



Strategy Science

Publication details, including instructions for authors and subscription information:
<http://pubsonline.informs.org>

Signposts for Problemistic Search: Reference Points and Adaptation in Rugged Landscapes

Axel Zeijen, Manuel Romagnoli, Luigi Marengo

To cite this article:

Axel Zeijen, Manuel Romagnoli, Luigi Marengo (2025) Signposts for Problemistic Search: Reference Points and Adaptation in Rugged Landscapes. *Strategy Science* 10(3):263-279. <https://doi.org/10.1287/stsc.2023.0072>

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License. You are free to download this work and share with others for any purpose, except commercially, and you must attribute this work as “*Strategy Science*. Copyright © 2025 The Author(s). <https://doi.org/10.1287/stsc.2023.0072>, used under a Creative Commons Attribution License: <https://creativecommons.org/licenses/by-nc/4.0/>.”

Copyright © 2025 The Author(s)

Please scroll down for article—it is on subsequent pages



With 12,500 members from nearly 90 countries, INFORMS is the largest international association of operations research (O.R.) and analytics professionals and students. INFORMS provides unique networking and learning opportunities for individual professionals, and organizations of all types and sizes, to better understand and use O.R. and analytics tools and methods to transform strategic visions and achieve better outcomes. For more information on INFORMS, its publications, membership, or meetings visit <http://www.informs.org>

Signposts for Problemistic Search: Reference Points and Adaptation in Rugged Landscapes

Axel Zeijen,^{a,*} Manuel Romagnoli,^b Luigi Marengo^c

^aDepartment of Management, Technology, and Economics, ETH Zurich, 8092 Zurich, Switzerland; ^bFaculty of Economics and Business Administration, Friedrich Schiller University Jena, 07743 Jena, Germany; ^cDepartment of Business and Management, LUISS University, 00197 Rome, Italy

*Corresponding author

Contact: azeijen@ethz.ch,  <https://orcid.org/0000-0002-5506-2140> (AZ); manuel.romagnoli@uni-jena.de,  <https://orcid.org/0009-0005-9951-7118> (MR); lmarengo@luiss.it,  <https://orcid.org/0000-0002-8299-7830> (LM)

Received: July 13, 2023

Revised: March 18, 2024;

September 30, 2024; February 20, 2025

Accepted: March 2, 2025

Published Online in Articles in Advance:
March 31, 2025

<https://doi.org/10.1287/stsc.2023.0072>

Copyright: © 2025 The Author(s)

Abstract. Reference points form an essential element of organizations' problemistic search and adaptation behavior. Yet, if search is triggered by shortfalls compared with peers but alternatives are discovered on the fly, it is not clear whether and when peer comparison leads to better search outcomes. We contribute to the literature by studying how reference points guide search and which outcomes they allow organizations to achieve. Specifically, we develop a model of search in complex landscapes in which agents' search behavior is guided by an upper (aspiration-level) social reference point and a lower (survival-point) social reference point. In our model, agents move across a subjective "terraced" landscape that is a simplified transformation of the "real" one. The vertical positions and shapes of these terraces are determined by the agents' reference points and change over time as a result of their own and their peers' performance evolution. In turn, these terraces define the search space that is navigated and the outcomes that can be reached. We show that the upper and lower bounds play fundamentally different roles in the search process, with the upper bound being more important in the short run and the lower bound more important in the long run. Studying heterogeneous populations, we find that reference points drive dynamic trade-offs between how easily decision makers can reach their aspiration level and how much they benefit from doing so. We highlight the importance of both internal fit between reference points and external fit with environmental factors.

 **Open Access Statement:** This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License. You are free to download this work and share with others for any purpose, except commercially, and you must attribute this work as "Strategy Science. Copyright © 2025 The Author(s). <https://doi.org/10.1287/stsc.2023.0072>, used under a Creative Commons Attribution License: <https://creativecommons.org/licenses/by-nc/4.0/>."

Supplemental Material: The online appendix is available at <https://doi.org/10.1287/stsc.2023.0072>.

Keywords: complex adaptive systems • simulations • organizational evolution

Introduction

A fundamental insight from the behavioral theory of the firm is that decision making cannot be reduced to simply choosing among given alternatives: "A theory of bounded rationality must incorporate a theory of search" (Simon 1979, p. 502). One important consequence is that a decision maker must always choose between picking among known alternatives or searching for a new, possibly better one (March 1991). Behavioral theory postulates that decision makers solve this problem by referring to reference points: subjectively defined levels of performance that guide search behavior (Cyert and March 1963, Gavetti et al. 2012). Notably, two pivotal reference points have been highlighted: *aspiration levels* and *survival points*. The former reflects Simon's idea of *satisficing* and the objective of problemistic search (Posen et al. 2018); organizations keep

searching until they find outcomes that deliver performance corresponding to aspirations. Survival points, meanwhile, are minimum performance levels below which agents will not accept a fall and where organizational survival is deemed at risk (March and Shapira 1992, Audia and Greve 2006).

Recent decades have seen much progress in our understanding of how organizations set these reference points as a function of past and peer performance (Bromiley 1991, Bromiley and Harris 2014) and how comparisons of performance to both aspiration levels and survival points drive search and risk-taking (March and Shapira 1987, 1992; Audia and Greve 2006, 2021; Hu et al. 2011; Tarakci et al. 2018). However, we know less about the *constraints* that reference points impose on search, about the portions of the search *space* that are made accessible, and finally, about the possible *outcomes*

to which such search may lead. As Posen et al. (2018, p. 219) put it, “research on problemistic search has largely blackboxed the search process itself.” A first reason for this, as Posen et al. (2018) also note, is the lack of counterfactual outcomes in observational studies and such studies’ variance-driven approach. A second is that modeling work on this topic has thus far employed a stylized view of search that neglects important aspects highlighted even within behavioral theory: that search is local (Cyert and March 1963), path dependent (Simon 1955, Nelson and Winter 1982), and subject to complex interdependencies (Levinthal 1997, Rahmandad 2019).

This issue is particularly salient when it comes to social reference points. If search is triggered by shortfalls compared with peers, it is not clear whether and when peer comparison leads to better search outcomes. This issue reflects an inherent tension in our view of search with reference points that has yet to be addressed. Firms search until desired performance is restored, but alternatives must be discovered on the fly. If search is local and path dependent, comparison with peers may spur search without necessarily guiding organizations to better outcomes, especially if those peers are on different trajectories. In this paper, we address this tension by modeling reference-point-guided search in a complex search space and by asking the following question. How do reference points shape the constraints and outcomes of problemistic search? We focus primarily on social reference points and

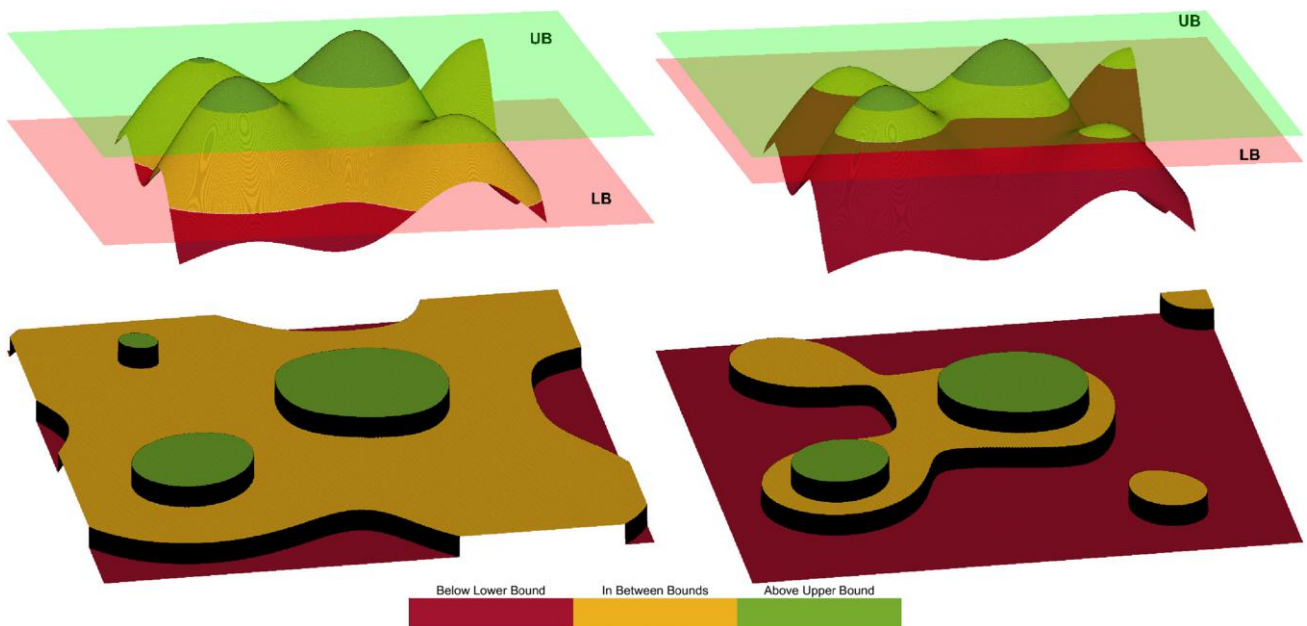
address past-based performance evaluation and absolute performance targets in alternative specifications.

