26
27 Experimental evidence also suggests that the softmax choice rule – which implies explorative
28
29 choice- is a good approximation to the neurological processes that underlie how humans
30
31 select actions in trial and error learning situations, in that exploration occurs, and does so
32
33 through inhibiting tendencies towards greedy (exploitative) action selection (Camerer and Ho,
34
35 1999; Daw et al, 2006; Laureiro-Martinez, Brusoni, Canessa, & Zollo, 2014). However, the
36
37 conditions under which people switch between exploration and exploitation are not yet fully
38
39 understood. Some recent studies have shown that people appear to enter uncertain choice
40
41 situations by exploring (Lea et al, 2011; Steyvers et al, 2009), and the rate of switching back
42
43 and forth between exploration and exploitation decreases as a function of experience.
44
45

46
47 Notwithstanding these broad consistencies, the parameters of the learning process, such as
48
49 exploration and learning rates, differ significantly between people. Recent research has
50
51 focused attention on the sources of these differences, using psychometric measures of
52
53 cognitive ability (Steyvers, et al, 2009), and neurological differences (Badre, Doll, Long, &
54
55 Frank, 2012), as well as differences based on dispositions to seek or avoid novelty (Payzan-
56
57 LeNestour, & Bossaerts, 2012). These findings provide empirical opportunities as the
58
59
60

1
2
3 individual differences can be measured and used as proxies (for parameters of the learning
4 process) to investigate whether predictions about how variations in the parameters of the
5 learning process produce expected results when coupled with particular properties of the task
6 environment.
7
8
9

10
11
12 In many organizational applications, it is the organizational unit as a whole that is assumed to
13 undertake reinforcement learning, with exploration. One way to justify this is to appeal to the
14 notion of a powerful key actor as representing the organization. In this case, the results at the
15 individual level can be directly used to justify the modelling assumptions.
16
17
18
19

20
21
22 The other is to impute learning behavior to the organizational level on the basis of
23 hypothesized internal dynamics such as the strengthening of coalitions, or the creation of
24 routines as Cyert and March (1963), and later Nelson and Winter (1982) did. There is a
25 significant body of work that indirectly documents organizational level reinforcement
26 learning through the phenomenon of organizational experience curves (e.g. Argote, 2012;
27 Zollo and Winter, 2002). These depict an improvement in performance with experience, but
28 with the rate of increase slowing down with experience. Since this pattern of performance
29 improvement with experience is a hallmark of reinforcement learning processes, the former
30 provides indirect evidence for the latter.
31
32
33
34
35
36
37
38
39
40
41
42

43 How organizations balance exploration and exploitation has also been an active area of
44 research since March (1991) first introduced the trade-off to the organizational literature.
45 Under the term “ambidexterity” (O’Reilly and Tushman, 2013) a significant empirical
46 literature now exists that includes cases as well as some large sample evidence on how
47 organizations attempt to balance exploration and exploitation- leaving little doubt that the
48 attempt is frequently and deliberately made, even though is not always successful (Brown and
49 Eisenhardt, 1997; Raisch et al., (2009), and also Gavetti et al., (2012) for a recent review of
50 the literature inspired by the behavioral theory of the firm).
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5 In sum, it would appear the behavioral principles that underlie multi-armed bandit models -
6 such as reinforcement learning and explorative choices- rest on reasonable empirical
7 foundations, whether used to model individual or organizational behavior in contexts where
8 experience is necessary to generate information about the value of alternatives. As the
9 research in psychology and organizations converges towards a better understanding of when
10 exploration is triggered, it may be possible to incorporate these refinements into the next
11 generation of models.
12
13
14
15
16
17
18
19

20 21 ***Key trade-offs and insights*** 22

23 Having examined the behavioral plausibility of its assumptions, we now turn to the key
24 underlying trade-offs and insights generated by this family of models. These insights are
25 worth investigating empirically. There are two key trade-offs that arise in the bandit model.
26
27 First, there is the canonical trade-off between exploitation of current knowledge and
28 exploration of new knowledge (Holland, 1975): The agent may rely on current knowledge by
29 selecting the arm known to produce a good outcome (exploitation) or choose an apparently
30 inferior arm in the hope of attaining a superior outcome and thereby improving future payoffs
31 (exploration). Because initial representations of agents usually do not correspond to the task
32 environment there is a positive value of exploration because it prevents agents from settling
33 on an inferior alternative. However, agents may also overinvest in exploration – that is, they
34 direct far too much effort on choosing inferior alternatives, without experiencing substantial
35 benefits that will improve their future (Fernie, & Tunney, 2006; Fridberg et al. 2010; Kjome
36 et al., 2010; Premkumar et al., 2008; Steingroever et al, 2012; Wood et al., 2005).
37
38
39
40
41
42
43
44
45
46
47
48
49
50

51 Figure 1 shows in a stylized manner the standard result on the value of exploration in a multi-
52 armed bandit model with a stable task environment. On the vertical axis, the performance of
53 agents is reported. Note that the performance measure used here is the cumulative payoffs
54 over all trials. This represents settings in which agents face persistent selection pressures
55
56
57
58
59
60

(Denrell and March, 2001; Levinthal and Posen, 2007). One could also simply report performance in the n -th trial, or the fraction of cases in which the best alternative has been found by trial n . The horizontal axis plots different values of the exploration parameter in the softmax choice rule, with low values corresponding to low levels of exploration. Figure 1 shows that agents underinvesting in exploration suffer performance shortfalls. They do not explore enough and thereby settle on an inferior alternative. To pick up again on the restaurant example, an agent settling for the first restaurant with decent food as her favourite place misses out on many superior establishments. In contrast, agents with a very high exploration level spend too much on exploring inferior alternatives. That is, she visits and spends money on many bad restaurants again and again in the hope that things improve. Agents that strike exactly the right balance between exploration and exploitation show superior performance because they identify the best alternative (by providing for sufficient exploration) and firmly hold on to it (by providing for enough exploitation).

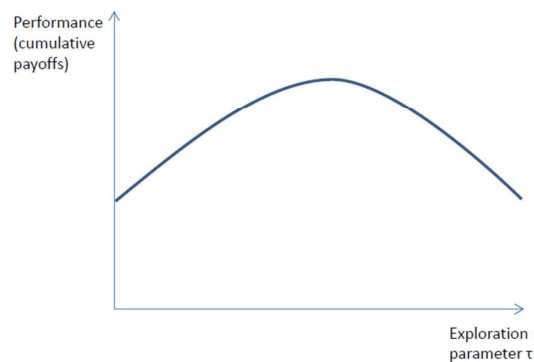


Figure 1: The optimal balance between exploitation and exploration

All models using the multi-armed bandit implicitly or explicitly feature the exploration-exploitation trade-off, and some may explicitly examine moderators that strengthen or weaken the forces generating this trade-off. For instance, an important paper that has analysed the location of the optimal level of exploration as a function of turbulence in the task environment is Posen and Levinthal (2012). They show that, optimally exploration should

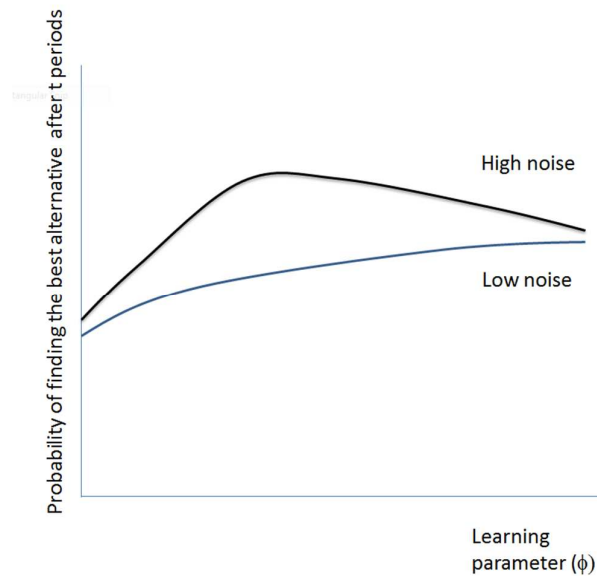
1
2
3 decline as the possibility of changes to the payoff distribution for the arms in the bandit
4
5 increases to high levels.
6
7

8
9 The second trade-off has to do with the potential pitfalls of learning too rapidly in
10 environments with noisy feedback (Denrell, 2007). Because feedback is noisy – payoffs are
11 drawn from a probability distribution – the agent is confronted by “false-positives” and
12 “false-negatives”, i.e., when the value of an action is believed to be higher than it really is and
13 vice versa. Clearly some learning is better than no learning at all, but as several organizational
14 models point out, rapid learning is more susceptible to superstitious learning and mistakes, in
15 which the agent may , selectthe wrong alternative by learning too rapidly from false-positives
16 and false-negatives. The agent can superstitiously associate purely chance outcomes with her
17 own chosen actions (Blanco, Matute, and Vadillo, 2011; Lave and March, 1975). (Consider
18 the proverbial “lucky socks” which get associated with positive outcomes by an individual
19 who happened to wear them to many good outcome situations).
20
21
22
23
24
25
26
27
28
29
30
31

32
33 Papers that have examined the consequence of this learning problem show that agents can
34 learn to become risk averse over time (March 1996; Denrell and March, 2001) and thereby
35 may systematically underinvest into higher-variance, more innovative outcomes. The related
36 “hot stove” effect (Denrell and March, 2001) identifies an important asymmetry between false
37 positive and false negative learning: agents will stop selecting alternatives associated with
38 false-negative outcomes and thereby may miss out on superior alternatives, whereas false
39 positives will likely be corrected for as the falsely attractive alternative is tried again. Thus a
40 bad restaurant that generates a lucky good meal (false positive) will be found out on being
41 visited again, but a good restaurant that produces an unlucky bad meal (false negative), may
42 never be visited again.
43
44
45
46
47
48
49
50
51
52

53
54
55 Figure 2 shows this second trade-off in a stylized manner. The vertical axis gives the
56 probability of having found the best alternative after a number of trials (obtained by
57
58
59
60

1
2
3 computing the fraction of agents in a simulation who have chosen the correct alternative), and
4 the horizontal axis shows increasing rates of learning.⁷ As the figure shows, in noisy
5 environments, rapid learning can do worse than slower learning.
6
7
8
9
10



11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31 **Figure 2: Learning rates and performance in different task environments**

32
33
34
35 The insights from the bandit model for the designers of learning systems, including
36 organizations, can be inferred from these trade-offs. The parameters of the learning process
37 crucially determine success when entering poorly understood and noisy task environments. In
38 particular, avoiding the extremes of very low or very high exploration, as well as very low or
39 very high learning rates, can help promote valid learning and successful choice of alternatives.
40
41
42
43
44
45
46

47 The trade-offs and insights generated by the multi-arm bandit problem illustrate the point we
48 made at the beginning that formal models have the ability to generate systematic insight in
49 ways that verbal theorizing may not be able to. As we will note in the final section, the
50
51
52
53
54

55 ⁷ A moderate level of exploration is assumed. At very high levels of exploration, the effects of learning
56 are suppressed, since choices do not discriminate among alternatives, regardless of what has been
57 learned about them.
58
59
60

1
2
3 precision of these predictions suggests the need and appropriateness of laboratory methods for
4 testing them.
5
6
7

8
9 ***b. Models of search on rugged landscapes***

10
11 In addition to the multi-arm bandit problem, modellers examine individual learning and
12 adaptation using the so-called ‘NK’ models. These models examine organizational decision
13 making problems that are less structured than the ones that multi-arm bandit problems
14 examine. Decisions about new products or processes, new business strategies, or new
15 organizational forms and designs that are often attributed to a recombination of existing
16 resources, technologies, or design elements (Fleming, 2001; Denrell et al., 2003; Lippman
17 and Rumelt, 2003) are examples of such problems. NK models differ from the bandit
18 problems in three important ways. First, unlike bandit problems, the NK models do not
19 assume that the set of possible actions is given and known by the agent. That is, the
20 representation of the task environment does not include all possible actions. Second, actions
21 that differ in payoffs are not equally spaced from one another. Third, in choice situations with
22 multiple pay-off relevant elements, the best option on one element often depends on choices
23 made about other elements. In other words, there is interdependence between different choice
24 elements.
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40

41 The NK model was originally developed in evolutionary biology (Wright, 1932; Kauffman
42 and Weinberger, 1989; Kauffman, 1993). Levinthal (1997) introduced the model to the
43 organizations and management literature to study the interrelationship between adaptation and
44 selection of organizational forms. The model takes its name from the two key parameters of
45 the model that establish the task environment: The N parameter of the model represents the
46 performance-relevant choice elements, while K is a parameter influencing the
47 interdependencies among these elements.
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 An everyday example helps to provide the key intuition behind most NK models. Suppose the
4 design of a car involves choices about four key design elements (the N in the model): engine
5 size, overall shape, placement of trunk and colour. If there are three choices for each design
6 element than the space of actions contains $3 \times 3 \times 3 \times 3 = 81$ points, each of which has an
7 associated performance level (this could be willingness to pay for an average car buyer). If
8 the best choice on each of these parameters is independent of the other element choices (e.g.
9 the best colour does not depend on the shape being chosen), then the performance landscape
10 has a single global peak. In modelling terms, this means that there are no interdependencies
11 and K is therefore 0. Even an extremely myopic search process, which never looks beyond
12 points in the immediate proximity of the current point, eventually finds this peak. Once we
13 find the best colour, for instance, we can hold that constant; find the best shape and so on,
14 sampling at most 12 points in this systematic manner.
15
16
17
18
19
20
21
22
23
24
25
26
27
28

29 However, if the best choice on one element depends on what has been chosen for the others –
30 the best colour actually does depend on the shape, which in turn depends on engine size and
31 so on, then the performance landscape will now have many points that are local peaks; they
32 are better than all their neighbouring points but not necessarily the best in the design space
33 overall. The K parameter regulates the degree of interdependencies and the higher the K, the
34 more local peaks and the more “rugged” is the design space. One will now need to sample
35 many more points and there may be no guarantee of finding the best point, unless all 81
36 points are sampled. Worse, a myopic search process that only looks at points in the proximity
37 of the current point (local search) will often get “stuck” on a local peak. The key difference
38 between the two situations lies in the existence of slopes; in the first case the slopes point
39 infallibly to the global peak (no interdependence, $K = 0$), but not in the second case
40 (interdependence, $K > 0$).
41
42
43
44
45
46
47
48
49
50
51
52
53
54