Specifically, we develop an extension of the NK framework (Kauffman 1993, Levinthal 1997) that captures the fundamental postulations of the aspirations literature (Simon 1955, 1979; Audia and Greve 2021). First, agents choose whether to search based on their evaluation of current performance against an aspiration level. Second, they evaluate potential outcomes against both their aspiration level—which they seek to attain—and a survival point—a threshold that they seek to stay above. Both aspiration levels and survival points are defined in our model in relative (or social) terms. For example, an agent may aspire to be among the top 10% of performers while refusing to take any action that would land her in the bottom 15%.

The intuition behind our extension is that a complex landscape is transformed if outcomes are evaluated against reference points rather than against current performance. Instead of a “rugged landscape” with peaks and valleys, decision makers face a subjective “terraced” landscape. These terraces reflect how outcomes are evaluated, and because they are defined in relative terms, their size and shape change as the population progresses in its search. In turn, these terraces capture how decision makers’ reference points affect the search space that can be navigated and the outcomes that can be reached.

Figure 1 illustrates this transformation. The upper panels of Figure 1 show a rugged landscape intersected

Figure 1. (Color online) Rugged Landscapes Transformed by Reference Points into Terraced Landscapes



Notes. (Upper panels) Landscapes representing a two-dimensional search space, with height representing fitness values. Horizontal planes display the fitness values associated with upper bounds (UBs; upper terrace) and lower bounds (LBs; lower terrace) and separate coordinates into different zones (above UB, below LB, and in between). (Lower panels) The corresponding partition of locations in the landscape into desirable (upper terrace), acceptable (middle terrace), and unacceptable (lower terrace).

by reference points, resulting in each position being evaluated as above, below, or in between reference points (lower panels of Figure 1). We will call the aspiration threshold the “upper bound” (UB) and the survival threshold the “lower bound” (LB). These thresholds are associated with absolute fitness values, forming the horizontal planes in Figure 1. As the fitness values associated with the reference points increase (from left to right in Figure 1), and the horizontal planes “rise” in the landscape, most peaks shrink and disappear, the space available to navigate (in between peaks and valleys) becomes smaller, and it may become impossible to reach the remaining peaks.

Figure 1 captures stylized accounts of how reference points may aid or constrain firms’ search and adaptation efforts. Consider, for example, the adaptation challenge of Research In Motion (RIM) (BlackBerry) in the late 2000s as it faced the rise of iPhone and Android smartphones.¹ RIM was situated on a different “local peak” than Apple and Google,² with its focus on physical keyboards, hardware more generally, and professional clients (as opposed to leveraging multitouch screens and external app developers for mass audiences). As it turned out, the peak of Apple and Google was higher (the plane associated with the upper bound had risen). This performance shortfall induced RIM to introduce a clickable screen and an app store and to focus on software improvements. However, it was reluctant to make other changes that it perceived as too detrimental for performance, such as fully relinquishing its keyboard, even as multiple competing phone makers, such as Samsung and HTC, joined the Android ecosystem and gained ground (the plane associated with the lower bound was also rising). This left RIM “trapped” in the wrong part of the landscape by its lower bound, and it never restored aspirational performance.

We analyze how reference points affect search through three sets of simulation experiments. First, documenting the search process in a single population of homogeneous agents, we show that the boundaries separating the terrace levels shift upward as better outcomes are found, guiding many agents to better outcomes but trapping some in inferior parts of the search space. The terraced landscape thus provides signals about attainable and desirable outcomes but has limited value for locating better ones. Second, exploring all combinations of reference points, we show that the upper and lower bounds play fundamentally different roles in the search process. The upper bound is more important in the short run, and the lower bound is more important in the long run. Notably, search with sufficiently low lower bounds can result in entire populations finding the best-possible outcomes in the long run, even in the face of complexity. Third, we examine the strategic aspect of search with reference

points. Is there an optimal set of reference points, and what does it depend on? Decision makers face a dynamic trade-off between being able to achieve desirable performance in the short run and this desirable performance being sufficient to compete in the long run. We find that there are distinct ways that a search strategy can fail. Specifically, we find inverse U-shape relationships over the levels of individual reference points as well as over the distance between the lower and upper bounds (the “width” of the central terrace).

Beyond our specific findings, we aim to contribute to existing lines of work studying reference points, problemistic search, and search in complex landscapes. First, our modeling approach allows us to evaluate the normative (performance) implications of the predominantly descriptive accounts of problemistic search offered by extant theory to date. Second, we contribute to modeling work on aspirations and search by incorporating the notion of space into the search process. This allows us to disentangle the performance implications of aspiration levels from those of survival points (March and Shapira 1992, Hu et al. 2011). Additionally, we go beyond analyzing only long-run performance by characterizing performance evolution—and the impact of search strategies—over time. Third, our model contributes to the literature on adaptive search in complex environments (Baumann et al. 2019) by connecting it to another fundamental pillar of the Carnegie tradition: research on aspirations and problemistic search. We show how a search strategy with reference points interacts with the underlying problem space in the same way as the metaphorical scissors of Simon (1990), producing what we call “terraced” landscapes. Our perspective contrasts with existing work focusing on cognitive limitations and heuristics (e.g., Gavetti and Levinthal 2000, Knudsen and Levinthal 2007) and instead, focuses on search rules, thus granting a larger role to the organization’s own strategic choice.

Theory

Our study is founded in two streams of research, both originating directly or indirectly from the behavioral theory of the firm (March and Simon 1958, Cyert and March 1963) and its foundational notion of bounded rationality (Simon 1955). The first focuses on the role of reference points in driving search, whereas the second studies the properties of search in complex spaces. Both rely on the idea that organizations do not know *ex ante* the whole set of possible strategies at their disposal but rather, discover eligible candidates while searching for them in the neighborhood of their current status quo. In other words, to improve their performance, firms search locally (Cyert and March 1963) and only consider changes to some aspects of their current strategy. For extensive reviews of both literatures,

see, for instance, Shinkle (2012), Posen et al. (2018), and Audia and Greve (2021) for reference points and Baumann et al. (2019) for search on complex landscapes. In this section, we focus on each literature in turn and consider the connections and complementarities between them, which inform our modeling strategy.

Search, Satisficing, and Reference Points

Because search is costly and risky, individuals and organizations do not keep searching endlessly but “satisfice” with respect to some aspiration level (Simon 1979). Aspiration levels have acquired a central position in the behavioral theory of decision making as a key “upper” reference point for determining when search is triggered and paused. March and Shapira (1987, 1992) further extended this perspective by arguing that managers also attend to a lower reference point called the “survival point,” which is perceived as the lowest performance level below which the organization will exit. Audia and Greve (2006) show that the extent to which firms pay attention to either reference point determines the extent of risk-taking in adaptive search. On a similar theme, Gimeno et al. (1997) study individual entrepreneurs’ minimum performance thresholds as critical levels of opportunity cost that govern their decisions to remain in the market. The existence and influence of reference points can also be found in adjacent literatures (e.g., Wallace and Etkin 2024), most notably in psychology through prospect theory (Kahneman and Tversky 1979).

The relationship between reference points and organizational search has attracted considerable interest, generating a stream of theoretical and empirical research. Organizations evaluate and interpret current and potential performance based on reference points (Lant et al. 1992, Fiegenbaum et al. 1996, Moliterno et al. 2014). In turn, these reference points determine whether organizations engage in risky actions (Miller and Bromiley 1990; Greve 1998, 2003b; Posen et al. 2018) or stop searching because their satisficing targets have been met (March and Simon 1958, Cyert and March 1963). Adaptation is thus the outcome of a two-step process in which the risk-seeking choice follows from the recognition that current performance is below aspirations.

Aspiration levels are not exogenously given but endogenously modified in the search process. They depend on experience with the environment. If the decision maker can easily find many good alternatives, aspirations tend to rise (Simon 1979). In addition to individual experience (Greve 2003a, Argote 2013), aspiration levels also depend on the social context and in particular, on the performance of a relevant comparison group: competitors (e.g., Baum and Dahlin 2007, Boyle and Shapira 2012, Moliterno et al. 2014, Kacperczyk et al. 2015). Thus, firms set an aspiration level

based on a weighted function of past and peer performance (Cyert and March 1963, Greve 1998). More recent research has investigated how aspirations are regulated by attention rules (i.e., the rules that govern the changes in the relative weights of the different components) (Bromiley and Harris 2014, Blettner et al. 2015, Berchicci and Tarakci 2022).

Since the stylized formulation of social comparison as driving aspirations by Cyert and March (1963), empirical and modeling research has sought to clarify the role of social comparison in driving problemistic search—with mixed perspectives, specifications, and results (Bromiley and Harris 2014, Posen et al. 2018). These mixed findings are arguably because of an implicit conflation of two signals provided by peer comparison: that the focal organization is not doing well enough (a risk-taking signal) and that it could do better (a guidance or learning signal). Both are encapsulated in the view of social comparison of Cyert and March (1963, p. 123) as the vicarious counterpart to direct experience.