55 The NK model has found wide applicability in organizational research, because many
56 managerial problems such as new product development, process innovation and
57
58
59
60

1
2
3 organizational routines, strategy formulation, or the design of organizational forms, activity
4 systems, and HRM systems appear to share a common structure: In all of these tasks,
5 managers must make choices over many elements that jointly and interdependently contribute
6 to overall performance, much as the design features in our car example do. In the following,
7 we outline how the NK model depicts the task environment, the agent's representation of the
8 task environment, the choice process, and the transformation of representations and choices
9 over time.
10
11
12
13
14
15
16
17
18

19 1. Task Environment:

20 The NK model depicts a complex combinatorial task (Kauffman & Weinberger, 1989). The N
21 parameter captures the number of performance-relevant task elements. For example, N could
22 represent product features, resources and capabilities, or distinct elements of a business policy.
23 A conventional simplifying assumption is that these elements are usually binary – they are
24 either present (1) or absent (0). To pick up on the car example above, the task of designing a
25 car could be described as choosing the car body configuration (hatchback or sedan), the
26 colour (black or white), and the engine type (internal combustion or electric motor). Each
27 possible action in this task environment is a configuration or combination of these elements
28 and the entire space of actions then contains 2^N combinations with distinct performance
29 consequences. For example, with $N = 3$, the entire task environments consists of 8 distinct
30 combinations, each associated with a unique binary string (e.g., [101], representing a sedan (1)
31 with black colour (0) and an electric motor (1)). Combinations represent alternative product
32 designs differing in features. Product [111] and product [000] offer radically different features,
33 while product [110] is more similar to the first product, differing in only a single element. To
34 ease exposition we will refer to a combination of choices such as [111] as a “policy”
35 alternative, or in brief a policy. A policy thus represents an action in the task environment,
36 and can be thought of as the equivalent to an arm in a multi-armed bandit model.
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 The NK model offers considerable flexibility so that task environments with varying levels of
4 interdependence between its elements can be represented. This is achieved through a
5 computational algorithm for the construction of the performance functions or “fitness
6 landscapes”; the mapping of policy alternatives onto performance values that accounts for
7 interdependencies among its elements. This provides the researcher with a simple way to vary
8 the complexity of the task environment (Simon, 1962) or problem difficulty (Page, 1996).
9 Specifically, parameter K regulates the level of interdependencies among task elements in the
10 payoff function. The payoff contribution of each element e_i in a policy is affected by both the
11 state of the element itself and the states of K other elements. Payoff contributions for
12 elements are drawn from a (usually uniform) probability distribution and the performance of
13 the entire policy alternative is then simply the average of all payoff contributions. For
14 example, with $K = 1$, payoff contribution of the car body also depends on the engine type-
15 indicating interactions between these choices – large engines mean car bodies with less
16 passenger room, etc.

17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33 Generally, with higher values of K, there are more local peaks and performance differences
34 among neighbouring policy alternatives, differing only in a single element, become more
35 pronounced (Kauffman, 1993; Rivkin, 2000). This gives the landscape the property of
36 ruggedness, with lots of peaks and surrounding valleys.

37
38
39
40
41 In many NK models, the interdependency structure is often randomly determined- any K
42 other elements may influence the payoff contribution of a choice. Alternatively, the modeller
43 may also choose to impose (by assumption) a specific structure on the task environment if the
44 purpose of the analysis is to study the implications of these differences. For instance, the
45 consequences of the degree of task decomposability or its hierarchical properties can be
46 explored in this way (e.g. Ethiraj and Levinthal, 2004; Rivkin and Siggelkow, 2007;
47 Ghemawat and Levinthal, 2008).

58 2. Representation:

1
2
3 As with any model of bounded rationality, the NK models assumes that agents do not have
4 accurate representations of their task environment. Given the complexity and structure of the
5 Task Environment, a crucial assumption in the application of the NK model to organizational
6 and management problems is that an agent only constructs a limited representation of a small
7 subset of the 2^N possible combinations and their associated outcomes *at any point in time*.

8
9 Usually this subset is generated through a search process defined in terms of the
10 neighbourhood of the status quo alternative. Thus, the agent's representation does not include
11 a full listing of all possible policy alternatives and their associated payoffs. For instance, the
12 designer of the automobile is assumed to not know the payoffs associated with all possible
13 combinations of design choices (such as black, sedan, medium engine, big trunk etc.).

14
15
16
17
18
19
20
21
22
23
24
25 In the majority of organizational NK models the agent only represents at any point in time the
26 current policy and a randomly selected proximate policy (which differs from current policy in
27 one element) , and the associated performance outcomes of these two policies. For example,
28 the car designer may consider a car configuration that changes the colour from white to black,
29 while keeping other elements constant. Thus the agent's representation at any point in time
30 may be thought of as consisting of $\{a_{1,t}, a_{-1,t}\}$ and a corresponding $\{\pi_{1,t}, \pi_{-1,t}\}$, where
31 these are the current policy alternative (location) and a randomly selected one-step neighbour
32 of the current policy, and their associated payoffs (which are accurately known). The
33 determination of the payoff of the neighbouring policy combination is often assumed to occur
34 in an (un-modelled) "offline" process, which one may imagine features either an experiment
35 within a "model" of the problem held by the agent, or an experiment that the agent actually
36 undertakes but which can be reversed.

37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52 There are some important exceptions to local search. For instance, one may allow for more
53 "distant" search or long jumps where multiple elements are changed at the same time (Cyert
54 and March, 1963: 170; Levinthal, 1997). In models with long jumps (e.g. Levinthal, 1997),
55
56
57
58
59
60

1
2
3 the representations may thus (also) include a randomly selected non-proximate combination
4 and an accurate associated payoff. For instance, the car designer may consider changing both
5 the colour and the engine type, while keeping the car body as it is. Gavetti and Levinthal
6 (2000) offer a depiction of agent's representations of a landscape in which the agent has an
7 initial coarse but unbiased representation of the landscape – they do not know the associated
8 payoff for every combination of elements (i.e. policy alternative), but they can see aggregate
9 payoffs for clusters of combinations of elements. To illustrate, the car designer might broadly
10 understand that streamlined shapes and red colour seem to do better when put together, but
11 not the various ways in which shape and colour interact with other elements (such as engine
12 size, trunk size, weight).

13
14
15 Relatedly, the representation could be inaccurate, as the associated payoff of the policy may
16 not reflect their true payoffs. Knudsen and Levinthal (2007) model a search process where the
17 associated payoffs in the representation may be subject to evaluation errors. The agent
18 generates a local alternative to the status quo, and the evaluation of its performance is subject
19 to an error function to capture the “screening” ability of the agent, with less able agents
20 making more mistakes in judging the payoff of an alternative. This may lead to adoption of
21 inferior alternatives in the search process.

22
23
24 More competent, less boundedly rational decision-makers might be modelled as having
25 knowledge of more than one proximate policy and their associated outcomes in addition to
26 the current policy in their representation. For instance, in the organizational models developed
27 by Rivkin and Siggelkow (2003; 2005; 2006) a more competent manager has a greater
28 number of policy alternatives in mind. This captures the degree of bounded rationality, i.e., a
29 more competent manager develops and evaluates a greater number of policy alternatives.

30
31
32 In sum, Representations of agents in NK models vary in terms of scope (how many and which
33 policy alternatives are in it at any point in time) and accuracy (whether the associated payoffs

1
2
3 are known and valid). However, it is important to emphasize that the notion of representation
4
5 is fairly limited in the NK model, because agents do not actively learn about interactions
6
7 among choice elements to guide search in the task environment, and represent only a very
8
9 small portion of the task environment at any point in time
10

11 12 13 3. Choice process

14
15 The choice process is usually assumed to follow a maximizing choice rule, selecting the
16
17 policy with the best payoff in the representation. As we have noted, this representation may
18
19 or may not be accurate, depending on whether evaluation errors occur (e.g. Knudsen and
20
21 Levinthal, 2007).
22

23 24 25 4. Transformation procedure:

26
27 The set of alternatives in the agent's representation is modified through a process of
28
29 recombinant search (Fleming, 2001), involving the sampling of new alternative policies by
30
31 recombining individual elements. The baseline assumption here is the notion of local
32
33 neighbourhood search – agents seek new alternatives in close proximity to and anchored on
34
35 the status quo. This is because local alternatives are easier to evaluate for boundedly rational
36
37 decision-makers. More distant search which involves changing multiple policy elements at
38
39 the same time may also be allowed for. The relative frequency of these types of search, and
40
41 their triggers are often the key features studied in NK models.
42

43
44
45 There is learning in the NK models, because agents change their representations and hence
46
47 actions as a consequence of previous actions taken and payoffs achieved. However, unlike
48
49 reinforcement learning in which agents obey the Law of Effect, in NK models agents learn
50
51 based on the principle of *(Local) Hill Climbing* - where they always search for improvements
52
53 relative to their current situation, often in their immediate neighbourhood (Levinthal, 1997).
54
55

56 57 58 *Behavioral plausibility of NK models*

1
2
3 While individual behavior has been extensively studied in experimental work on human
4 problem solving (c.f., Newell and Simon, 1972; Kotovsky et al., 1985; Langley et al., 2014),
5 these investigations have generally been restricted to problems that have clearly defined goals
6 and the agents know when these have been accomplished. In contrast, the NK model assumes
7 that the solution – the global optimum – is not known- neither its location nor its value (c.f.
8 Selten, 2001). This leads to two key behavioral assumptions of the NK model of individual
9 agent adaptation - **local search** and **exploratory choice**.
10
11
12
13
14
15
16
17
18

19 Experimental studies by Busemeyer et al. (1986) and Busemeyer and Myung (1987) on
20 resource allocation tasks with interdependencies found strong evidence for local hill-climbing
21 behavior in both studies. Further, participants tend to select a similar direction in allocating
22 resources if performance increased, while failure induced search in a new direction. Billinger,
23 Stieglitz and Schumacher (2014) studied search behavior in NK task environments with
24 various levels of complexity. They found strong evidence for the primacy of local search. Yet,
25 they also found evidence for a form of adaptive, **problemistic search**: Individuals gradually
26 search more distant, exploratory alternatives as a response to poor performance. These results
27 suggest that the search for new alternatives may indeed be sensitive to performance feedback,
28 as much of the empirical research in the organizational learning literature has assumed (Greve,
29 2003), but which is a notion that has not been explored to any great depth in the NK
30 modelling literature. The studies also showed that the search process does not seem to
31 produce a deep understanding of the overall landscape, as is implicitly assumed in the
32 principle of local search.
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48

49 Finally, the broader experimental evidence (not based on NK tasks) also supports the idea of
50 local search. There is evidence that (1) the starting point matters because it defines the region
51 in which people will begin searching (2) the more favourable the initial feedback is the more
52 localized search becomes, such that the region of search becomes smaller, and finally that (3)
53 local search reduces the region in which new experiences are sampled, and this prevents the
54
55
56
57
58
59
60

1
2
3 decision maker from reaching their optimal preference point (Ariely, Loewenstein, & Prelec,
4
5 2003; Hoeffler, Ariely, & West, 2006; Klayman and Ha, 1987).
6
7

8
9 Gavetti et al. (2012) note in their recent review of the literature that there is extensive
10
11 evidence that organizational change is triggered by performance feedback and that low
12
13 performance stimulates search (also see Greve, 1998; 2003). We have already noted the
14
15 significant empirical literature on organizational ambidexterity, and the evidence it produces
16
17 for explorative behavior at the organizational level. However, there is limited direct evidence
18
19 of the local nature of organizational search and the conditions under which it becomes less
20
21 local (e.g. Stuart and Podolny, 1996; Fleming and Sorenson 2004; Katila and Ahuja, 2002;
22
23 Rosenkopf and Nerkar, 2001).
24
25

26
27 In sum, the key behavioral assumptions of the NK model of individual agent adaptation-
28
29 namely local search and exploratory choice appear to rest on good empirical foundations.
30
31 However, the experimental and empirical literature also finds to a larger extent than assumed
32
33 in the models that search is adaptive and guided by performance feedback. The results here
34
35 speak in favour of the notion of problemistic search as proposed by Simon (1955), Cyert and
36
37 March (1963), and Greve (2003), and suggest that NK models could take on this property as a
38
39 modelling feature
40
41

42 43 ***Key trade-offs and insights***

44
45 After examining behavioral plausibility we now discuss the trade-offs and insights from the
46
47 NK model that we believe offer promising avenues for empirical research.
48
49

50
51 It is well known that finding the global optimum on an NK landscape is “NP hard” (Rivkin,
52
53 2000) which essentially means that it is impossible for a decision-maker to actively try out all
54
55 possible configurations. For example, the problem of designing cars obviously entails many
56
57 more elements than just deciding on car body, colour, and engine type. Thus organizational
58
59

1
2
3 scholars are most interested in the search process that unfolds in NK landscapes as well the
4 extent to which agents in the model succeed at finding better choices. The major results from
5 models of search in NK landscapes centre on the question of how properties of the Task
6 Environment – in particular the ruggedness of the performance landscape – impact the
7 effectiveness and outcomes of various search procedures. In a simple landscape with no
8 interdependence between choices ($K = 0$) local search inevitably leads all agents to the global
9 optimum. For rugged landscapes with multiple local peaks ($K > 0$) agents facing the same
10 task environment are likely to end up at different local optima (which may differ in
11 performance properties). The initial Representation (i.e. the starting location) influences
12 where in the landscape an agent searches locally for performance improvements. This leads
13 to strong path-dependency in the subsequent search process and the associated
14 Transformation of the initial Representation. Because of these path-dependencies,
15 heterogeneity in policies and in performance emerges as stable outcomes in rugged
16 landscapes (Levinthal, 1997).
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32