The risk-taking line is developed in March and Shapira (1987) and has been followed by much empirical research (e.g., Kacperczyk et al. 2015). Because of data limitations, however, such work reveals relatively little about what exactly can be learned from peer comparison. Usually, external performance feedback is measured as average or median industry performance levels (Greve 1998, Baum et al. 2005, Massini et al. 2005, Iyer and Miller 2008; see also Bromiley and Harris 2014 on measurement and operationalization of performance signals).

The guidance line is explicitly developed in the study of Baum and Dahlin (2007), which combines performance feedback with learning from organizations’ own and others’ experience (e.g., Thornton and Thompson 2001). This line of work considers more explicitly how comparisons with reference groups provide organizations with potentially useful information. Importantly, it has informed modeling studies starting from the model in March (1988) that seek to tie reference points to risky search and performance (March and Shapira 1992). Knudsen (2008) shows that the performance of the reference group guides risk-taking and that peer comparisons thus lead to higher performance than past-based reference points do. This argument is also echoed in later work (e.g., Dong 2021). Hu et al. (2011) explicitly consider strategies for setting reference groups and link social comparison with variance in search outcomes.

An additional value of the modeling approach lies in its normative contribution; despite the focus on performance evaluation against reference points in observational studies, it is hard to establish, without counterfactual outcomes, whether a performance level considered good enough for the organization is

appropriate for the external selection environment. It is not obvious whether and when higher reference points or targets ultimately result in higher performance (Posen and Levinthal 2012, Denrell et al. 2023); to preview our results, we find that often they do not.

This tradition of search models, however, views search as random draws and in doing so, falls short of accounting for two additional features of search that behavioral theory has emphasized: locality and path dependency. Locality means that decision makers typically can only access alternatives in the vicinity of what they already know and have mastered; radical changes are rare. Path dependency follows from the observation that search is subject to constraints, and a decision maker cannot readily access the entire space of alternatives. As search proceeds, these constraints typically become narrower, and large portions of the search space become inaccessible.

Search in Complex Landscapes

Locality and path dependency are captured most naturally by models of search in complex landscapes and in particular, by NK fitness landscape models (Kauffman 1993, Levinthal 1997; see Csaszar 2018 for a reading guide for a general audience and Baumann et al. 2019 for a recent survey). In these models, search unfolds in a space with a notion of adjacency or distance (allowing for the study of locality). In fitness landscape models, path dependence is a product of complexity (Simon 1996). If the performance implications of different choice dimensions interact, a fitness landscape becomes “rugged,” and local search will only lead to a “local peak.” Moreover, once an agent adopts a new alternative, she can access all of its neighbors but loses access to foregone alternatives.

Much of the NK literature in management since Levinthal (1997) studies organizations’ search strategies on rugged landscapes to investigate the implications of how the search strategies of boundedly rational decision makers interact with complex problems to produce search patterns and outcomes—the Simon (1990) “scissors” argument. These strategies are related to our study in that they have the effect of “simplifying” the complex search problem. Gavetti and Levinthal (2000), Knudsen and Levinthal (2007), and Yi et al. (2016) study search with coarse evaluation of the impact of choices. Csaszar and Levinthal (2016) directly study mental representations as simplified fitness landscapes. Podolny (2018), with the twist of exploring exceptional rather than bounded search capabilities, studies a form of “discerning” search with performance thresholds of strategy elements.

The integration of reference points in a search strategy also transforms the search problem but in a different way. Although the approaches listed above consider evaluation against current outcomes, they

contain no notion of aspirations or the problemistic search generated by subaspirational performance. (A notable exception is found in the NK-based experiment in Billinger et al. (2021), which considers whether decision makers search in the first place.) Search with reference points does not directly affect the search space but rather, the decision maker’s evaluation criteria or performance metric. To illustrate, a decision maker may be perfectly aware that two options are distinct and associated with different actual performance outcomes, but nevertheless, the decision maker evaluates them as equally valuable if neither provides aspirational performance. If because of myopia-based search limitations (Cyert and March 1963, Levinthal and March 1993), she cannot determine which outcome is more likely to lead to above-aspirational performance in the future, she would be indifferent between the options despite their (known and observed) performance difference.

In this paper, we examine indifference as produced by evaluation against relative or social reference points. Indifference between outcomes has been studied previously as a result of other mechanisms. Lobo et al. (2004) and Jain and Kogut (2014) study true fitness neutrality between outcomes; that is, if sets of outcomes are equivalent in a selection environment, organizations can leverage this neutral space to resolve the exploration-exploitation dilemma (March 1991) as they can introduce changes while maintaining performance. Multiarmed bandit models analyzing the exploration-exploitation problem also have some connection to our approach as the search strategies that they employ (e.g., a SoftMax algorithm as in Posen and Levinthal 2012) may produce neutral selection—although bandits, similar to random-draw models, contain no notion of space. In our model, neutrality derives from the behavioral implications of reference points. When reference points are defined in relative terms, our neutral spaces are not given exogenously but endogenously modified as agents search for locations with higher performance.

Our modeling strategy aims to reconnect the above-mentioned strands and analyze how reference points shape adaptive search behavior in complex settings. We do so by integrating specifications regarding reference points as drivers of action and filters for performance evaluation into the NK framework. Exploiting the benefits of agent-based models in theory development (Lave and March 1975, Levinthal and Marengo 2016), we seek to derive both descriptive and prescriptive findings regarding how reference points shape adaptive search processes and outcomes and how organizations with certain reference points fare in different environments and in competition with each other. We describe our model in more detail below.

Model

We take as our starting point the standard NK model, wherein agents navigate a fitness landscape in search of high-performance outcomes. There are N binary choice dimensions or components, leading to 2^N possible configurations or locations that individually map onto a fitness value and collectively form a fitness landscape. The correlation between fitness values of adjacent configurations (i.e., those differing by only one bit) is determined by the second model parameter K , which represents the extent of interdependencies among choice dimensions and thus, governs the complexity of the system. Parameter K can range from zero (each component's contribution to performance depends solely on its own value) to $N-1$ (such a contribution also depends on the value taken by all of the other $N-1$ components). "Simple" ($K=0$) landscapes are single peaked and "smooth," and a simple adaptive local search process will quickly converge on the unique optimal configuration. As K increases, the landscape becomes increasingly "complex" and "rugged," with an increasing number of local optima. A local search process performs poorly in complex landscapes as it quickly reaches the closest local optimum and gets stuck there.

The search patterns and outcomes in NK models follow from the assumptions about the search process. One common assumption, which we follow, is that fitness values can be observed without errors and before the configuration is adopted. The assumption from which we depart is the agents' acceptability criterion, which indicates whether, given its value, a new configuration would be acceptable and is adopted. The typical assumption is that agents search perpetually for new configurations as they are never "satisfied" with their current performance and always seek improvement, whereas they never accept a performance decline. Thus, standard NK models implicitly assume that (1) reference points are defined in absolute rather than social terms (i.e., agents only look at their own past performance without comparing it with others' performance) and that (2) decision makers consider an infinitely high aspiration level and a survival point equal to current performance.

We modify the acceptability criterion by explicitly specifying an upper bound (above which a decision maker is satisfied and explores no further) and a lower bound (below which she regards any outcome as unacceptable). We also assume that these two bounds are defined "socially" as shares of the population. For instance, $UB = 0.9$ and $LB = 0.05$ indicate that the agent is satisfied only when her performance is in the top 10% of the population and will not consider any action that would leave her among the 5% worst performers. In general, a higher *upper bound* means

that the decision maker is more ambitious; a higher *lower bound* means that she is less tolerant of poor performance.

In our model, these reference points govern the salience of potential outcome changes; the decision maker is sensible only to those outcome differences that reach or cross reference points (i.e., those that would take her organization from *below* a reference point to *above* it or vice versa) and is indifferent to differences that do not (e.g., March 1988, Lant and Shapiro 2008, Audia and Greve 2021).³ Thus, reference points shape two search heuristics; they determine whether the organization will engage in search in the first place (Posen et al. 2018) and if so, which changes it will adopt.

This specification implies that the objective function of the searching agent is transformed from trying to achieve higher fitness values to achieving a certain ranking among all agents. As a result, the two reference points jointly divide the underlying fitness landscape into three subspaces to form a "terraced" ranking landscape. The three levels consist of locations whose fitness is equal to or above the aspiration level (the set of desirable and search-halting locations), locations whose fitness is below the survival point (the set of unacceptable locations), and locations whose fitness lies between survival and aspiration (the set of acceptable but not satisficing locations). Figure 1, introduced above, visualizes this transformation; in the upper panels of Figure 1, rugged landscapes are shown with the two reference points as horizontal planes; the lower panels of Figure 1 display the corresponding terraced landscape.⁴

The search process is modeled as follows.⁵ In every simulation run, a population of agents is seeded in random positions in a randomly generated landscape, and all act simultaneously. In each time period, every agent compares her current performance with her reference points. If (and only if) the agent finds herself below the fitness level associated with her UB , she observes all local search options. If any of these would take her to or above the UB , she selects one of these at random. If not but any would take her to or above the LB , she selects one of these at random. If all search options are below the LB , she keeps the current configuration.

Our modeling assumptions are by no means the only ones possible, but in our view, they reflect the fundamental features of problemistic search. Our assumption that search stops when performance is greater than or equal to UB is in line with the theory of satisficing behavior by Simon (1955), although others have argued that when aspiration levels are met, the probability of change decreases but does not vanish altogether (e.g., Greve 1998). The assumption that agents keep searching as long as performance is

below the *UB* (i.e., until aspirational performance is achieved) also corresponds to a stylized account of the problemistic search (Posen et al. 2018).

This search process produces indifference between outcomes on the same performance “terrace” and implies a form of drift of an agent located on the middle terrace (i.e., between the upper and lower bounds). That is, when an agent in our model cannot attain aspirational performance with any local search option, she will choose to switch to a new configuration that may make aspirational outcomes available in future search steps. For example, a firm unable to match competitors’ performance benchmarks may adopt a new business process that does not increase performance in the short run but may eventually deliver aspirational performance if further changes are made that leverage the new process. The survival point then governs the decrease in performance that such a firm is willing to incur. This specification reflects the idea that the merits of an organizational change are not limited to its immediate performance implications but also, include the new opportunities that it opens up.

By combining features of problemistic search with locality and path dependence, our model complements existing modeling research on reference points and search, which operationalizes search as random draws (March 1988, March and Shapira 1992, Hu et al. 2011, Dong 2021, Denrell et al. 2023). Such random-draw processes also imply indifference between outcomes below the aspiration level but lack any notion of path dependence as core to understanding adaptation (Simon 1955, Nelson and Winter 1982, Levinthal 2021).

Finally, it is important to note that this transformation only affects the landscape “as searched by” the agents. This captures the essence of the performance-feedback perspective in that “it is not performance per se that influences behavior but rather how performance is assessed” (Audia and Greve 2021, p. 5). The actual landscape and its contours remain unaffected, and we consider and report performance on this actual fitness landscape; hazard rates in the external environment are not dependent on how performance is assessed internally but rather on performance per se.

Results

We subject our models to three sets of simulation experiments.⁶ The first aims to characterize search with social reference points. It is well known that in landscape models, agents remain on the first local peak that they encounter (local optimum). Population-wide search quickly ceases, and any initial heterogeneity among agents is only partly diminished as agents remain scattered across the numerous local peaks of the landscape. With social reference points, in contrast, we find that agents can access and explore much larger

portions of the landscape (i.e., the central terrace of locations that deliver performance between the survival point and the aspiration level). Agents ultimately converge on the locations with highest performance because the performance levels associated with both reference points keep rising as agents search. The final outcome is a higher rate of convergence of agents on locations with the highest performance.

Second, we analyze how different sets of reference points lead to different search outcomes. To this end, we generate different (homogeneous) populations, each endowed with a set of reference points, and study which sets lead to the highest collective performance. We find that although aspiration levels matter most in the short run, long-run performance is only determined by the survival point. Notably, the highest collective (and individual) performance in the long run is achieved by populations with very low survival points and any aspiration level above approximately 0.4. This counterintuitive result is explained by the fact that an agent will search for higher performance even when her own aspiration level is fairly low because the rest of the population also improves. This search effort is frustrated when agents are “trapped” on local optima. If agents’ survival point is low and their aspiration level is sufficiently high, the “neutral” space on the central terrace will be relatively large, making it unlikely that agents will remain trapped.

In our final set of experiments, we consider heterogeneity in reference points within the population. Organizations may differ in terms of reference points, and these differences may drive heterogeneity in performance in competitive environments. In these simulation experiments, we additionally introduce a selection mechanism of competitive pressure, so badly performing agents are at risk for exiting the industry. We show that when agents differ in reference points and in the presence of competitive pressure, the highest performance is achieved by agents with moderate aspiration levels ($UB = 0.75$) and moderate survival points ($LB = 0.25$). The benefit of moderate reference points follows from the trade-off between adopting targets associated with high performance and these targets being easily attained. Moreover, we show that the relationship between performance and the distance between reference points follows an inverted U-shape. Finally, we show that the effects of social reference points depend on both the characteristics of the rest of the population and competitive pressure; those reference points that do well in competitive settings lead to considerably worse outcomes in homogeneous populations.

All results shown below are averages of 10,000 simulations, each generating a new landscape, with landscape fitness values rescaled to the zero-one interval. Time is displayed on a logarithmic scale to highlight dynamics in the short and medium runs.

Search with Reference Points

We simulate a population of agents that all have the same reference points. These agents evaluate their performance (and that of search options) in relation to their reference points and accept outcomes according to their acceptability criteria. To characterize the search process, Figure 2 displays four time trends for a population with $UB = 0.75$ and $LB = 0.25$. Panel (a) of Figure 2 displays average performance (black line with diamond markers) and average performance conditional on being located above (green line with round markers), in between (yellow line with triangular markers), or below (red line with square markers) the two bounds. Panel (b) of Figure 2 displays the fitness values associated with the population's upper (green line with round markers) and lower (red line with square markers) bounds. Panel (c) of Figure 2 displays how many agents are above (green dotted area), in between (yellow striped area), or below (red cross-hatched area) the bounds. Finally, panel (d) of Figure 2 displays how searching agents (i.e., those in between bounds) view the landscape. We consider all locations in the landscape and determine the share of locations that are desirable (green dotted area in Figure 2(d)), acceptable (yellow striped area in Figure 2(d)), unacceptable (red crosshatched area in Figure 2(d)), and inaccessible (blue area in Figure 2(d)). Figure 1 serves as an illustrative guide to the search process described here; Online Appendix A presents a corresponding small-world example in an actual NK landscape.

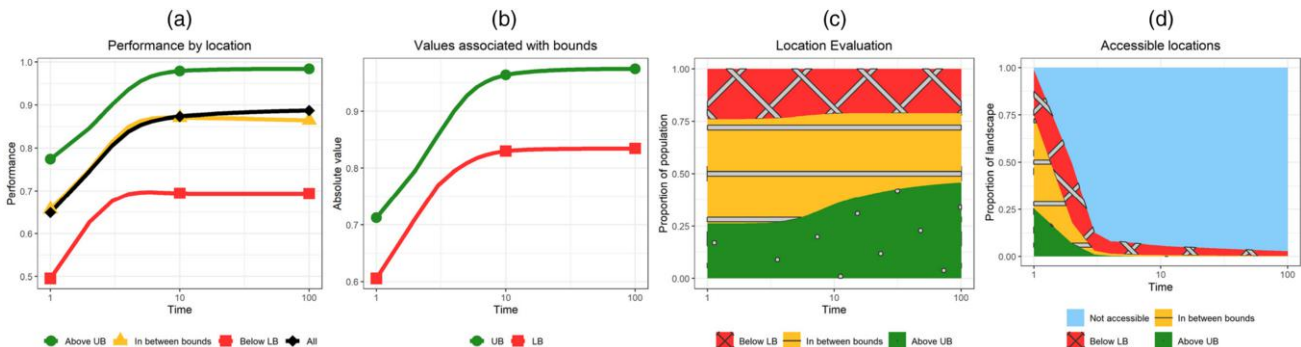
The search process unfolds as follows. At the start of the simulation, a quarter of the agents are satisfied and will not search (because $UB = 0.75$). The other agents do search; some discover search options that are desirable (at or above UB) and cease search, whereas others move to acceptable but not desirable positions (above LB but below UB) and continue searching. Still others cannot find any positions above the LB and remain in

place. During this search process, the fitness levels associated with both the reference points shift up over time. The *absolute* performance needed to perform well compared with others increases. As a result, the “terraced landscape” that agents navigate shifts over time; the part of the landscape that is considered unacceptable (below LB) grows, whereas the part that is considered desirable (at or above UB) shrinks until it consists only of the highest few “peaks” of the landscape (see the right panels of Figure 1).

The search process at play here is markedly different than how search under complexity is usually considered. Search guided by reference points does not result in agents scattered among local peaks, getting stuck in the contours of the landscape. Rather, it results in agents racing to find a dwindling set of good outcomes and getting stranded along the way by a “rising tide” of competition. During their search for good outcomes, agents can make use of the distance between their reference points and the space in the landscape corresponding to this “span of indifference.” For this population, all outcomes between the 25th and 75th percentiles of population performance can be accessed; they form the middle “terrace” in Figure 1 that connects much of the landscape. Panel (d) of Figure 2 displays the portion of the landscape that can be observed from this middle terrace on which the agent is located. It shows that the middle terrace quickly shrinks, even though the difference between the fitness levels associated with the reference points remains substantial. Panel (d) of Figure 2 shows that initially, the entire landscape can be accessed, with half of the locations considered acceptable (as in the left panels of Figure 1, the middle terrace connects most of the search space). After only a few periods, however, most of the landscape becomes inaccessible; search ceases when the search space runs out.