33 A central challenge for boundedly rational agents in rugged NK landscapes is therefore to
34 prevent getting stuck on a low-performing local optimum and to increase the likelihood of
35 identifying the global optimum. The models in the literature have identified two broad
36 conditions to achieve this. First, an agent must be able to broaden search beyond the
37 immediate neighbourhood of an existing alternative (e.g., develop more elaborate
38 representations that go beyond the neighbourhood, in order to avoid trapping on local peaks).
39 Second, the agent needs to be able to stabilize search around a better-performing alternative in
40 choice selection. Readers will recognize that in essence, this is a version of the exploration-
41 exploitation trade-off that was introduced in our discussion of bandit models. In both the NK
42 and multi-armed bandit, exploration refers to a process of seeking alternatives that are not
43 indicated by the current representation, though the manner in which the task environment and
44 its representation are modelled are of course very different.
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 Figure 3 offers a stylized representation of this trade-off in an NK model. Early in the search
4 process, local search rapidly increases performance from the starting position, because the
5 task environment typically offers some refinement potential. However, over time, local search
6 exhausts refinement potential as agents get stuck on local peaks. Further performance
7 improvements based on local search are then no longer possible. In contrast, an adaptive
8 search rule that triggers gradually more distant search as a response to failure and reverts back
9 to local search after success (Billinger et al. 2014) improves performance more slowly, but
10 tends to outperform in the long run as distant search helps agents to move off low-performing
11 local peaks.
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

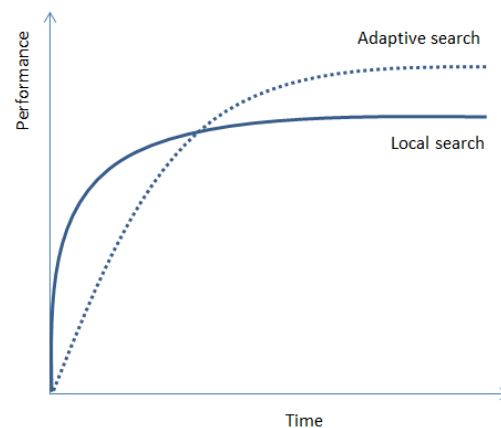


Figure 3: Performance of local and adaptive search in complex task environments

The modelling literature has explored several individual and organizational attributes that may help to stimulate adaptive search. Indeed, it may be fair to say that the philosophy behind most of the models in this tradition has been to take particular organizational features (or their consequences) and translate them into behaviorally plausible search processes that break free of local hill climbing, such as analogical reasoning – in which solutions that have worked in one domain are transplanted to another (Gavetti and Levinthal, 2000; Gavetti, Rivkin and Levinthal, 2005); divide and conquer heuristics in which a problem is broken down into smaller sub-problems, on the assumption that each can be solved in isolation (Ethiraj and

1
2
3 Levinthal, 2004); and means-end analysis, in which one works backwards from the desired
4
5 result to current best action (Denrell, Fang and Levinthal, 2004).
6
7

8
9 For instance, Gavetti and Levinthal (2000) show that simple, low dimensional
10
11 Representations – limited knowledge about the rugged performance landscape – help agents
12
13 to identify more attractive starting positions for further local search. In Rivkin and
14
15 Siggelkow's (2003) model, less boundedly rational managers have better representations of
16
17 the task environment in the neighbourhood of the current policy. Knudsen and Levinthal
18
19 (2007) propose an approach in which agents evaluate a new policy in the neighbourhood with
20
21 error and thereby sometimes adopt a lower-performing action. This allows them to move off
22
23 low-performing peaks. This is beneficial in rugged landscapes. However, errors in the
24
25 selection of actions should not be too large, since agents may then be incapable of holding on
26
27 to superior choices.
28
29

30
31 As with the multi-armed bandit model, the insights from the NK model for the designers of
32
33 learning systems, including organizations can be inferred from this central trade-off between
34
35 local and non-local search. When entering poorly understood complex task environments, the
36
37 parameters of the search process determine success. In particular, avoiding being trapped by
38
39 local peaks with low payoffs, while also avoiding falling off peaks with good payoffs, is
40
41 critical. The models suggest that organizational processes that support analogical reasoning,
42
43 the use of divide and conquer heuristics, and sampling practices in surrounding environments,
44
45 produce this balance and therefore offer advantages in the search process. These ideas can and
46
47 should be empirically tested.
48
49

50
51 We turn next to models of mutual adjustment.
52
53

54 55 56 **2. Models of Mutual adjustment** 57 58 59 60

1
2
3 Models of mutual adjustment between agents capture important organizational phenomena
4 such as inter-unit or inter-agent collaboration, socialization and knowledge exchange,
5 imitation, and delegation and control. In fact, any organizational process in which multiple
6 agents influence each other's adaptation processes would qualify as a candidate to be
7 analysed using models of mutual adjustment. For instance, managers of different product
8 divisions within a company seeking to realize synergies by coordinating their marketing
9 campaigns; teams of engineers developing different sub-systems; new employees adjusting to
10 an organization's culture and vice versa. Organizational models of mutual adjustment either
11 explicitly or implicitly build on models of individual adaptation. Many recent models of
12 mutual adjustment in organizations have relied on either NK or multi-armed bandit models to
13 capture individual agent behavior, but also using some form of mutual adjustment
14 assumptions between multiple agents.
15
16
17
18
19
20
21
22
23
24
25
26
27
28

29 In each of these models, two or more interacting agents jointly face either a bandit or an NK
30 task; therefore the discussion of Representations, Choice and Transformation processes from
31 the family of individual agent adaptation models carries over exactly here (and we do not
32 repeat those). The key difference lies in the nature of the Task Environment that each agent
33 faces, which now includes the other agent(s)- possibly without the focal agent being aware of
34 this.
35
36
37
38
39
40
41
42

43 The discussion of behavioral plausibility of mutual adjustment models is also largely the same
44 as in individual adaptation models, with the important exception of the task environment,
45 which as noted above now includes the other agent(s). The social psychological evidence
46 shows that we treat risky interactions with nature differently than we do interactions with
47 other people, implying that social risk is different from non-social risk (Bohnet and
48 Zeckerhauser, 2004; Blount, 1995). However, without exception, the mutual adaptation
49 models that we review below assume that the agents act exactly as they would in a task
50 environment in which they were alone.
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5 There are two justifications for this. First from a modelling point of view, it is plausible that
6
7 the overall learning or search process of an agent is the same regardless of whether their
8
9 interaction is with Nature or another agent, but the parameters of the search process may
10
11 change. For instance, agents may explore less, learn more slowly or search more locally
12
13 when they know there are other agents mutually adjusting to them, but the basic processes of
14
15 local search and reinforcement learning would not change. Second, one may interpret the
16
17 model as pertaining to situations of such task environment complexity as to render the agent
18
19 effectively incapable of meaningfully recognizing the existence of other agents (for instance
20
21 if there are a large number of other agents one cannot easily communicate with). As Simon
22
23 (1990) famously remarked, the task environment and the agent's internal processes work
24
25 jointly like the blades of a scissor to produce observable behavior. Above certain levels of
26
27 environmental complexity, the agents' behavior, whatever the internal processes, may look
28
29 the same as it does in simpler contexts.
30
31

32
33 Besides the inclusion of other agents, there are other secondary variations in the task
34
35 environments (namely interdependence, communication and imitation possibilities) of these
36
37 models such that it is useful to discuss the key trade-offs and results from models of coupled
38
39 learning/search, social learning, and imitation separately within this model family.
40
41

42
43
44 i. Key insights from coupled learning/ search models

45
46 One important class of mutual adjustment models, which we call "coupled learning/search
47
48 models", assumes that agents are engaged in a search for combinations of good
49
50 interdependent choices, but do so under communication constraints, arising from
51
52 specialization, spatial distance, concerns about alignment of interests, etc. Communication
53
54 constraints are a critical element of these models. Absent such constraints, coordination of
55
56 search is a trivial problem.
57
58
59
60

1
2
3 On practical grounds these models of mutual adjustment are equivalent to models of
4 individual adaptation in a task environment that just so happens to include another agent (but
5 the focal agent typically doesn't really "know" this). The feedback each agent sees in any
6 given time period is thus partly dependent on the other agent's (unobserved) actions, resulting
7 in the possibility of learning inappropriate lessons from feedback. Thus, the challenge for the
8 agents is to coordinate their search processes, in a multi-period situation with feedback in
9 every period, given interdependence and communication constraints.
10
11
12
13
14
15
16
17
18

19 Lounamaa and March (1987) presented an early model that explicitly examines mutual
20 adjustment processes in search and learning, and pointed to the hazards of rapid learning
21 (similar to that discussed in the section on bandit models of individual adaptation). Lave and
22 March (1975) also reported the hazards of rapid learning in matching coordination games and
23 offer similar explanations (i.e., superstitious learning- in which agents learn inappropriately
24 from false negatives and false positives.
25
26
27
28
29
30
31
32

33 Puranam and Swamy (Forthcoming) used a coupled multi-armed bandit model (i.e., the
34 payoff to each agent from selecting an arm also depends on what arm the other agent has
35 selected) to show that even incorrect initial Representations of the Task environment held by
36 both agents may be better than representations that do not discriminate at all between the
37 alternative choices. This is because the latter are more likely to generate "false negatives"-
38 suppose the first agent takes the correct action accidentally, but if the second agent does not,
39 the feedback on joint actions indicates a failure, which causes the first agent to also discard
40 their own correct choice. This effect grows stronger as the learning rates of agent's increase
41 (as the agents learn more rapidly from feedback that includes false negatives), and weaker as
42 the communication between the agents improves (so that they recognize a false negative as
43 such). Knudsen and Srikanth (2014) developed this line of argument further by showing that
44 coupled learning processes needs to balance efforts at reducing mutual confusion by aligning
45 Representation of the Task Environment (and thereby restricting exploration) and at reducing
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 joint myopia by slower learning or reduced knowledge transfer (and thus promoting
4 exploration). These results are a manifestation of the trade-off reported in Figure 2 for the
5 multi-armed bandit model.
6
7

8
9
10
11 Lee and Puranam (Forthcoming) model coupled learning in a hierarchy, in which a
12 superior defines strategy and a subordinate implements it. The superior thus has a
13 view on which arm to choose in a bandit task, transmits this view to a subordinate,
14 who then selects the arm but with some probability of error (i.e. imprecise
15 implementation). The superior then revises beliefs based on observed outcomes,
16 without knowing the actual action chosen by the subordinate. The central insight in
17 their model is that given this separation between belief (superior's) and action
18 (subordinate's), effective implementation by the subordinate has benefits beyond the
19 well-known effect of enabling exploitation of good strategies. Curtailing exploration
20 by the subordinate paradoxically enables the discovery of better strategies by the
21 organization, as it allows more effective learning from feedback on the value of
22 current strategies. It therefore makes sense, the authors argue, to insist on precise
23 implementation of strategies in organizational hierarchies even when their validity is
24 unknown.
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43

44 Rivkin and Siggelkow have studied coupled search processes on rugged landscapes using the
45 NK model, with agents engaged in local search (2003; 2005; 2006). They find that
46 mechanisms that cause interdependent agents to take each other's payoffs into account (such
47 as broad, inter-unit incentives) prevent one agent from taking actions that harm the
48 performance of the other; but interestingly the absence of such instruments can help prevent
49 premature settlement of the dyad onto a local peak. Siggelkow and Rivkin (2009) extended
50 the analysis to a particular form of a hierarchical coupled learning problem, where higher
51
52
53
54
55
56
57
58
59
60

1
2
3 level choices partially constrain lower level choices, to show that such a coupled search
4
5 process can obscure the real impact of the higher level choices. Thus the impact of choices
6
7 about the first layer of organizational structure (i.e. divisional vs. functional) may be obscured
8
9 by adaptation at the departmental level. For example, the performance over time of an
10
11 organization depends not only on the higher level structure (e.g. divisional vs functional) but
12
13 also on the manner in which lower level units adapt to their own task environments. It cannot
14
15 therefore be used purely as a signal of the effectiveness of the former.
16
17

18
19 In summary, the key insights from this set of models are that in situations of interdependence,
20
21 learning confronts certain unique challenges as well as opportunities. As a consequence,
22
23 certain features of the mutual adjustment process are expected to play a key role in shaping its
24
25 success, such as the rates of learning, the nature of initial beliefs, and the hierarchical
26
27 differences between the agents mutually adjusting to each other. These conjectures can be
28
29 tested empirically.
30
31

32 33 ii. Key insights from social learning models

34
35 A second class of mutual adjustment models allow for communication between independent
36
37 agents, but focus instead on the rates at which agents learn from each other – in particular
38
39 how they may copy or share their representations of the task environment. The actions of the
40
41 agents themselves are not interdependent in these “social learning” models (Bandura, 1977),
42
43 instead each agents payoff depends on their own action, but since all agents are in the same
44
45 environment, they can learn from the successful actions of other agents. The classic model by
46
47 March (1991) on exploration and exploitation is of this form: agents essentially learn from
48
49 superior, better-performing agents in their Task environment, by copying their actions.
50
51

52
53 The models of social learning also yield the learning trade-off obtained in the coupled
54
55 learning literature, i.e., moderate rates of individual learning lead to better representations of
56
57 reality. The reasoning is the same, which is that rapid learning can give rise to superstitious
58
59

1
2
3 learning (while very slow learning may not take advantage of learning opportunities at all).
4
5 This commonality is not surprising given that the key building block for both kinds of models
6
7 is an individual level reinforcement learning process. Subsequent work in this tradition has
8
9 deepened insights around these social learning processes. Miller et al. (2006) and Lazer and
10
11 Friedman (2007) added a social network perspective to the underlying mechanism, with
12
13 individuals only learning from each other if they are socially connected. Fang, Lee and
14
15 Schilling (2010) and Schilling and Fang (2013) developed the social network perspective
16
17 further by showing how subgroup isolation and information distortion enhance exploration in
18
19 organizational learning. Isolated groups with few direct ties between them preserve individual
20
21 variations in representations longer. Isolated groups thereby slow down the learning process,
22
23 leading to more correct representation of the task environment. Likewise, Information
24
25 distortion generates variation as agents learn from each other and has a similar effect.
26
27 In summary, who one is connected to, and at what rate one learns from others as opposed to
28
29 from one's own experiences are factors that should shape the effectiveness of learning among
30
31 a community of learners.
32

33 34 35 iii. Key insights from Imitation/Competitive Learning models

36
37 Finally, several models of mutual adjustments have considered a particular form of social
38
39 learning, which occurs when the interests are not aligned between teacher and learner and
40
41 indeed they may be competing- e.g. the case when laggards in an industry try to imitate the
42
43 leaders. Using the NK model as a framework, the central insight here is that complexity may
44
45 serve as an effective barrier to imitation (Rivkin, 2001; Lenox et al., 2007). Specifically, the
46
47 finding is that partial or imperfect imitation of complex practices (perfect imitation is unlikely
48
49 by definition when complexity is high) may not necessarily improve performance (Rivkin,
50
51 2000).
52

53
54
55 Csaszar and Siggelkow (2010), on the other hand, demonstrate some value to imitation, albeit
56
57 in low complexity environments. They do this by pinpointing two distinct mechanisms
58
59
60

1
2
3 influencing the value of imitation. First, imitators benefit directly from copying superior
4 practices, especially in low-complexity task environments. Second, imitation even of inferior
5 practices is also an instrument for dislodging agents from inferior local peaks and thereby
6 promotes more distant search. Relatedly, Posen et al. (2013) develop the intriguing insight
7 that imperfect imitation may be better than perfect imitation in some situations. Specifically,
8 imperfect imitation allows the imitator to preserve and build upon unique and valuable
9 practices currently not implemented by superior firms.
10
11
12
13
14
15
16
17
18

19 A central element of adaptive rationality is the presence of aspirations that determine when
20 search and imitation stops. Aspirations of an agent may not just be informed by the individual
21 search and learning process –their own historical aspirations (Lant, 1992; Greve, 1998) – but
22 they may also be conditioned socially on the performance of others (Festinger, 1954; Greve,
23 1998). When aspirations are socially formed, they induce a unique mutual adjustment
24 dynamic, since agents that uncover superior alternatives raise the reference point for the entire
25 ecology of learners (Levitt and March, 1988; Lant and Mezias, 1992). This can, in turn,
26 reignite search and learning by others, thereby creating a self-reinforcing dynamic of mutually
27 adjusting aspirations and increasing performance. Thus, competition serves an important role
28 in spurring innovation and imitation, not only because it raises the expected returns to search
29 (as assumed in traditional rational choice models), but also because the success of others
30 raises the aspiration of the focal firm. This is crucial in a task environment in which no one
31 may know the upper bound on what is truly achievable (i.e. the height of the global peak is
32 unknown).
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49

50 These results point to possibilities for empirical examination of the effects of complexity on
51 competition and innovation.
52
53
54

55 **3. Information Aggregation**

56
57
58
59
60

1
2
3 Like mutual adaptation models, in information aggregation models agents are part of each
4 other's task environments. As a family, these models provide a general framework to
5 understand how individual evaluations are aggregated into organizational level outcomes
6
7 (Knudsen & Levinthal 2007; Christensen & Knudsen 2010; Csaszar 2013; Csaszar and
8
9 Eggers, 2013). In the earlier versions of these models, while the agents interact with each
10
11 other, they do not necessarily adapt to each other- their Representations and Choice processes
12
13 typically do not undergo any change over time. Recent papers assume a transformation
14
15 process for changing the agent's representations and choice processes, qualifying this family
16
17 of models for inclusion in our review of adaptive rationality models.
18
19
20
21
22

23 Individual evaluations of options in these models are opinions about the absolute viability or
24 relative attractiveness of alternatives. (If an individual's evaluation is the final one and not
25 subject to appeal, we may call it a decision by the organization). For instance, when
26
27 individual managers in an organization evaluate projects, investment alternatives, or
28
29 candidates to hire, they are generating evaluations (which may or may not be the final
30
31 decision of the organization, depending on the distribution of decision rights, e.g., is the
32
33 relevant manager the CEO?).
34
35
36
37
38

39 The models we discuss in this section study the impact of different ways in which individual
40 evaluations are aggregated into an organizational level decision. These different decision
41 architectures are of course a very important part of organizational structure, and so these
42
43 models may be seen as formalizations of the consequences of different organizational
44
45 structures to aggregate evaluations. The models draw on the seminal work of Sah and Stiglitz
46
47 (1985, 1988) and depict decision-makers as holding fallible beliefs about the value of an
48
49 alternative. The emphasis is therefore clearly on the evaluation, rather than the generation of
50
51 alternatives (Knudsen and Levinthal, 2007).
52
53
54
55
56
57
58
59
60

1
2
3 As with all models of bounded rationality, the representations the agents hold (in this case,
4 about what makes an alternative worthy of positive evaluation) may be flawed. In general,
5 fallible decision-makers could make two types of errors: They may reject a superior
6 alternative (Type I error) or accept an inferior alternative (Type II error). An everyday
7 example may help to convey the intuition. Imagine a group of employees who are screening
8 credit card applications. Each makes an individual evaluation- should an application be
9 accepted or rejected? Each employee individually is fallible, and may make either a Type I or
10 a Type II error. But how does the structure of the group influence its aggregate performance?
11 In other words, should we get each application seen independently by an employee who
12 makes a final decision on it, or should we get an application seen by multiple employees who
13 must all agree before it is accepted? These different decision architectures will have very
14 different aggregate Type I (reject superior alternative) and Type II (accept inferior alternative)
15 error rates.
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30

31 . A decision architecture consists of a set of rules that influence information flows and
32 decision rights in an organization. For example, in a three-member linear hierarchy, a new
33 project is first evaluated by a lower-level manager. Only if she accepts the project, it gets
34 passed on to a middle manager. Otherwise, the project is rejected and not evaluated further.
35 The middle manager, in turn, then also decides to reject it or accept it to send it on to the
36 senior manager for final approval. In contrast, in a decentralized polyarchy, all three
37 managers would evaluate the project and acceptance by one is sufficient for project approval.
38 These building blocks describe a wide range of possible organizational architectures, with
39 many hybrid forms such as committee voting falling between the extreme forms of the
40 hierarchy (all agents must accept and pass on to the next) and the polyarchy (any agent may
41 accept) (Knudsen and Levinthal, 2007; Christensen and Knudsen, 2010; Csaszar, 2013;
42 Csaszar & Eggers, 2013).
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57

58 Task Environment:
59
60

1
2
3 The task environment that each agent faces is one in which alternatives arise typically
4 sequentially rather than simultaneously, and in which there are no interdependencies between
5 alternatives. Each agent faces a stream of projects with some associated payoff value.
6
7 Evaluations involve characterizing these into good and bad projects. For example, in
8
9 considering investment projects, a project could be evaluated as worthy of acceptance if it
10
11 offers a positive net present value.
12
13
14
15

16 17 Representation and Choice process:

18
19 In these models, the representation and choice process are combined into a single modelling
20
21 element called a screening function. Decision-makers may differ in their ability to represent
22
23 accurately the relationship between the attributes of the project and its value, denoting
24
25 differences in skills, experience, or talent. This differential ability is usually modelled as a
26
27 screening function. This is a function that gives us the probability of a project being accepted
28
29 as a function of some underlying characteristic. For instance the characteristics could be
30
31 quality, and agents may differ in their ability to correctly assess this. An ideal screener is one
32
33 who accepts the project with probability=1 if it exceeds the quality threshold, else rejects it.
34
35 Fallible screeners, however, will accept even below threshold quality projects with some
36
37 probability, and reject above threshold quality projects with some probability.
38
39
40

41 42 Transformation process:

43
44 In many models, the screening function of an agent remains constant over time. Christensen
45
46 and Knudsen (2013) introduced a model in which the screening ability of agents develops as a
47
48 function of the decision architecture they are in. Agents are assumed to learn from the
49
50 alternatives that they evaluate. The assumption here is that alternatives represent learning
51
52 opportunities and that more experience translates into a higher screening ability.
53
54

55 56 ***Behavioral plausibility***

1
2
3 Models of information aggregation make few psychological assumptions besides the idea that
4 individual evaluations are often imperfect because of bounded rationality, but may improve
5 with experience. This assumption appears to have strong empirical support (c.f., Kahneman
6 and Klein, 2009). The central behavioral idea here is that agents identify and learn about
7 relationships between objectively identifiable project cues and the associated value of
8 incoming projects. For example, the default risk of a credit application by a company is
9 associated with certain cues (such as current profitability, growth, investments etc.). The
10 behavioral research shows that learning leads to improved screening ability when the
11 relationship between cues and project value are stable, when feedback is rapid and
12 unequivocal (i.e., less noisy), and when opportunities exist for prolonged practice with a
13 similar set of projects (Osman, 2010; 2014; Newell and Shanks, 2014).
14
15
16
17
18
19
20
21
22
23
24

25 26 27 ***Key trade-offs and insights*** 28

29 The models provide many insights into how different aggregation structures convert
30 individual evaluations into organizational decisions.
31
32
33
34

35 Figure 4 offers a stylized depiction of the core result from this class of models. Acceptance
36 probability, here, is a function of the project value: a perfect evaluator would reject all
37 projects with a value below zero and accept all projects with a value above zero. A single
38 agent with imperfect screening ability is depicted as a straight line with a positive slope,
39 indicating that there is a positive probability of rejecting (accepting) good (bad) projects.
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

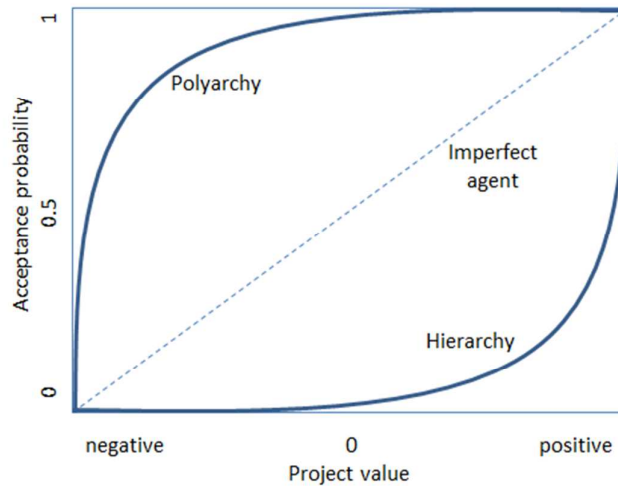


Figure 4: Individual ability and information aggregation in the polyarchy and the hierarchy

In a hierarchy, every agent has to accept an alternative for the entire organization to adopt it. This lowers the probability of accepting bad projects and minimizes Type II error, but this way of aggregating beliefs comes at the costs of a higher rate of Type I errors. In contrast, the acceptance of an alternative by a single agent is all that is required for its adoption in a polyarchy. This results in a higher rate of Type II errors and less Type I errors. Put differently, the polyarchy in the aggregate is more explorative and less risk-averse than the hierarchy (Csaszar, 2012; 2013).

Csaszar and Eggers (2013) extended the information aggregation perspective by also considering how delegation, majority voting, and the averaging of opinions impact information aggregation and the decision-making quality. They show that delegation is the most effective structure when there is diversity of expertise, when accurate delegation is possible (i.e., when it is known who has the expertise), and when there is a good fit between the firm's knowledge and the knowledge required by the environment. In other words, trusting the expert is a good option under the very restrictive conditions identified above. Otherwise, voting or averaging may be the most effective structure. This result formalizes and specifies the notion of wisdom of crowds (Surowiecki, 2005).

1
2
3
4
5 Christensen and Knudsen's (2013) model with transforming screening functions allows
6
7 agents to learn from experience and get better at screening. The basic model structure is
8
9 further enriched by allowing agents not only to accept or reject an alternative, but also to pass
10
11 on an alternative when they do not feel confident about their own judgment. These advances
12
13 therefore mix problems of mutual adjustment and information aggregation to get a richer
14
15 understanding of organizational decision-making over time. A key insight is that if an agent
16
17 only sees projects already screened by others based on his current location in the decision
18
19 architecture, her screening function may transform over time to one which would be
20
21 inappropriate if she were placed elsewhere in the structure.
22
23
24

25 **4. Linkages between Empirical Analysis and Models**

26
27
28 In this section, we conclude by summarizing and discussing the linkages between empirical
29
30 analysis and models. These occur in two directions. First, there is the linkage between
31
32 modelling assumptions and the evidence on actual adaptive behavior by individuals and
33
34 organizations. Second, there is the linkage between model generated results and testable
35
36 empirical implications
37
38
39

40 i. The behavioral plausibility of modelling assumptions

41
42 For each model family, we examined the structure of the models in terms of four common
43
44 elements, the Task Environment, Representations, Choice processes, and Transformation (of
45
46 Representation) processes- with the goal of understanding to what extent the assumptions in
47
48 these models are consistent with the evidence about how individuals and organizations
49
50 behave. Our review paid close attention to the key behavioral assumptions in these models of
51
52 adaptive rationality, and these may be summarized as follows:
53

- 54 a) *Complexity of the Task Environment* to a point where the agent's Representations are
55
56 necessarily incomplete/incorrect; as a consequence agents may not represent the
57
58
59
60

1
2
3 entire search space, or possess knowledge about the payoffs associated with actions,
4
5 or be aware of the existence and actions of other agents that shape their own Task
6
7 Environments, or see the true value of alternatives. As we have noted, this
8
9 assumption has more to do with the translation of an empirical context appropriately
10
11 into a model; certainly there are organizational contexts where these assumptions are
12
13 highly plausible, and others where they are less so.
14
15

16
17 The remaining assumptions pertain to individual psychology or organizational behavior
18
19 relating to the nature of adaptive rationality:
20