Finally, consider the search outcomes. In the first periods, all searching agents have ample opportunities

Figure 2. (Color online) Search with Reference Points



Notes. $UB = 0.75$, $LB = 0.25$, and there are 100 agents; the horizontal time axis is logarithmic. (a) Average fitness conditional on whether agents are located above, below, or in between the upper and lower bounds. (b) Absolute fitness values associated with upper and lower bounds. (c) The proportion of the population located above, in between, or below upper and lower bounds. (d) The fitness landscape as perceived by agents in between bounds (the proportion of the search space accessible from the current location).

to find high-ranking outcomes, and overall fitness increases as agents overtake each other. After just 10 periods, however, those agents that have found the highest-performing outcomes set a fitness benchmark for all of the others (that is already close to the global peak). Whether others can also achieve such high-performance outcomes now depends on whether agents searching in the indifference space can reach aspirational outcomes or whether they are stuck. This can be seen in the colored lines in panel (a) of Figure 2. The fitness of agents conditional on their terrace is stable; aggregate fitness increases are driven by agents shifting from the middle to the top terrace (i.e., from performance below the upper bound to above it). This can also be seen in panel (c) of Figure 2; gradually, agents on the middle terrace find the few highest peaks and join the agents already there.⁷

It is worth re-emphasizing that search with reference points means that agents satisfice—and the agents that we study aim no higher than to be in the top 25% of the population. Yet, this search strategy implies that eventually only top performance satisfies aspirations and that many agents achieve it. This result is because of the cues that are provided by the perceived landscape. It informs the decision maker about the range of possible outcomes and how well she needs to perform to exceed benchmarks. Importantly, however, it does not tell her *how* to achieve desirable performance outcomes and thus, does not help her improve performance directly. The decision maker's subjective terraced landscape, generated by her reference points and collective performance scores, forms a sort of "signpost" for problemistic search; it determines where search can unfold and the extent of exploration.

Reference Points and Collective Outcomes

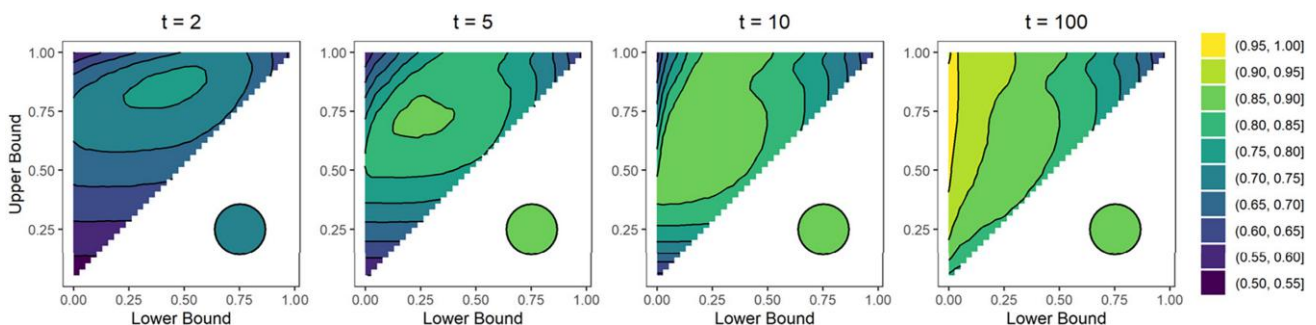
In the following analyses, we study how different reference points lead to different collective outcomes. Figure 3 displays performance outcomes of the same simulation experiment for all combinations of upper

and lower bounds. Each combination of reference points is plotted as coordinates on a grid in Figure 3, with colors representing the height of average performance (that is, the average performance reported in panel (a) of Figure 2 is one "pixel" in each panel of Figure 3). We limit the analysis to cases where $UB > LB$, leading to a triangle of data points, and for comparison, we plot results of the standard NK model in circles in the blank space in the plots in Figure 3. In Figure 3, the panels from left to right display snapshots of the same simulation. Online Appendix B presents more detailed results for selected combinations of reference points as well as for the standard NK model.

Figure 3 shows two main findings. First, there is a large range of outcomes, with some reference-point combinations not aiding search at all, even in the long run, but also, many combinations resulting in the entire population finding the maximum performance (i.e., the highest peak(s) of the landscape). Agents who aim for only the highest performance but are also unwilling to tolerate bad performance during the search process do the worst; for these agents, reference points constrain search more than they aid it. Highest performance is instead achieved by populations of agents characterized by a very low LB for any UB equal to or greater than approximately 0.4. It requires only a very low benchmark for all agents to perform very well.⁸

Second, the difference in performance between reference-point combinations varies over time, and the UB and the LB play different roles over time. In the short run, the horizontal stripes in Figure 3 are a sign that the upper bound (vertical axis) is the main determinant of performance, although the circular core of higher performance indicates that a medium-low (and decreasing) value of LB is required to maximize performance. Even after 10 periods, the stripes already tend to turn vertical, and highest performance moves toward low values of LB . Finally, in the long run, the pattern becomes markedly characterized by vertical

Figure 3. (Color online) Reference Points and Performance over Time



Notes. Panels indicate snapshots over time. Coordinates within panels (pixels) reflect reference point combinations for homogeneous populations of 100 agents each; shade reflects average population performance. Circles placed in the blank space represent performance in the standard NK model.

stripes in Figure 3, implying that LB is the main determinant of performance.

This analysis thus sheds more light on how search outcomes are shaped by reference points, particularly by the distance between reference points and the “width” of the central terrace that this distance generates. Treating more outcomes as acceptable (lower LB) enlarges the search space and allows agents to converge on the global peak. However, this comes at a cost in terms of speed. Increasing the UB provides guidance mainly in the short run.

Our next analysis relates the importance of the search space width to complexity. A well-known finding from the organizational search literature is that complexity severely hinders adaptability; the fitness landscape becomes more “rugged,” with more local peaks, and local search leaves organizations stuck in suboptimal solutions. This problem can be mitigated by being more tolerant of short-run fitness loss in order to achieve superior outcomes in the medium-long run (Knudsen and Levinthal 2007, Jain and Kogut 2014).

Figure 4 extends our baseline experiment for environments with varying levels of complexity and different levels of LB . We plot the development of average performance over time, organized by LB , for low ($K=1$, left panel), medium ($K=3$, central panel), and high ($K=7$, right panel) complexity. As expected, complexity hinders the discovery of high-performance outcomes, but a large distance between reference points (i.e., with $LB=0.01$, solid lines with dotted markers) helps organizations overcome competency traps, albeit at the expense of speed.

Competition on Reference Points

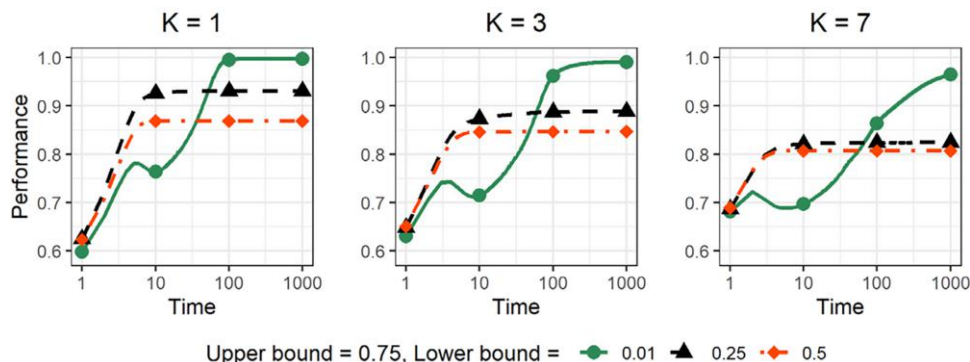
Thus far, we have simulated populations of agents with homogeneous reference points. We now turn to an investigation of heterogeneous agents to explore the interaction between decision makers with different upper and lower bounds. This type of analysis is informative for multiple reasons. First, it allows us to understand the implications of reference-point-guided search strategies more generally. Rather than studying

such strategies in an environment where all decision makers have the same strategies, we can observe whether the results hold when others follow different strategies. For example, we find above that given enough time, even decision makers with low upper and lower bounds are all able to collectively find the best outcomes. Yet, is such a search strategy also appropriate if the organization is surrounded by competitors pursuing higher targets and may not have the luxury of time to achieve the best outcomes? Similarly, we observed that the width of the search space is a function of others’ performance. If these relevant others follow different search strategies, this likely affects focal decision makers’ search opportunities.

Second, this analysis allows us to derive normative implications. In a large population of heterogeneous agents, simulation experiments are able to identify which agents perform best—and consequently, which search strategies would have led any one agent to superior outcomes. This aspect is particularly instructive given the lack of counterfactual outcomes in observational studies. Revisiting an earlier example, if we interpret RIM’s difficulty in adapting to the iPhone’s success as because of its too-high *lower* bound, would it have been better off with a lower LB ? Our results below suggest that indeed it would have been but only if this change were also coupled with the adoption of a low(-er) UB (i.e., no longer aiming to be the dominant smartphone manufacturer).