- 21 b) *Exploration in choice*: Choice processes within the Representation stray from pure
22
23 exploitation (i.e. maximization or greedy action selection) and feature a degree of
24
25 exploration.
26
27 c) *Problemistic search* to modify representations– search and exploration is triggered by
28
29 performance falling below aspiration levels.
30
31 d) *Local hill climbing to modify representation*: the discovery of new alternatives
32
33 occurs primarily through the principle of local improvements; (a relaxation of this
34
35 tendency and the willingness to engage in distant search may be interpreted as a form
36
37 of exploration).
38
39 e) *Reinforcement learning to modify representations*: the discovery of the associated
40
41 values for known alternatives occurs through the Law of Effect- alternatives with
42
43 better than expected outcomes are more likely to be tried again, those with worse than
44
45 expected outcomes are less likely to be tried again.
46
47
48

49
50 The good news for researchers in the organizational sciences, modelers and otherwise, is that
51
52 the evidence for these assumptions at both the individual and organizational level is in general
53
54 supportive. We have documented the empirical evidence, both experimental and non-
55
56 experimental, that gives credibility to these assumptions in the relevant sections where we
57
58 discussed each model.
59
60

1
2
3
4
5 To be sure, there are places where the evidentiary base is stronger than others, and much
6 replication is necessary before these assumptions can be safely granted the status of axioms.
7
8 For instance, individual level evidence on problemistic search, and organizational level
9 evidence of hill-climbing behavior both can be strengthened. It is still largely a mystery as to
10 when both individuals and organizations display broadly similar (though certainly not
11 identical) adaptive search and learning behavior, and when they do not since it is apparent
12 that the strong unitary actor (e.g. the heroic CEO with fully centralized decision powers) is
13 such a rarity.
14
15
16
17
18
19
20
21
22

23 The aggregation processes within organizations that produce broad similarity across scales
24 still need to be carefully investigated, and is a fascinating research agenda in itself.⁸ As we
25 will note below, scholars interested in this particular research question can draw on work
26 done in social psychology and organizational behavior that has compared group versus
27 individual decisions.
28
29
30
31
32
33
34

35 But in answer to the first question that motivated this review- namely are the assumptions that
36 underlie models of bounded rationality in organizations behaviorally plausible?- we must
37 answer, with cautious optimism, in the affirmative. The behavioral plausibility of modelling
38 assumptions could serve as an encouragement for researchers to put in effort at further
39 refining tests of behavioral regularities so that modellers can strengthen and make more
40 precise the assumptions in their models. Likewise, empirical researchers in organization
41 science, we suggest, should be receptive to at least engaging with these models as sources of
42 testable predictions.
43
44
45
46
47
48
49
50
51
52
53

54
55 ⁸ We mean similarity beyond that induced by the language of description. For instance, if we want to
56 compare individual and organizational screening functions, we must look for similarities beyond those
57 generic to any screening function.
58
59
60

1
2
3 ii. The empirical implications of model results
4

5 If organizational models of (adaptive) bounded rationality fare well in terms of their use of
6 behaviorally plausible assumptions, we should be optimistic that their predictions find
7 empirical support. Yet surprisingly, attempts to empirically test the predictions of these
8 models have been rather limited. In part this may be because of the subtlety involved in
9 interpreting the results of such a test, even if conducted. To elaborate on this, we consider two
10 kinds of tests.
11
12
13
14
15
16
17

18
19 First, when we test the implications drawn from, say a multi-armed bandit or an NK model of
20 individual adaptation, we are testing the predictions resulting from the interaction between a
21 Task Environment with certain assumed properties, and a set of behavioral rules (captured in
22 the nature of Representations, Choice processes and Transformation process). These are of
23 course the famous twin blades of Simon's scissors- the environment and the agent's internal
24 processes jointly produce observable behavior (Simon, 1956; 1990). However since the latter
25 (behavioral rules) were drawn from an existing evidentiary base, the test is first of all about
26 the former (i.e. assumptions made about the task environment), and whether these correspond
27 to the empirical context being used to test the model predictions.
28
29
30
31
32
33
34
35
36
37
38

39 Further, one may also test if variations in the parameters of the search or learning process
40 produce expected results when coupled with particular properties of the task environment.
41 Laboratory experiments, where one can carefully construct environments, appear to be a good
42 method to test model predictions. To be useful and correspond to real world problems, these
43 laboratory based studies could be combined with carefully constructed case studies (c.f.,
44 Edmondson & McManus, 2007 about the utility of mixed methods in fields that are neither
45 very mature nor very new).
46
47
48
49
50
51
52
53
54

55
56 Second, when the Task Environment includes other agents (in models of mutual adjustment
57 and information aggregation), then our tests of model predictions are implicitly testing if the
58
59
60

1
2
3 environments are indeed such as to cause agents in them to be unable to distinguish other
4 agents immersed in the same environment from the environment itself. Put simply, we are
5 testing the assumption that the agents behave non-strategically with respect to each other in
6 the task environment of interest, since that is what the behavioral rules assume in the models
7 of mutual adjustment and information aggregation. The limited data that examines differences
8 in agent behavior when interacting with human actors compared to non-humans (e.g., games
9 where human agents play with a machine or with nature) suggests that this assumption is not
10 valid in all task environments (e.g., Blount, 1995; Bohnet & Zeckerhauser, 2004).

11 Confronting the predictions of these models with such empirical data may lead to a
12 recognition that the behavioral rules used by individuals in simple task environments that they
13 recognize as containing other agents (e.g. Blume, Duffy, Franco, 2009) may be different from
14 more complex situations or in which the other agents fade, as it were, into the task
15 environment (Osman and Ananiadis-Basias, 2013). The empirical literature is somewhat
16 limited in that the task environment faced by decision makers has hitherto not been very
17 complex. It is not clear that the observed difference will generalize to very complex
18 environments. Thus, how much iterative reasoning about other's intentions is both plausible
19 (and tractable) will then be a key issue for modelers and empiricists to consider together, by
20 varying task environments systematically.

21
22 With these observations in mind, we outline the broad contours of the kinds of empirical tests
23 one could conduct (and are indeed being conducted) of the results from the existing models.
24 We outline the correspondence between model parameters and phenomena in naturally
25 occurring data, and these are also summarized in Table 1.

26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52 ----- INSERT TABLE 1 ABOUT HERE -----
53
54

55
56 iii. Suggestions for empirical studies
57
58
59
60

1
2
3 For models of individual adaptation based on multi-armed bandits, the key properties of the
4 task environment are noise and changes in feedback conditions. The key behavioral
5 parameters of the learning process are the extent of exploration and the rate of learning.
6
7 Empirical tests could thus, for example, involve comparing differences in performance
8 between adaptive agents (individuals or organizations) in task environments with varying
9 rates of changes or degrees of noise in feedback (e.g. industries or investment opportunities
10 with varying levels of environmental change or noise in performance feedback) as a function
11 of the agent's tendencies to make choices other than the past successful ones (i.e. exploratory
12 choices) and responsiveness to feedback (i.e. learning rates).
13
14
15
16
17
18
19
20
21
22

23 The key property of the task environment in NK models is the interdependency parameter K ,
24 which creates ruggedness in the performance landscape. The search parameters of interest are
25 those that affect the balance between making exploratory "leaps" on the landscape instead of
26 local hill climbing. Thus, tests of these models would require us to exploit variations in task
27 environments in terms of their interdependence structure (e.g., highly decomposable task
28 environments, such as the production of software, vs. less decomposable ones such as the
29 development of an advertising campaign), along with variations in properties of the decision-
30 making agent that would lead to more or less adherence to local hill climbing behavior.
31
32
33
34
35
36
37
38
39
40

41 A few studies have already been conducted that offer support for the predictions from the NK
42 model. Using patent data, Fleming and Sorenson (2001) found that high levels of technical
43 interdependencies reduce the usefulness of inventive efforts (i.e., tracking model results that
44 show that exploitation is not a useful strategy in complex environments). In a related study,
45 Fleming and Sorenson (2004) showed that investments in basic science are particularly
46 valuable in technological fields with high interdependencies. In such task environments, local
47 search quickly traps inventors on local peaks. Science provides a cognitive representation
48 (Gavetti and Levinthal, 2000) of the technological search space and thereby helps inventors to
49 identify more attractive regions in the landscape. Lenox, Rockart, and Lewin (2010) used
50
51
52
53
54
55
56
57
58
59
60

1
2
3 survey data to directly test their own model that enriches the basic NK setup with
4
5 oligopolistic competition among organizations (Lenox, Rockart, Lewin, 2006). They found
6
7 that average profitability is highest with moderate interdependency, and that the dispersion
8
9 and skewness of profits is especially pronounced with high levels of interdependencies. These
10
11 contributions provide empirical support for NK model predictions about how complexity
12
13 affects managerial and organizational behavior and performance outcomes.
14
15

16
17 Models of mutual adjustment through coupled learning based on the multiple-armed bandit or
18
19 the NK framework create an organization comprising multiple agents. The search process of
20
21 this organization depends on the variance in the parameters of the following features of its
22
23 internal structure: the nature of reward interdependence between the agents, their relative
24
25 rates of learning or exploration, degree of heterogeneity of its networks, and the distribution
26
27 of decision rights between them. These structural features of organizations are observable. By
28
29 jointly examining these attributes as well as variations in the noise and interdependence in
30
31 task environments, the model predictions can be empirically tested. Social learning models,
32
33 through their increasing use of network structure as a key feature of the models, also point to
34
35 an obvious empirical testing strategy: to link network structure to innovation and knowledge
36
37 heterogeneity across agents (Fang et al., 2010). Models of imitative/competitive learning
38
39 make the structure of the task environment a key feature of the models, and so lend
40
41 themselves to empirical testing by examining the heterogeneity of agent's stable sets of
42
43 choices as a function of the interdependence in their task environments (Csaszar and
44
45 Siggelkow, 2010).
46
47
48

49
50 In models of information aggregation, the key aspects of the task environment pertain to the
51
52 rate of arrival and value of projects to be evaluated. Organization design theories have built
53
54 extensive typologies of environmental conditions such as turbulence and munificence that
55
56 capture such variations (e.g., Aldrich, 1979; Dess and Beard, 1984; McCarthy et al. 2010).
57
58 These typologies and the instruments used to measure variance on these dimensions can be
59
60

1
2
3 used to empirically verify model predictions about how rate of arrival of projects and the
4 value of projects to be evaluated lead to particular decisions of organizations. Also, the
5 aggregation structures that combine individual evaluations (the key element of these models)
6
7 are related to the distribution of decision rights in organizations- as visible in authority
8
9 hierarchies and committees. Measuring differences in authority hierarchies and committees
10
11 and combining them with differences in environmental munificence and turbulence can be
12
13 and indeed has been one method used to test predictions derived from information
14
15 aggregation models (c.f., Csaszar, 2012; Reitzig and Maciejovsky, 2014).
16
17
18
19
20

21
22 An alternative to looking for naturally occurring data is to set up experiments in the
23 behavioral laboratory⁹ to test predictions. As we noted at the outset of the review, there is a
24 natural complementarity between modeling and experimental work in that they both deal
25
26 with interaction structures among a small number of agents, in carefully controlled task
27
28 environments (Billinger et al, 2014; Reitzig and Maciejovsky, 2014). Experiments thus
29
30 provide a very direct test of the predictions emerging from models, as indeed has been
31
32 discovered to their great advantage by researchers in behavioral game theory (Camerer,
33
34 2003).
35
36
37
38
39
40

41 To make our idea about complementarity concrete, we consider some possible
42
43 experimental paradigms that may be useful starting points (if not destinations) on such a
44
45 journey.
46
47
48
49

50 As we have outlined, the m-armed bandit and the NK landscape are two basic model
51
52 families that have been used to understand adaptation by unitary decision makers.
53
54

55
56 ⁹ Field experiments would be even stronger if feasible to set up in a way to allow for precise tests of
57 mechanisms (e.g. Bloom et al. 2013).
58
59
60

1
2
3 Fortunately experimental tasks exist that correspond quite closely to the task environments
4 of these models. The IOWA gambling task (Bechara et al., 1997) is one such example.
5
6
7

8
9
10 Bandit tasks and the IOWA gambling task share similar structural properties. Both present
11 decision-makers with a fixed number of choice alternatives, and in both each choice
12 alternative have a fixed rate of reward which is unknown to the decision maker. From trial
13 to trial the decision-maker receives information (outcome feedback and/or reward
14 information) from their choice between the alternatives, and their job is to reliably select
15 sequentially from the alternatives so that they maximize their rewards. The appeal of bandit
16 tasks, and by extension variants of the IOWA gambling task, is that they lend themselves
17 very well to examining both psychological phenomena (such as risk taking, impulsivity,
18 and exploration behavior) and modelling including a determination of what is optimal
19 behavior in different choice environments. This laboratory task has many features which
20 can be carefully manipulated (e.g., payoff structure, feedback, framing, training length,
21 cost of information search, and time pressures) that allow researchers to understand the
22 types of strategies that individuals develop while exploring and then exploiting their
23 environment.
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40

41 As outlined above, the NK model has emerged as a primary modelling approach to study
42 complex management problems such as new product development (e.g. Almirall and
43 Casadesus-Masanell, 2010; Claussen et al., 2014), organization design (e.g. Rivkin and
44 Siggelkow, 2003), and strategy making (e.g. Gavetti et al., 2005). It may also be a useful
45 experimental platform to study search, learning, and decision-making in complex task
46 environments. Busemeyer et al. (1986) used a simple resource allocation task in which
47 participants have to divide up a fixed number of resources among distinct alternatives
48 (such as allocating time to research and teaching). This task can be made more or less
49 complex by allowing for payoff interactions among the alternatives. The resource
50
51
52
53
54
55
56
57
58
59
60

1
2
3 allocation task thereby shares some common features with the NK model, because
4
5 experimental researchers can vary the number of alternatives (equivalent to the N
6
7 parameter in NK models) and the presence of interaction (K parameter) and can be used to
8
9 study predictions from NK models. Billinger et al. (2014) proposed an experimental
10
11 product design task building directly on the NK model. Participants, in their experimental
12
13 setup, must design a product by (re-) combining various attributes and then receive
14
15 performance feedback about specific combinations. The task allows for the easy
16
17 manipulation of, for example, the number of design elements (the N parameter in the NK
18
19 model), interactions among them (the K parameter), different reward structures, or the
20
21 information available to participants (e.g. prior knowledge about the task environment such
22
23 as specific interactions among elements or the payoff of the global optimum, information
24
25 about the designs and current performance of other participants etc.). These experimental
26
27 tasks may allow for the detailed study of how human decision-makers grapple with
28
29 complex problems in various settings such as new product development, process
30
31 improvements, or the discovery and modification of new organization designs or business
32
33 models. Likewise, getting a better understanding how individuals and groups search, learn,
34
35 and choose in complex situations might lead to a refinement of existing formal models and
36
37 theories.
38
39
40
41