To undertake this analysis, we divide the total set of agents (here, 90) into groups of 10 and assign each group a unique set of upper and *lower* bounds. We employ a three-by-three factorial design, with lower bounds of 0.01, 0.25, or 0.5 and *upper* bounds of 0.51, 0.75, or 1, and we present results including the very long run (1,000 periods). Our simulation experiment proceeds as described above. However, given that agents differ in their reference points, the same ranking position in the population may be desirable for some agents but merely acceptable or even unacceptable for others, and therefore, it may produce different

Figure 4. (Color online) Performance by Reference Point Distance and Environmental Complexity



search behavior for different agents. In this simulation experiment, we are interested in comparative performance implications. To this end, five agents per period exit and are replaced by new entrants using a roulette-wheel algorithm. Note that performance outcomes under competitive pressure are of course conditional on performing well enough to survive (see also the related discussion in Levinthal and Posen 2007). For this reason, we only display survival; average performance is reported along with corresponding results in the absence of competitive pressure in Online Appendix B.

Intuitively and based on the preceding analysis, when agents compete with different reference points, they face a trade-off when setting their aspiration level. If it is set too low, the agent will be too easily satisfied, stop searching, and may be overtaken by agents with higher aspirations. Indeed, in the short run, the performance associated with low-to-moderate aspiration levels does not exceed average performance, putting these agents at risk for exit.⁹ On the other hand, if an agent's aspiration level is set too high, she is unlikely to find such high-performing alternatives. We showed above that, in the long run, the probability of reaching high-performance locations mainly depends on having a low survival point. However, in the presence of competitive pressure, such agents are at risk for exit during search. The agents with the highest probability

of survival are those who strike a balance between moderate aspiration levels ($UB = 0.75$) and moderate survival points ($LB = 0.25$).

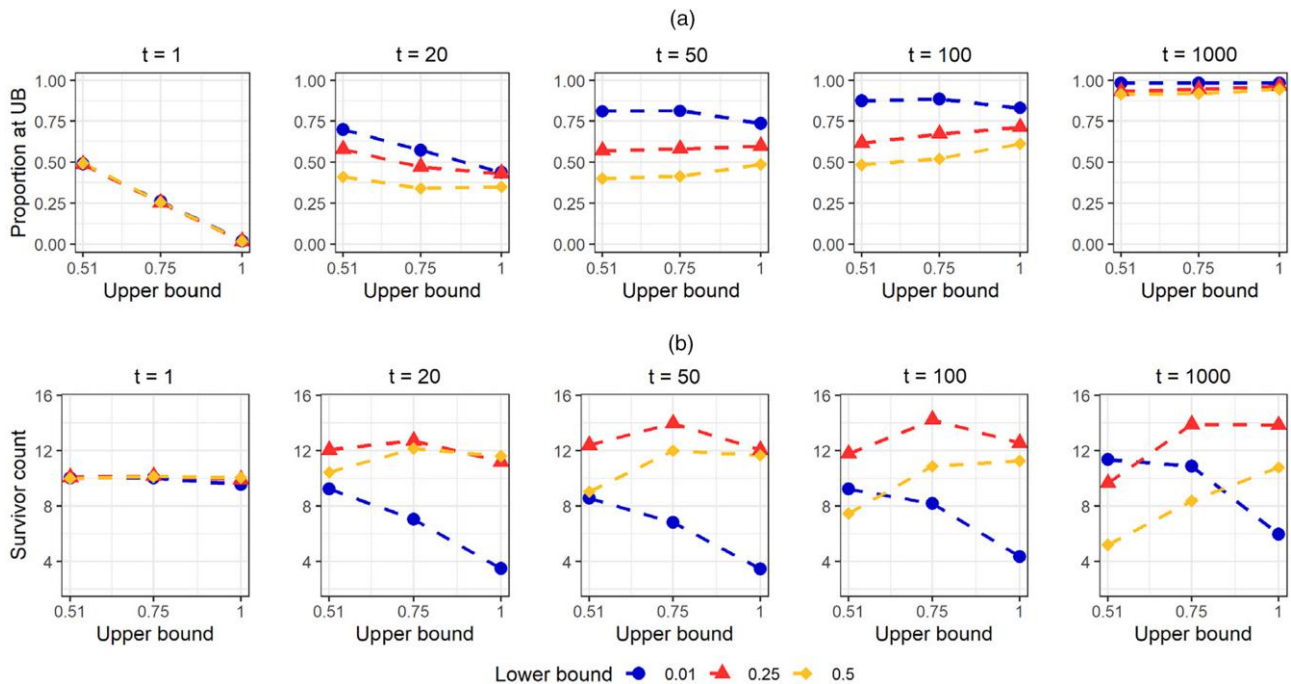
In Figure 5, we show how different combinations of reference points help decision makers reach desired outcomes and the implications for survival. Figure 5 shows snapshots of the state of the population at different points in time. Figure 5(a) shows the proportion of agents who have achieved their UB; Figure 5(b) shows the count of surviving agents.

Figure 5 shows two outcomes. First, there is indeed a trade-off between the UB serving as a suitable target and reference points jointly serving as a guide to reach this target. Moreover, this trade-off is dynamic over time. Initially, firms are mostly guided by the UB; a high UB provides better performance if reached, whereas a low UB is easier to reach. After a while, however, it is the LB that determines whether agents are able to reach the UB.

Figure 5(b) additionally shows that agents with moderate UBs and LBs ($UB = 0.75$, $LB = 0.25$) perform best. This specific result as well the overall pattern is robust to a wide range of exit and entry rates. Only when competitive pressure is very low (one agent exiting per period) do agents with very low UBs and LBs ($UB = 0.51$, $LB = 0.01$) eventually dominate: from around period 500.¹⁰

Note that this best-performing cohort is precisely the type analyzed in our initial simulation experiments

Figure 5. (Color online) Upper-Bound Achievement and Survival



Notes. (a) Proportion of agents at the upper bound by group. (b) Survival by group. Nine groups of 10 agents with roulette-wheel replacement of 5 agents per period. Panels indicate snapshots over time. Upper bounds are plotted on the horizontal axis, and lower bounds are indicated by lines and markers.

(Figure 2) and that it does not perform as well as groups with other reference points in homogeneous populations (see Figure 3). The mechanism explaining this difference is the interaction between agents' reference points and the performance of all agents. This is further demonstrated by Figure 6, which displays the performance of agents with the same reference points ($UB = 0.75$, $LB = 0.25$) in different populations. The solid green line with dotted markers in Figure 6 displays homogeneous populations (this line corresponds to the agents displayed in Figure 2). The dotted blue line with triangle markers in Figure 6 displays agents with $UB = 0.75$ and $LB = 0.25$ in a heterogeneous population with other worse-performing agents.¹¹ These worse-performing peers effectively maintain a wide distance between fitness levels associated with the reference points, which guides search for longer. Finally, the dashed red line with diamond markers in Figure 6 (corresponding to the agents displayed in Figure 5(b)) displays the additional effect of competition in driving performance up.

These results show that the appropriateness of reference points (their external fit) is contingent not only on complexity but also, on the degree of competitive pressure and on the performance of others (which are, in turn, shaped by their reference points).

Second, Figure 5 shows that there is an element of *internal* fit (Siggelkow 2002); the panels in Figure 5 display an inverted U-shape effect of the distance between the upper and lower bounds on performance outcomes. Figure 5(a) shows that those agents with a high LB (yellow lines with diamond markers) are more likely to eventually achieve their UB if the UB is *higher* (points on the right side in each panel)—in contrast to intuition and short-run results. Conversely, agents

with a low LB (blue lines with dotted markers in Figure 5(a)) are consistently less likely to achieve their UB if they have a high UB . To understand why, recall from the previous analysis how the reference points jointly guide search; if the LB is very low and the UB is very high, agents regard many options as acceptable but not desirable and hence, search too much. Conversely, agents with less distance between reference points consider only a few search options as acceptable and thus, search too little. This inverted U-shape pattern is even more evident when we evaluate survival outcomes in Figure 5(b). Agents with the smallest and largest search spans ($UB = 0.51$, $LB = 0.5$ and $UB = 1$, $LB = 0.01$, respectively) perform worst.

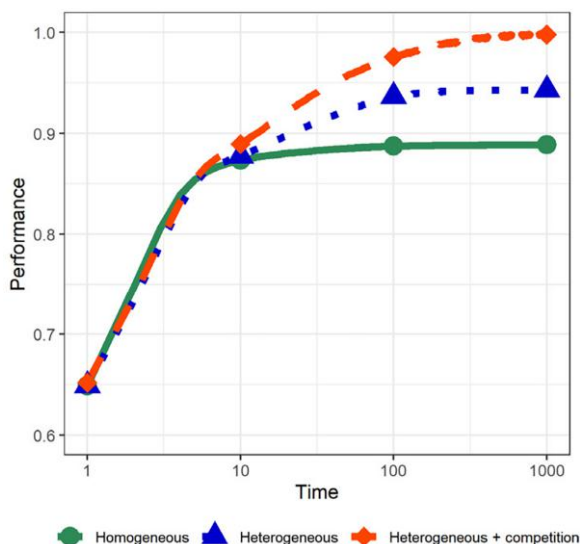
Robustness and Alternative Specifications

We have minimized our adaptations to the standard NK model to isolate the core tensions of interest (Knudsen et al. 2019) and only investigated social reference points with coarse decision rules. One can, of course, consider different assumptions about how reference points are formed and expressed, how accurately they are perceived, and how strictly they guide the search process. Online Appendix D contains results with different specifications and generally supports our central result that decision makers require a moderate “middle terrace” to navigate the landscape. Furthermore, these additional results serve to make more connections between our simplified model and the existing approaches and results in the literature.