42 Experimental tests for the aggregation models can take their inspiration from research that
43
44 uses games to investigate how groups differ from individuals as decision making units. There
45
46 are many elegant demonstrations of differences between groups and individuals using
47
48 classical economic games such as the prisoner's dilemma, dictator, ultimatum, trust, centipede
49
50 and principal-agent games (Kugler et al, 2012; Wildschut et al, 2003). In the prisoner's
51
52 dilemma game (PDG) two players choose between a cooperative and a non-cooperative
53
54 option, and the combination of choices both players make have different payoffs. In adapted
55
56 versions a unitary decision to cooperate or defect is made by a group, which can be compared
57
58
59
60

1
2
3 with decisions made by individuals. Wildschut et al (2003) and others since have shown that
4 groups are more competitive than individuals in the PDG. Ways of amplifying and
5
6 attenuating this discontinuity effect have been connected to key factors associated with group
7
8 dynamics (e.g., strong conflicts of interest, distributed responsibility of decisions taken,
9
10 relaxed attitudes to violations of norms of fairness, level of communication). Moreover,
11
12 when looking across studies comparing groups with individuals in economics tasks the
13
14 emerging pattern seems to be that groups make unitary decisions that are closer to rational in
15
16 game theoretical terms (Kugler et al, 2012; Maciejovsky et al., 2013). Using tasks such as the
17
18 PDG to look at aggregation issues is extremely valuable, because it is a simple starting point
19
20 from which to examine different aggregation structures, such as consensus, voting, or
21
22 hierarchy in order to compare how groups using these rules act, and if there are any
23
24 predictable differences to how individuals do so. Another approach is to use the group versus
25
26 individual method used in the PDG with experiments that use paradigms developed to test
27
28 behaviors in NK environments. Some emerging research in this area examines how groups
29
30 differ from individuals on local versus distant search. Using the experimental NK task,
31
32 Billinger et al. (2015) examined how two-person teams search for performance improvements
33
34 in rugged landscapes. They find that teams search less locally and perform better than
35
36 individuals in identical tasks.
37
38
39
40
41

42 A second, and as yet largely uncharted area of exploration is to construct task
43
44 environments that allow the study of mutual adjustment processes. Experimental tasks in
45
46 which pairs (or more numbers) of agents take interdependent actions and adjust to
47
48 feedback are of course staple in social psychology and behavioral game theory (e.g., Roth
49
50 and Murnighan, 1978; Murnighan and Roth, 1983); however tasks that map closely to
51
52 bandit or NK models are yet to be constructed. In sum, we can think of the space of
53
54 opportunities here along these two dimensions: 1) Individual vs. multi-agent decision
55
56 making entity and 2) Adjustment to environment vs. adjustment to another decision
57
58
59
60

1
2
3 making entity (i.e., mutual adjustment). The quadrant that involves groups interacting with
4
5 groups is sparse at this point, but is perhaps the closest to inter-organizational phenomena
6
7 of interest such as competition and collaboration.
8
9

10
11 A challenge that organizational scientists may face in relying on experiments is the skepticism
12
13 that readers and reviewers may display about whether interactions between small numbers of
14
15 agents (often two) may have anything to say about “real” organizations, as opposed to say
16
17 “merely” teams, dyads or groups. (This is in addition to the usual concerns about external
18
19 validity that are justifiably flagged with any experimental work). A thorough consideration of
20
21 the issue is more than we can provide in this paper, but we will note the following: a review
22
23 of the most influential definitions of organization reveals certain commonly occurring criteria
24
25 for an entity to be considered an organization, such as the existence of goals whose attainment
26
27 depends on the efforts of multiple agents, but the number of agents (beyond the “more than
28
29 one” requirement) is never among these criteria (Puranam, Alexy, Reitzig, 2014). In our view,
30
31 there is no basis (besides convention) on which one can say that a three-person firm is an
32
33 organization but a four-person team is not; to the extent these are goal directed multi-agent
34
35 systems, they are both organizations, albeit solving their problems of division of labor and
36
37 integration of effort in different ways (also see Raveendran, Puranam and Warglien,
38
39 Forthcoming).
40
41
42

43
44 There are of course important and completely legitimate questions to be addressed on whether
45
46 the manner in which a three-person organization works is the same as an n-person
47
48 organization (scaling), or whether the behavior of the system changes when each agent is
49
50 itself an organization (recursion). However, we believe we can make a start at answering
51
52 them by taking the first steps into the lab, possibly guided by intuitions derived from models.
53
54 The prospect for fruitful collaborations between experimental researchers and modelers in
55
56 advancing organization science seems bright.
57
58
59
60

5. Conclusions

We began the paper by suggesting that the literature on adaptively rational models of organizations based on Simon's (1916-2001) pioneering work on bounded rationality was deserving of critical examination on account of its central claim that it is based on 'correct' behavioral assumptions. We also suggested that the extensive research in this tradition has generated many precise insights that could be empirically tested. To facilitate this examination and provide organizational scholars with a tractable typology of models we classified and examined model families. Specifically, we reviewed three important families of models of adaptive rationality in organizations- individual adaptation, mutual adaptation, and information aggregation. These do not exhaust the model families prevalent in organizations research, but represent conceptually important and influential ones. And finally, we discussed opportunities for research that strengthen the links between formal modelling in organizations research, and its behavioral foundations.

Our review suggests that the adaptively rational models are based on empirically validated behavioral assumptions in the main. We have identified a few assumptions that bear further testing. We have also identified the key insights and predictions generated by different models and provided information about the empirical tests (albeit limited) of these predictions. Finally, we discussed opportunities for future research and outlined how empirical tests can be constructed to take advantage of these opportunities.

Our hope is that this review will facilitate greater dialogue between modelers and empiricists and that organizational theories and empirical studies will attain greater precision and provide better predictions as a result.

References

(Note: Contributions are coded for model families for easy reference: BA = Multi-armed Bandit Model, NK = NK model; MA = Mutual Adjustment Model; IA = Information Aggregation model)

Aldrich, H. (1979). *Organizations and environments*. Stanford University Press.

Adner, R., Polos, L., Ryall, M., & Sorenson, O. (2009). The case for formal theory. *Academy of Management Review*, 34(2), 201-208.

Ariely, D., Loewenstein, G., & Prelec, D (2003). Coherent Arbitrariness: Stable Demand Curves Without Stable Preferences. *Quarterly Journal of Economics*, 118(1), 73-105. (NK)

Argote, L. (2012). *Organizational learning: Creating, retaining and transferring knowledge*. Springer.

Arthur, W. B. (1991). Designing economic agents that act like human agents: A behavioral approach to bounded rationality. *The American Economic Review*, 81(2), 353-359. (BA)

Almirall, E., & Casadesus-Masanell, R. (2010). Open versus closed innovation: A model of discovery and divergence. *Academy of Management Review*, 35(1), 27-47. (NK)

Badre, D., Doll, B. B., Long, N. M., & Frank, M. J. (2012). Rostrolateral prefrontal cortex and individual differences in uncertainty-driven exploration. *Neuron*, 73(3), 595-607. (BA)

Bandura, A. (1977). *Social learning theory*. (MA)

Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (1997). Deciding advantageously before knowing the advantageous strategy. *Science*, 275(5304), 1293-1295. (BA)

Billinger, S., Stieglitz, N., & Schumacher, T. R. (2014). Search on rugged landscapes:: an experimental study. *Organization Science*, 25(1), 93-108. (NK)

Billinger, S., Kretschmer, T., Schumacher, T.R., Stieglitz, N. (2015): Searching in teams: An experimental study of aggregation and incentives. Working Paper. (NK)

Blanco, F., Matute, H., & Vadillo, M. (2011). Making the uncontrollable seem controllable: The role of action in the illusion of control. *The Quarterly Journal of Experimental Psychology*, 64(7), 1290-1304.

Bloom, N., Eifert, B., Mahajan, A., McKenzie, D., & Roberts, J. (2013). Does Management Matter? Evidence from India. *The Quarterly Journal of Economics*, 128(1), 1-51.

Blount, S. (1995). When social outcomes aren't fair: The effect of causal attributions on preferences. *Organizational behavior and human decision processes*, 63(2), 131-144. (NK)

- 1
2
3 Blume, A., Duffy, J., & Franco, A. M. (2009). Decentralized organizational learning: an
4 experimental investigation. *The American Economic Review*, 1178-1205. (MA)
5
6 Bohnet, I., & Zeckhauser, R. (2004). Trust, risk and betrayal. *Journal of Economic Behavior
7 & Organization*, 55(4), 467-484. (MA)
8
9 Busemeyer, J. R., Swenson, K. N., & Lazarte, A. (1986). An adaptive approach to resource
10 allocation. *Organizational Behavior and Human Decision Processes*, 38(3), 318-341.
11 (NK)
12
13 Busemeyer, J. R., & Myung, I. J. (1987). Resource allocation decision making in an uncertain
14 environment. *Acta Psychologica*, 66(1), 1-19. (NK)
15
16 Bush, R. R., & Mosteller, F. (1955). *Stochastic models for learning*. (BA)
17
18 Brand, H., Sakoda, J. M., & Woods, P. J. (1957). Effects of a random versus pattern
19 reinforcement instructional set in a contingent partial reinforcement situation.
20 *Psychological Reports*, 3(3), 473-479. (BA)
21
22 Brand, H., Wood, P. J., & Sakoda, J. M. (1956). Anticipation of reward as a function of
23 partial reinforcement. *Journal of Experimental Psychology*, 52(1), 18-22. (BA)
24
25 Brown, S. L., & Eisenhardt, K. M. (1997). The art of continuous change: Linking complexity
26 theory and time-paced evolution in relentlessly shifting organizations. *Administrative
27 Science Quarterly*, 42(1), 1-34.
28
29 Camerer, C. F. (2003). *Behavioral game theory: Experiments in strategic interaction*.
30 Princeton University Press.
31
32 Camerer, C.F., & Ho, T. (1999). Experience-weighted Attraction Learning in Normal Form
33 Games. *Econometrica*, 67(4), 827-874. (BA)
34
35 Carley, K. (1992). Organizational learning and personnel turnover. *Organization Science*,
36 3(1), 20-46.
37
38 Christensen, M., & Knudsen, T. (2010). Design of decision-making organizations.
39 *Management Science*, 56(1), 71-89. (IA)
40
41 Christensen, M. & Knudsen, T. (2013). Confidence and Joint Learning, Working paper,
42 University of Southern Denmark.
43
44 Claussen, J., Kretschmer, T., & Stieglitz, N. (2014). Vertical Scope, Turbulence, and the
45 Benefits of Commitment and Flexibility. Forthcoming in *Management Science*.
46
47 Cohen, M. D., March, J. G., & Olsen, J. P. (1972). A garbage can model of organizational
48 choice. *Administrative Science Quarterly*, 17(1), 1-25.
49
50 Csaszar, F. A. (2012). Organizational structure as a determinant of performance: Evidence
51 from mutual funds. *Strategic Management Journal*, 33(6), 611-632. (IA)
52
53 Csaszar, F. A. (2013). An efficient frontier in organization design: Organizational structure as
54 a determinant of exploration and exploitation. *Organization Science*, 24(4), 1083-1101.
55 (IA)
56
57
58
59
60

- 1
2
3 Csaszar, F. A., & Eggers, J. P. (2013). Organizational decision making: An information
4 aggregation view. *Management Science*, 59(10), 2257-2277. (IA)
5
6 Csaszar, F. A., & Siggelkow, N. (2010). How much to copy? Determinants of effective
7 imitation breadth. *Organization Science*, 21(3), 661-676. (NK, MA)
8
9 Cyert, R. M., & March, J. G. (1963). *A behavioral theory of the firm*. Englewood Cliffs, NJ.
10
11 Davis, J. P., Eisenhardt, K. M., & Bingham, C. B. (2007). Developing theory through
12 simulation methods. *Academy of Management Review*, 32(2), 480-499.
13
14 Davis, J. P., Eisenhardt, K. M., & Bingham, C. B. (2009). Optimal structure, market
15 dynamism, and the strategy of simple rules. *Administrative Science Quarterly*, 54(3),
16 413-452.
17
18 Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B., & Dolan, R. J. (2006). Cortical
19 substrates for exploratory decisions in humans. *Nature*, 441(7095), 876-879. (BA)
20
21 Denrell, J. (2007). Adaptive learning and risk taking. *Psychological Review*, 114(1), 177.
22 (BA)
23
24 Denrell, J., Fang, C., & Levinthal, D. A. (2004). From T-mazes to labyrinths: Learning from
25 model-based feedback. *Management Science*, 50(10), 1366-1378. (BA)
26
27 Denrell, J., Fang, C., & Winter, S. G. (2003). The economics of strategic opportunity.
28 *Strategic Management Journal*, 24(10), 977-990.
29
30 Denrell, J., & March, J. G. (2001). Adaptation as information restriction: The hot stove effect.
31 *Organization Science*, 12(5), 523-538. (BA)
32
33 Dess, G. G., & Beard, D. W. (1984). Dimensions of organizational task environments.
34 *Administrative Science Quarterly*, 52-73.
35
36 Domjan, M. (2010) *The Principles of Learning and Behavior*. 6th edition, Stamford, CT.
37
38 Edmondson, A. C., & McManus, S. E. (2007). Methodological fit in management field
39 research. *Academy of management Review*, 32(4), 1246-1264.
40
41 Erev, I., & Roth, A. E. (1998). Predicting how people play games: Reinforcement learning in
42 experimental games with unique, mixed strategy equilibria. *American Economic Review*,
43 88(4), 848-881. (BA)
44
45 Ethiraj, S. K., & Levinthal, D.A. (2004). Bounded rationality and the search for
46 organizational architecture: An evolutionary perspective on the design of organizations
47 and their evolvability. *Administrative Science Quarterly*, 49(3), 404-437. (NK)
48
49 Fang, C. (2012). Organizational learning as credit assignment: A model and two experiments.
50 *Organization Science*, 23(6), 1717-1732. (BA)
51
52 Fang, C., Lee, J., & Schilling, M. A. (2010). Balancing exploration and exploitation through
53 structural design: The isolation of subgroups and organizational learning. *Organization*
54 *Science*, 21(3), 625-642. (MA)
55
56
57
58
59
60

- 1
2
3 Fang, C., & Levinthal, D. (2009). Near-term liability of exploitation: exploration and
4 exploitation in multistage problems. *Organization Science*, 20(3), 538-551. (BA)
5
6 Fernie, G., & Tunney, R. J. (2006). Some decks are better than others: The effect of reinforcer
7 type and task instructions on learning in the Iowa Gambling Task. *Brain and Cognition*,
8 60(1), 94-102. (BA)
9
10 Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, 7(2), 117-
11 140. (MA)
12
13 Fleming, L. (2001). Recombinant uncertainty in technological search. *Management Science*,
14 47(1), 117-132. (NK)
15
16 Fleming, L., & Sorenson, O. (2001). Technology as a complex adaptive system: evidence
17 from patent data. *Research Policy*, 30(7), 1019-1039. (NK)
18
19 Fleming, L., & Sorenson, O. (2004). Science as a map in technological search. *Strategic*
20 *Management Journal*, 25(8-9), 909-928. (NK)
21
22 Fridberg, D. J., Queller, S., Ahn, W.-Y., Kim, W., Bishara, A. J., Busemeyer, J. R. Stout, J. C.
23 (2010). Cognitive mechanisms underlying risky decision-making in chronic cannabis
24 users. *Journal of Mathematical Psychology*, 54(1), 28-38. (BA)
25
26 Gavetti, G., & Levinthal, D. (2000). Looking forward and looking backward: Cognitive and
27 experiential search. *Administrative science Quarterly*, 45(1), 113-137. (NK)
28
29 Gavetti, G., Levinthal, D. A., & Rivkin, J. W. (2005). Strategy making in novel and complex
30 worlds: the power of analogy. *Strategic Management Journal*, 26(8), 691-712. (NK)
31
32 Gavetti, G., Levinthal, D., & Ocasio, W. (2007). The Carnegie School's Past, Present, and
33 Reconstructing for the Future. *Organization Science*, 18(3), 523-536.
34
35 Gavetti, G. (2012). Toward a behavioral theory of strategy. *Organization Science*, 23(1), 267-
36 285.
37
38 Gavetti, G., Greve, H. R., Levinthal, D. A., & Ocasio, W. (2012). The behavioral theory of
39 the firm: Assessment and prospects. *The Academy of Management Annals*, 6(1), 1-40.
40
41 Ghemawat, P., & Levinthal, D. (2008). Choice interactions and business strategy.
42 *Management Science*, 54(9), 1638-1651. (NK)
43
44 Gilboa, I., & Schmeidler, D. (2001). A theory of case-based decisions. Cambridge University
45 Press.
46
47 Gittins, J. C. (1979). Bandit processes and dynamic allocation indices. *Journal of the Royal*
48 *Statistical Society. Series B (Methodological)*, 41(2), 148-177. (BA)
49
50 Gittins, J.C., & Jones, D.M. (1974). A dynamic allocation index for the sequential design of
51 experiments. Gan, J., Saikad, K., & Vince, L. (eds), *Progress in Statistics*, North Holland,
52 Amsterdam, 241-246. (BA)
53
54 Greve, H. R. (1998). Performance, Aspirations, and Risky Organizational Change.
55 *Administrative Science Quarterly*, 43(1), 58-86.
56
57
58
59
60

- 1
2
3 Greve, H. R. (2003). Organizational learning from performance feedback: A behavioral
4 perspective on innovation and change. Cambridge University Press.
5
- 6 Gruber, M., MacMillan, I. C., & Thompson, J. D. (2008). Look before you leap: Market
7 opportunity identification in emerging technology firms. *Management Science*, 54(9),
8 1652-1665.
9
- 10 Gulati, R., & Puranam, P. (2009). Renewal through reorganization: The value of
11 inconsistencies between formal and informal organization. *Organization Science*, 20(2),
12 422-440.
13
- 14 Harrison, J. R., Lin, Z., Carroll, G. R., & Carley, K. M. (2007). Simulation modeling in
15 organizational and management research. *Academy of Management Review*, 32(4), 1229-
16 1245.
17
- 18 Hoeffler, S., Ariely, D., & West, P. (2006). Path dependent preferences: The role of early
19 experience and biased search in preference development. *Organizational Behavior and*
20 *Human Decision Processes*, 101(2), 215-229. (NK)
21
- 22 Holland, J. H. (1975). *Adaptation in natural and artificial systems: An introductory analysis*
23 *with applications to biology, control, and artificial intelligence*. U Michigan Press.
24
- 25 Hu, S., Blettner, D., & Bettis, R. A. (2011). Adaptive aspirations: performance consequences
26 of risk preferences at extremes and alternative reference groups. *Strategic Management*
27 *Journal*, 32(13), 1426-1436.
28
- 29 Hug, S. (1999). Nonunitary Actors in Spatial Models How Far Is Far in Foreign Policy?.
30 *Journal of Conflict Resolution*, 43(4), 479-500.
31
- 32 Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk.
33 *Econometrica*, 47(2), 263-291.
34
- 35 Kahneman, D., & Klein, G. (2009). Conditions for intuitive expertise: a failure to disagree.
36 *American Psychologist*, 64(6), 515.
37
- 38 Katila, R., & Ahuja, G. (2002). Something old, something new: A longitudinal study of
39 search behavior and new product introduction. *Academy of management journal*, 45(6),
40 1183-1194.
41
- 42 Kauffman, S. A. (1993). *The Origins of Order: Self-Organization and Selection in Evolution*.
43 Oxford University Press. (NK)
44
- 45 Kauffman, S. A., & Weinberger, E. D. (1989). The $\langle i \rangle$ NK model of rugged fitness
46 landscapes and its application to maturation of the immune response. *Journal of*
47 *Theoretical Biology*, 141(2), 211-245. (NK)
48
- 49 Klayman, J., & Ha, Y. W. (1987). Confirmation, disconfirmation, and information in
50 hypothesis testing. *Psychological review*, 94(2), 211. (NK)
51
- 52 Kjome, K. L., Lane, S. D., Schmitz, J. M., Green, C., Ma, L., Prasla, I., .Moeller, F. G.
53 (2010). Relationship between impulsivity and decision making in cocaine dependence.
54 *Psychiatry Research*, 178(2), 299-304. (BA)
55
56
57
58
59
60

- 1
2
3 Knudsen, T., & Levinthal, D. A. (2007). Two faces of search: Alternative generation and
4 alternative evaluation. *Organization Science*, 18(1), 39-54. (NK, IA)
5
6 Knudsen, T., & Srikanth, K. (2014). Coordinated Exploration Organizing Joint Search by
7 Multiple Specialists to Overcome Mutual Confusion and Joint Myopia. Forthcoming in
8 *Administrative Science Quarterly*. (MA)
9
10 Kotovsky, K., Hayes, J. R., & Simon, H. A. (1985). Why are some problems hard? Evidence
11 from Tower of Hanoi. *Cognitive Psychology*, 17(2), 248-294.
12
13 Kretschmer, T., & Puranam, P. (2008). Integration through incentives within differentiated
14 organizations. *Organization Science*, 19(6), 860-875.
15
16 Kugler, T., Kausel, E. E., & Kocher, M. G. (2012). Are groups more rational than
17 individuals? A review of interactive decision making in groups. *Wiley Interdisciplinary*
18 *Reviews: Cognitive Science*, 3(4), 471-482.
19
20 Langley, P., Pearce, C., Barley, M., & Emery, M. (2014). Bounded rationality in problem
21 solving: Guiding search with domain-independent heuristics. *Mind & Society*, 13(1), 83-
22 95.
23
24 Lant, T. K. (1992). Aspiration level adaptation: An empirical exploration. *Management*
25 *Science*, 38(5), 623-644.
26
27 Lant, T. K., & Mezas, S. J. (1992). An organizational learning model of convergence and
28 reorientation. *Organization Science*, 3(1), 47-71.
29
30 Laureiro-Martínez, D., Brusoni, S., Canessa, N., & Zollo, M. (2014). Understanding the
31 exploration–exploitation dilemma: An MRI study of attention control and decision-
32 making performance. Forthcoming in *Strategic Management Journal*. (BA)
33
34 Lave, C. A., & J. G. March. 1975. *An Introduction to Models in the Social Sciences*. Harper
35 & Row, New York. (BA, MA)
36
37 Lazer, D., & Friedman, A. (2007). The network structure of exploration and exploitation.
38 *Administrative Science Quarterly*, 52(4), 667-694. (NK, MA)
39
40 Lea, S. E., McLaren, I. P., Dow, S. M., & Graft, D. A. (2012). The cognitive mechanisms of
41 optimal sampling. *Behavioral Processes*, 89(2), 77-85. (BA)
42
43 Lee, E., Puranam, P. (Forthcoming). The implementation imperative: Why one should
44 implement even imperfect strategies perfectly, *Strategic Management Journal*. (BA,
45 MA)
46 Lenox, M. J., Rockart, S. F., & Lewin, A. Y. (2006). Interdependency, competition,
47 and the distribution of firm and industry profits. *Management Science*, 52(5), 757-772.
48 (NK, MA)
49
50 Lenox, M. J., Rockart, S. F., & Lewin, A. Y. (2007). Interdependency, competition, and
51 industry dynamics. *Management Science*, 53(4), 599-615. (NK, MA)
52
53 Lenox, M. J., Rockart, S. F., & Lewin, A. Y. (2010). Does interdependency affect firm and
54 industry profitability? An empirical test. *Strategic Management Journal*, 31(2), 121-139.
55 (NK, MA)
56
57
58
59
60

- 1
2
3 Levinthal, D. A. (1997). Adaptation on rugged landscapes. *Management Science*, 43(7), 934-
4 950. (NK)
5
6 Levinthal, D., & Posen, H. E. (2007). Myopia of selection: Does organizational adaptation
7 limit the efficacy of population selection?. *Administrative Science Quarterly*, 52(4), 586-
8 620. (NK)
9
10 Levitt, B., & March, J. G. (1988). Organizational learning. *Annual Review of Sociology*, 319-
11 340.
12
13 Lippman, S. A., & Rumelt, R. P. (2003). The payments perspective: micro-foundations of
14 resource analysis. *Strategic Management Journal*, 24(10), 903-927.
15
16 Loch, C. H., & Kavadias, S. (2002). Dynamic portfolio selection of NPD programs using
17 marginal returns. *Management Science*, 48(10), 1227-1241. (BA)
18
19 Lounamaa, P. H., & March, J. G. (1987). Adaptive coordination of a learning team.
20 *Management Science*, 33(1), 107-123. (MA)
21
22 Luce, D. R. (1959). *Individual Choice Behavior*. Harper & Row, New York.
23
24 Maciejovsky, B., Sutter, M., Budescu, D. V., & Bernau, P. (2013). Teams make you smarter:
25 How exposure to teams improves individual decisions in probability and reasoning tasks.
26 *Management Science*, 59(6), 1255-1270.
27
28 March, J. G. (1991). Exploration and exploitation in organizational learning. *Organization*
29 *Science*, 2(1), 71-87. (MA)
30
31 March, J. G. (1996). Learning to be risk averse. *Psychological Review*, 103(2), 309. (BA)
32
33 McCarthy, I. P., Lawrence, T. B., Wixted, B., & Gordon, B. R. (2010). A multidimensional
34 conceptualization of environmental velocity. *Academy of Management Review*, 35(4),
35 604-626.
36
37 Miller, K. D., Zhao, M., & Calantone, R. J. (2006). Adding interpersonal learning and tacit
38 knowledge to March's exploration-exploitation model. *Academy of Management Journal*,
39 49(4), 709-722. (MA)
40
41 Murnighan, J. K., & Roth, A. E. (1983). Expecting Continued Play in Prisoner's Dilemma
42 Games A Test of Several Models. *Journal of Conflict Resolution*, 27(2), 279-300.
43
44 Minniti, M., & Bygrave, W. (2001). A dynamic model of entrepreneurial learning.
45 *Entrepreneurship Theory and Practice*, 25(3), 5-16. (BA)
46
47 Nelson, R. R., & Winter, S. G. (1982). *An evolutionary theory of economic change*. Belknap:
48 Cambridge, Ma.
49
50 Newell, B. R., & Shanks, D. R. (2014). Unconscious influences on decision making: A
51 critical review. *Behavioral and Brain Sciences*, 37(1), 1-19.
52
53 Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-
54 Hall.
55
56
57
58
59
60