Specifically, we modify how reference points are formed and influence search strategy using five alternative specifications. First, we loosen the assumption that decision makers are truly indifferent between outcomes in between the two bounds by employing a SoftMax decision rule anchored on the two reference points (Luce 1959). This rule is common in the bandits literature (see, for example, Posen and Levinthal 2012 and Puranam et al. 2015). This extension generates “relative indifference” and effectively reduces the span of indifference; outcomes close to the UB are likely to be considered preferable, whereas outcomes close to the LB are unlikely to be selected. As a result, search outcomes are better in the short run but worse in the long run.

Second, we investigate absolute rather than rank-based reference points (see also Denrell et al. 2023 on this contrast). As noted above, this analysis is similar to earlier work on neutrality in NK landscapes (Lobo et al. 2004, Jain and Kogut 2014) and largely replicates its findings. Third, we investigate reference points that are a weighted function of past and peer performance (Cyert and March 1963, Greve 2003a, Argote 2013, Bromiley and Harris 2014). This specification tunes the degree to which indifference between outcomes shapes the search process. Indeed, as past performance

Figure 6. (Color online) Performance of Agents with $UB = 0.75$ and $LB = 0.25$ in Different Populations



drives the reference points more dominantly, the model becomes equivalent to the standard NK analysis (Levinthal 1997). However, performance increases for all types of agents if past performance is considered; by basing search strategies on both social and historical reference points, agents are guided in a complementary way by cues about potential outcomes globally (information from peers) and locally (information from the past). This finding recalls the insights in Gavetti and Levinthal (2000) regarding the value of employing both forward- and backward-looking search and the interaction between memory and neutrality discussed in Jain and Kogut (2014).

Fourth, we follow the approach in Knudsen and Levinthal (2007) and investigate imperfect evaluation. Specifically, we introduce noise into the evaluation of current performance and of search options in relation to reference points. The basic results remain unchanged. Finally, we loosen the assumption that agents remain in place when they cannot find outcomes above the *LB* and do not move from one unacceptable location to another equally unacceptable one. Specifically, we assume that in this case, agents choose a local adaptation at random. In the long run, this change ensures that all agents reach the best outcomes; in competitive environments, it gives an edge to agents with a high *LB* who are less at risk for getting trapped, but it does not change the overall pattern of results.

Discussion

Reference points continue to occupy a central place in behavioral explanations of search and adaptation. In this paper, we advance theory that has mostly focused on search triggers and risk-taking by analyzing how social reference points shape the search space and possible outcomes. We do so by explicitly modeling space and complexity in problemistic search—which entails locality and path dependency. So-called “rugged search landscapes” are transformed into subjective and endogenously changing “terraced” landscapes when agents evaluate outcomes against social reference points instead of their own current performance. We show that the shape of these terraces and in particular, the “width” of the central terrace determine where and how much decision makers search and which outcomes they can eventually reach. Specifically, we show that the upper and lower bounds play different roles in shaping search over time and that the lower bound is the main driver of long-term performance. Furthermore, we show that the suitability of reference-point combinations is contingent on external factors (complexity of the search problem, competitors’ reference points, and competitive pressure) and that there are downsides to employing a search space that is either too wide or too narrow. In

the following sections, we outline our contributions to the literatures on search and strategy.

Contribution to the Literature on Search

Our main contribution to the literature on problemistic search is that we unpack the search process as guided by reference points in greater detail. Our perspective considers that the choice set available to organizations is shaped by their reference points, their current activity profile (and its current performance), and the performance of others. Moreover, it links these choice sets to performance outcomes: the likelihood that problemistic search will result in above-aspirational performance in the short and long runs and the likelihood that above-aspirational outcomes will be sufficient to survive in competitive environments. This connection has been recognized as important in the literature (Lant and Shapira 2008, Posen et al. 2018) but has been hard to establish with observational data (lacking counterfactuals that allow nonavailable and nonpursued opportunities to be studied) or with search models based on random-draw performance outcomes. By linking opportunities to outcomes, our approach allows us to append normative implications to the mostly descriptive account of problemistic search offered by extant theory.

A central finding in the literature considering both aspiration levels and survival points is that firms attend to both, with relative salience based on which reference point is closer to current performance (Audia and Greve 2006). Our model, with agents navigating a space bounded by desirable and unacceptable outcomes (see again Figure 1), produces a compatible perspective; agents close to either “terrace boundary” are more likely to encounter outcomes that traverse that boundary. Our perspective complements the existing literature in that it connects the height of and distance between reference points to space in the landscape and shows how reference points shape the portion of the search space that is accessible.

Our results suggest a much more important role for the lower bound or survival point than previous literature has assigned it. In the literature, however, the lower bound often has a qualitatively different interpretation; the setting of a survival point is often considered an assessment of the organization’s available stock of resources (although it can also have other antecedents—see, e.g., Gimeno et al. 1997). Our model incorporates the downside of low performance by making hazard a function of comparative performance, but future work could examine organizational-level differences in available resources, potentially also linking our findings to slack search (Chen and Miller 2007).

Past research on reference-group selection has focused on performance levels (Knudsen 2008, Moliterno et al. 2014)—different-sized “ponds” that organizations choose

to swim in (Frank 1985). Our model suggests that distance (in technology, practices, or policies) is an important second dimension of comparison. Performance signals from peers on a distant “peak” may hinder organizations more than they help. Recent advances in topic modeling (e.g., Arts et al. 2023) may allow future work to incorporate this dimension in work on social aspirations (see, for example, Lechler et al. 2023).

We show that reference points, far from necessarily being a source of suboptimal choices, allow for more exploration and reduce lock-in into local peaks. This insight relates aspiration levels to the evolvability of organizations (Ethiraj and Levinthal 2004). The importance of indifference has already been highlighted by prior work on neutrality in complex landscapes (Kimura 1983, Wagner and Altenberg 1996, Wagner 2005) and has been studied in the management literature of Lobo et al. (2004) and Jain and Kogut (2014). This paper contributes to that line of work as it considers neutrality not as a given and stable feature of the *external* selection environment but rather, as an emergent and endogenous feature of the *internal* selection environment (Burgelman 1994, Levinthal 2021). This means that neutrality is a function not merely of coarse evaluation but also, of the decision maker’s choices and those of the rest of the population. A terraced fitness landscape is derived from an agent’s search strategy, and its shape is endogenously modified as the focal agent and the rest of the population keep searching. The contours of the fitness landscape and its degree of neutrality will differ between organizations at any point in time and will differ over time for the same organization.

Our perspective of reference-point-guided search complements existing perspectives of adaptive search in complex environments both within the NK literature and more generally. To the literature on search using NK models, we contribute by developing an extension of the canonical NK model (e.g., Levinthal 1997). This extension relies on heuristics for whether to search and how to evaluate results. It reveals that the canonical model is a rather peculiar special case where agents focus only on their own past performance, have infinitely high ambition, and will not tolerate any performance decrease. Our model embeds a richer set of elements from the Carnegie tradition to show that social reference points are important drivers of organization-level search and adaptation patterns as well as population-level outcomes. We hope that our model adds to the value of the NK model as a tool for developing theory in the behavioral tradition (Baumann et al. 2019). Beyond its focus on reference points, it contributes to a small but growing line of work using the NK model that considers agents’ choices and performance outcomes as dependent on others’ choices (Lenox et al. 2006, Giustiziero et al. 2022).

More specifically, our perspective on search complements existing models that incorporate ideas from the Carnegie tradition. Such models, often under the heading of “mental representations” (Csaszar and Levinthal 2016), consider how searching agents act upon a (typically simplified or coarse) representation rather than the actual fitness landscape (Gavetti and Levinthal 2000, Knudsen and Levinthal 2007). In contrast to this line of work, we consider a mechanism that does not necessarily rest on cognitive limitations but rather, rests on search rules, implying a larger role for strategic choice by the organization.

Beyond the NK model, our findings can be useful for the broader literature on adaptation and inertia (e.g., Eggers and Park 2018). Our illustrative case of RIM’s (BlackBerry) adaptation problem has been previously explained (in Argyres et al. 2019) as (because of prohibitive comparative adjustment costs) an explanation focused on the contours of the fitness landscape. Other accounts of inertia have focused on firms’ search strategies (e.g., Tripsas and Gavetti 2000). Our perspective links these explanations by studying search strategies in tandem with the nature of the search problem (Simon 1990).

Empirical implications may be drawn in relation to different (inter-)organizational settings and serve as the basis for observational studies. For example, our homogenous populations could be interpreted as exemplifying organizations’ expectation setting (e.g., our setting $UB = 1$, $LB = 0.5$ is one where top performance is rewarded but deviation from best practices is discouraged) or industrial environments (e.g., setting $UB = 1$, $LB = 0.01$ characterizes a more entrepreneurial culture where decision makers aim for exceptional performance and risk-taking is encouraged; populations with $UB = 0.51$, $LB = 0.5$, on the other hand, may represent mature industries with clear benchmarks and no upside to surpassing them). Our results may help future work understand, for example, how individuals’ or firms’ reference points translate into collective outcomes or how organizations fare if environmental conditions change while aspirations remain constant. Our analysis of heterogeneous populations can be used to study industries in which firms have different approaches to (problemistic) search.