- 1
2
3 Noda, T., & Bower, J. L. (1996). Strategy making as iterated processes of resource allocation.
4 Strategic Management Journal, 17(S1), 159-192. (BA)
5
6 Nowak, M., & Sigmund, K. (1993). A strategy of win-stay, lose-shift that outperforms tit-for-
7 tat in the Prisoner's Dilemma game. Nature, 364(6432), 56-58. (BA)
8
9 Otto, A. R., Markman, A. B., Gureckis, T. M., & Love, B. C. (2010). Regulatory Fit and
10 Systematic Exploration in a Dynamic Decision-Making Environment. Learning, Memory,
11 36(3), 797-804. (BA)
12
13 O'Reilly, C., & Tushman, M. (2013). Organizational ambidexterity: Past, present and future.
14 The Academy of Management Perspectives, 27(4), 324-338.
15
16 Osman, M. (2010). Controlling uncertainty: a review of human behavior in complex dynamic
17 environments. Psychological Bulletin, 136(1), 65-86.
18
19 Osman, M.(2014). Future-Minded: The Psychology of Agency and Control. Basingstoke:
20 Palgrave-MacMillan.
21
22 Osman, M., & Ananiadis-Basias, A. (2013). Context and animacy play a role in dynamic
23 decision- making. Journal of Entrepreneurship, Management and Innovation. (2), 61-78.
24
25 Page, S. E. (1996). Two measures of difficulty. Economic Theory, 8(2), 321-346. (NK)
26
27 Payzan-LeNestour, E., & Bossaerts, P. (2012). Do not bet on the unknown versus try to find
28 out more: estimation uncertainty and “unexpected uncertainty” both modulate
29 exploration. Frontiers in Neuroscience, 6(150). (BA)
30
31 Posen, H. E., Lee, J., & Yi, S. (2013). The power of imperfect imitation. Strategic
32 Management Journal, 34(2), 149-164. (MA)
33
34 Posen, H. E., & Levinthal, D. A. (2012). Chasing a moving target: Exploitation and
35 exploration in dynamic environments. Management Science, 58(3), 587-601. (BA)
36
37 Premkumar, P., Fannon, D., Kuipers, E., Simmons, A., Frangou, S., & Kumari, V. (2008).
38 Emotional decision-making and its dissociable components in schizophrenia and
39 schizoaffective disorder: A behavioral and MRI investigation. Neuropsychologia, 46(7),
40 2002–2012. (BA)
41
42 Prietula, M., Carley, K., & Gasser, L. (1998). Simulating organizations: Computational
43 models of institutions and groups. The MIT Press.
44
45 Puranam, P., & Swamy, M. (Forthcoming)). How initial representations shape coupled
46 learning processes. Organization Science. (BA, MA)
47
48 Puranam, P., Alexy, O., & Reitzig, M. (2014). What's “New” About New Forms of
49 Organizing?. Academy of Management Review, 39(2), 162-180.
50
51 Puranam, P., & Vanneste, B. S. (2009). Trust and governance: Untangling a tangled web.
52 Academy of Management Review, 34(1), 11-31.
53
54 Rahmandad, H. (2008). Effect of delays on complexity of organizational learning.
55 Management Science, 54(7), 1297-1312. (BA)
56
57
58
59
60

- 1
2
3 Raisch, S., Birkinshaw, J., Probst, G., & Tushman, M. L. (2009). Organizational
4 ambidexterity: Balancing exploitation and exploration for sustained performance.
5 *Organization Science*, 20(4), 685-695.
6
- 7 Raveendran, M., Puranam, P., & Warglien, M (Forthcoming), Object salience in division of
8 labor: experimental evidence, *Management Science*
9
- 10 Reitzig, M., & Maciejovsky, B. (2014). Corporate hierarchy and vertical information flow
11 inside the firm a behavioral view. Forthcoming in *Strategic Management Journal*. (IA)
12
- 13 Rivkin, J. W. (2000). Imitation of complex strategies. *Management Science*, 46(6), 824-844.
14 (NK)
15
- 16 Rivkin, J. W. (2001). Reproducing knowledge: Replication without imitation at moderate
17 complexity. *Organization Science*, 12(3), 274-293. (NK)
18
- 19 Rivkin, J. W., & Siggelkow, N. (2003). Balancing search and stability: Interdependencies
20 among elements of organizational design. *Management Science*, 49(3), 290-311. (NK,
21 MA)
22
- 23 Rivkin, J. W., & Siggelkow, N. (2007). Patterned interactions in complex systems:
24 Implications for exploration. *Management Science*, 53(7), 1068-1085.(NK, MA)
25
- 26 Rosenkopf, L., & Nerkar, A. (2001). Beyond local search: boundary-spanning, exploration,
27 and impact in the optical disk industry. *Strategic Management Journal*, 22(4), 287-306.
28
- 29 Roth, A. E., & Murnighan, J. K. (1978). Equilibrium behavior and repeated play of the
30 prisoner's dilemma. *Journal of Mathematical Psychology*, 17(2), 189-198.
31
- 32 Rubinstein, A. (1998). *Modeling bounded rationality*. MIT Press.
33
- 34 Sah, R. K., & Stiglitz, J. E. (1985). Human fallibility and economic organization. *The*
35 *American Economic Review*, 75(2), 292-297. (IA)
36
- 37 Sah, R. K., & Stiglitz, J. E. (1988). Committees, hierarchies and polyarchies. *Economic*
38 *Journal*, 98(391), 451-470. (IA)
39
- 40 Schilling, M. A., & Fang, C. (2013). When hubs forget, lie, and play favorites: Interpersonal
41 network structure, information distortion, and organizational learning. Forthcoming in
42 *Strategic Management Journal*. (MA)
43
- 44 Schoemaker, P. J. (1982). The expected utility model: Its variants, purposes, evidence and
45 limitations. *Journal of Economic Literature*, 20(2), 529-563.
46
- 47 Selten, R. (2001). What is bounded rationality? In: Gigerenezzer, G., & Selten, R. (eds):
48 *Bounded Rationality: The Adaptive Toolbox*, Cambridge, MA: MIT Press, 13-36).
49
- 50 Siggelkow, N. (2002). Misperceiving interactions among complements and substitutes:
51 Organizational consequences. *Management Science*, 48(7), 900-916.
52
- 53 Siggelkow, N., & Rivkin, J. W. (2005). Speed and search: Designing organizations for
54 turbulence and complexity. *Organization Science*, 16(2), 101-122. (NK, MA)
55
- 56
57
58
59
60

- 1
2
3 Siggelkow, N., & Rivkin, J. W. (2006). When exploration backfires: Unintended
4 consequences of multilevel organizational search. *Academy of Management Journal*,
5 49(4), 779-795. (NK, MA)
6
- 7 Siggelkow, N., & Rivkin, J. W. (2009). Hiding the evidence of valid theories: how coupled
8 search processes obscure performance differences among organizations. *Administrative*
9 *Science Quarterly*, 54(4), 602-634. (NK, MA)
10
- 11 Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of*
12 *Economics*, 69(1), 99-118.
13
- 14 Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological*
15 *Review*, 63(2), 129.
16
- 17 Simon, H. A. (1979). Rational decision making in business organizations. *The American*
18 *Economic Review*, 493-513.
19
- 20 Simon, H. A. (1990). Invariants of human behavior. *Annual Review of Psychology*, 41(1), 1-
21 20.
22
- 23 Simon H. A. (1996). *The Sciences of the Artificial*. Cambridge, MA: MIT Press. 3rd ed.
24
- 25 Simon, H. A. (2007). *An empirically-based microeconomics*. Cambridge University Press.
26
- 27 Smith, V. L. (1982). Microeconomic systems as an experimental science. *The American*
28 *Economic Review*, 923-955.
29
- 30 Spiegler, R. (2011). *Bounded rationality and industrial organization*. Oxford University Press.
31
- 32 Steingroever, H., Wetzels, R., Horstmann, A., Neumann, J., & Wagenmakers, E. J. (2012).
33 Performance of Healthy Participants on the Iowa Gambling Task. *Psychological*
34 *Assessment*, 1-14. (BA)
35
- 36 Steyvers, M., Lee, M. D., & Wagenmakers, E. (2009). A Bayesian analysis of human
37 decision-making on bandit problems. *Journal of Mathematical Psychology*, 53(3), 168-
38 179. (BA)
39
- 40 Stieglitz, N., Knudsen, T., & Becker, M. Adaptation and inertia in dynamic environments.
41 Working Paper. (BA)
42
- 43 Stuart, T. E., & Podolny, J. M. (1996). Local search and the evolution of technological
44 capabilities. *Strategic Management Journal*, 17(S1), 21-38.
45
- 46 Sutton, R.S., & Barto, A. G. (1998). *Reinforcement learning: An introduction*. MIT Press.
47 (BA)
48
- 49 Surowiecki, J. (2005). *The wisdom of crowds*. Random House LLC.
50
- 51 Thorndike, E. L. (1911). *Animal intelligence: Experimental studies*. Macmillan.
52
- 53 Wildschut, T., Pinter, B., Vevea, J. L., Insko, C. A., & Schopler, J. (2003). Beyond the group
54 mind: a quantitative review of the interindividual-intergroup discontinuity effect.
55 *Psychological Bulletin*, 129(5), 698-715.
56
57
58
59
60

- 1
2
3 Wood, S., Busemeyer, J., Kolling, A., Cox, C. R., & Davis, H. (2005). Older adults as
4 adaptive decision makers: Evidence from the Iowa Gambling Task. *Psychology and*
5 *Aging*, 20(2), 220–225. (BA)
6
- 7 Wright, S. (1932). The Roles of Mutation, Inbreeding, Crossbreeding and Selection in
8 Evolution, *Proceedings of the Sixth Annual Congress of Genetics* 1: 356-366. Reprinted
9 in William B. Provine (1986), *Sewall Wright: Evolution: Selected Papers* : Chicago:
10 University of Chicago Press: 161-177. (NK)
11
- 12 Yechiam, E., & Busemeyer, J. R. (2005). Comparison of basic assumptions embedded in
13 learning models for experience-based decision making. *Psychonomic Bulletin & Review*,
14 12(3), 387-402. (BA)
15
16
- 17 Zollo, M., & Winter, S. G. (2002). Deliberate learning and the evolution of dynamic
18 capabilities. *Organization Science*, 13(3), 339-351.
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table 1: Summary of model families

<i>Model</i>	<i>Task environment</i>	<i>Representation of task environment</i>	<i>Choice process</i>	<i>Transformation</i>	<i>Key trade-offs</i>	<i>Key insights</i>
Models of individual adaptation						
Bandit model	<p>Actions are discrete alternatives with uncertain and unknown outcomes</p> <p><u>Key parameters:</u> Noise and change in payoff (e.g. some investment opportunities produce clear signals of value, others do not)</p>	<p>Subjective beliefs about outcomes (e.g. discriminative between choices, and bias of representation)</p>	<p>Probabilistic choice based on representation</p> <p><u>Key parameter:</u> Propensity to explore by taking actions other than those currently believed to be best (e.g. deviation from past successful actions, novelty seeking)</p>	<p>Reinforcement learning to update representation</p> <p><u>Key parameter:</u> Responsiveness to feedback (e.g. incentives for success and failure)</p>	<p>Exploitation of current representation versus exploration to update representation</p> <p>Learning without being misled by positive and negative feedback</p>	<p>Value of exploration (Sutton & Barto, 1998)</p> <p>Dangers of learning too rapidly from noisy feedback (Denrell, 2005; Posen & Levinthal, 2012)</p> <p>Hot stove effect (Denrell & March, 2001)</p> <p>Risk aversion as an outcome of learning (March, 1996)</p>
NK model	<p>Actions are complex combinatorial alternatives with certain but unknown outcomes</p> <p><u>Key parameter:</u> Interdependence between choice elements (e.g. decomposability of design space or value chain)</p>	<p>Limited representation of available alternatives and of structure of task environment</p> <p><u>Key parameters:</u> Quality of initial representation (e.g. coarseness, discrimination among choices, and bias of representation)</p>	<p>Maximizing choice with potentially imperfect evaluation</p> <p><u>Key parameter:</u> Errors in evaluation (e.g. experience or expertise of evaluating agent)</p>	<p>Recombinant search to create and test new alternatives</p> <p><u>Key parameter:</u> Propensity to change many elements at a time (e.g. uncoordinated search with specialists vs. coordinated search, political deadlocks leading to local search)</p>	<p>Local versus distant search</p> <p>Perfect versus imperfect alternative evaluation</p>	<p>Value of combining local and distant search (Levinthal, 1997; Siggelkow & Levinthal, 2003)</p> <p>Efficacy of cognitive search (Gavetti & Levinthal, 2000)</p> <p>Value of imperfect evaluation for exploration (Knudsen & Levinthal, 2007)</p>

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49

<i>Models of mutual adjustment</i>					
Coupled learning/ search models	Outcomes of alternatives depends on the choices of other agents	Draws on models of individual adaptation	Agents take actions that are interdependent, adapt to feedback <u>Key (additional) parameters:</u> Reward interdependence between agents, relative rates of learning, initial representations, communication constraints, distribution of decision rights	Slow versus rapid learning from own choices in updating representations Coordination versus exploration	Value of slow learning (Lounamaa & March, 1987) Benefits of incorrect representations (Puranam & Swamy, 2014)
Social learning	Presence of other agents in a collaborative setting		<u>Key (additional) parameters:</u> Connectivity to other agents, relative rates of learning	Slow versus rapid learning from choices of other agents	Value of weak ties (Miller et al., 2006; Fang et al. 2009) and information distortion (Schilling & Fang, 2012) to preserve slow learning
Imitation / Competitive Learning	Presence of other agents in a competitive setting		Agents update representations by observing the choices of others in a competitive setting <u>Key (additional) parameters:</u> Observability, imitation breadth	Imitation versus innovation	Complexity as imitation barrier (Rivkin, 2000; 2001) Imitation to reignite exploration (Czaszar & Siggelkow, 2010) Value of imperfect imitation (Posen et al., 2013)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49

<i>Models of information aggregation</i>						
Decision structures	<p>Actions correspond to projects to be selected</p> <p><u>Key parameters:</u> Rate of flow and outcomes of alternatives</p>	<p>Decision-makers differ in ability to evaluate outcomes of alternatives</p> <p><u>Key parameters:</u> Experience and expertise in alternative evaluation</p>	<p>Maximizing choice with imperfect evaluation</p> <p><u>Key parameters:</u> Experience and expertise in alternative evaluation (note representation and choice are combined in these models)</p>	<p>Updating based on reinforcement learning</p> <p><u>Key parameter:</u> Responsiveness to feedback (e.g. incentives for performance)</p>	<p>Type I (rejection of superior alternative) versus Type II error (acceptance of inferior alternative)</p>	<p>Design of reliable decision structures (Christensen & Knudsen, 2010)</p> <p>Decision structures and exploration (Knudsen & Levinthal, 2007; Csaszar, 2012)</p> <p>Limits to majority voting and averaging of opinions (Csaszar & Eggers, 2012)</p>