Contribution to the Strategy Literature

Our model complements insights from related modeling work in the strategy literature. Consider, for instance, the long-standing topic of firm heterogeneity. Previous literature adopting the NK model has emphasized the structural factors underpinning the emergence and persistence of such heterogeneity. Notably, past research has seen heterogeneity as a function of landscape complexity (Levinthal 1997) or of interdependencies between a firm’s strategy and positioning (Adner et al. 2014). Our

model complements these insights by showing how firms' decisions are the outcome of the interaction between environmental characteristics and the decision maker's evaluation of those characteristics. We show that the evolution of heterogeneity depends not only on external factors but also, on the criteria that managers adopt to orientate themselves within the landscape.

Our study also connects with work on the behavioral implications of incentives for firm performance. For example, our model can be seen as a continuation of the behavioral analysis of industrial policies (Li and Csaszar 2021), in which social reference points are interpreted as regulations anchoring *targets* to population dynamics (for example, to achieve certain benchmarks for reducing emissions rather than constraining organizations' decisions). In our approach, the policymaker does not need to know the precise structure of the landscape.

Our results have normative implications for the suitability of performance targets (to the extent that they can be controlled). At the industry level (e.g., for policymakers), our work has implications for the design of incentives and selection mechanisms in circumstances where what matters is the collective performance of a population of organizations. We show that conditional on the complexity of the task environment and competitive pressure, high relative targets may be unnecessary in the long run. At the organizational level, our work points at distinct ways that a search strategy may fail; organizations may search too broadly or too narrowly depending on the distance between reference points, and reference-point sets that deliver temporary advantage in the short run may hinder adaptation in the long run. Furthermore, we show that the suitability of reference-point sets for focal organizations depends on environmental complexity, competitive pressure, and the reference points adopted by competitors as external contingencies.

Limitations and Future Research Opportunities

Our model embodies a simplified operationalization of search with reference points. We have explored extensions that incorporate other findings in the aspirations and search literature but only to the extent that they influence search with aspirations and reference points. Future work could expand our model in different ways—for example, by incorporating the problematization of reference groups (Hu et al. 2011, Moliterno et al. 2014). It could also broaden our perspective—for example, by also considering belief-based search rather than aspiration-based search as outlined in Keil et al. (2023). The NK model as an adaptive search model faces inherent limitations in studying forward-looking strategies. Other approaches beyond NK could integrate our perspective into the growing line of work that considers how decision makers adjust their choices

to assumptions about peer actions and outcomes (Menon 2018, Menon and Yao 2024).

Conclusion

In this paper, we have developed a model that integrates problemistic search with social reference points and adaptation in complex environments. We study how peer comparison endogenously transforms the search space, with implications for search opportunities and outcomes. Our results show that upper and lower reference points shape population search processes differently, and they highlight the importance of internal and external fits of reference points for organizational adaptability. We contribute to the understanding of organizational search processes—specifically, by providing a model that can inform future research on the connections between organizational adaptation and population dynamics.

Acknowledgments

The authors thank Jose Arrieta, Jerker Denrell, Thorbjørn Knudsen, and Dan Levinthal for useful comments and suggestions on earlier drafts. The authors also thank seminar participants at SCANCOR at Stanford University, Solvay Brussels School of Economics and Management, Southern Denmark University, New York University Stern School of Business, the University of Strasbourg, and ETH Zurich and conference audiences at the 2022 Theoretical Organization Models Society Annual Meeting, DRUID2022, the Strategic Management Society Annual Conference 2022, the 2023 Carnegie School of Organizational Learning Conference, and the 2023 Strategy Science Conference. The authors are grateful to Felipe Csaszar and two anonymous referees for their valuable comments and suggestions. All errors remain the authors' own.

Endnotes

¹ See McNish and Silcoff (2015) as well as coverage in news outlets: <https://www.ft.com/content/184d536a-856a-11e6-a29c-6e7d9515ad15> and <https://www.theguardian.com/commentisfree/2013/nov/05/why-blackberry-failed>. As summarized in *The Guardian* in 2013, RIM had opportunities to adapt but “took the wrong direction.” Note that as we will discuss later, RIM's failure to adapt has been explained previously in “landscape terms” (Argyres et al. 2019). Our interpretation complements this explanation by explaining both the trigger for change and the tolerance for performance decrease as shaped by social comparisons. Note also that a standard “NK” explanation would not account for RIM trying to adapt in the first place.

² Google provided Android as an open-source operating system to device makers via its Open Handset Alliance.

³ We tune this salience in alternative specifications.

⁴ Of course, this representation does not accurately capture the N -dimensional task environment that we model; in Online Appendix A, we present the same transformation and an example search process in an $N = 3$ environment with terraced landscapes.

⁵ A pseudocode, a decision tree, and an example of the search criterion are provided in Online Appendix C. A replication package is available at <https://doi.org/20.500.11850/725148>.

⁶ For simplicity and comparability, we keep all model parameters constant and equal to what we call the “standard NK” model from Levinthal (1997) where possible.

⁷ Agents aim to rank at least as high as their upper reference point, which allows for ties. Prohibiting ties is possible as an alternative specification but has rather peculiar implications. For example, if all agents aim for best performance (UB = 1), this will lead to one agent occupying the global peak (the first to find it) and all others perpetually searching for a nonexistent better outcome. Moreover, if a second agent was to traverse the global peak as an “acceptable but not desirable” location, both agents would immediately consider the global peak to be below aspirations and move away from it. With few agents, the idea that peaks can be “occupied” is interesting (see, for example, Giustiziero et al. 2022). Implications of firms overtaking each other in a random-draw model (i.e., one without any peaks) have been analyzed in Denrell et al. (2023).

⁸ A similar result can be found in the model developed by Denrell et al. (2023). In a model focused on aspirations and performance-perturbing shocks, the highest long-term collective performance is reached when agents have very low aspirations. This is because only the worst-performing few agents search, and aggregate performance slowly “ratchets up” over time.

⁹ This can occur because reference points are expressed as population shares. See Online Appendix B for these results.

¹⁰ These results are shown in Online Appendix B.

¹¹ The analysis of heterogeneous populations without competition can be found in Online Appendix B.

References

- Adner R, Csaaszar FA, Zemsky PB (2014) Positioning on a multiattribute landscape. *Management Sci.* 60(11):2794–2815.
- Argote L (2013) *Organizational Learning: Creating, Retaining and Transferring Knowledge*, 2nd ed. (Kluwer Academic Publishers, Norwell, MA).
- Argyres N, Mahoney JT, Nickerson J (2019) Strategic responses to shocks: Comparative adjustment costs, transaction costs, and opportunity costs. *Strategic Management J.* 40(3):357–376.
- Arts S, Cassiman B, Hou J (2023) Position and differentiation of firms in technology space. *Management Sci.* 69(12):7253–7265.
- Audia PG, Greve HR (2006) Less likely to fail: Low performance, firm size, and factory expansion in the shipbuilding industry. *Management Sci.* 52(1):83–94.
- Audia PG, Greve HR (2021) *Organizational Learning from Performance Feedback: A Behavioral Perspective on Multiple Goals* (Cambridge University Press, Cambridge, UK).
- Baum JAC, Dahlin KB (2007) Aspiration performance and railroads' patterns of learning from train wrecks and crashes. *Organ. Sci.* 18(3):368–385.
- Baum JAC, Rowley TJ, Shipilov AV, Chuang Y-T (2005) Dancing with strangers: Aspiration performance and the search for underwriting syndicate partners. *Admin. Sci. Quart.* 50(4):536–575.
- Baumann O, Schmidt J, Stieglitz N (2019) Effective search in rugged performance landscapes: A review and outlook. *J. Management.* 45(1):285–318.
- Berchicci L, Tarakci M (2022) Aspiration formation and attention rules. *Strategic Management J.* 43(8):1575–1601.
- Billinger S, Srikanth K, Stieglitz N, Schumacher TR (2021) Exploration and exploitation in complex search tasks: How feedback influences whether and where human agents search. *Strategic Management J.* 42(2):361–385.
- Blettner DP, He ZL, Hu S, Bettis RA (2015) Adaptive aspirations and performance heterogeneity: Attention allocation among multiple reference points. *Strategic Management J.* 36(7):987–1005.
- Boyle E, Shapira Z (2012) The liability of leading: Battling aspirations and survival goals in the *Jeopardy!* Tournament of Champions. *Organ. Sci.* 23(4):1100–1113.
- Bromiley P (1991) Testing a causal model of corporate risk taking and performance. *Acad. Management J.* 34(1):37–59.
- Bromiley P, Harris JD (2014) A comparison of alternative measures of organizational aspirations. *Strategic Management J.* 35(3):338–357.
- Burgelman RA (1994) Fading memories: A process theory of strategic business exit in dynamic environments. *Admin. Sci. Quart.* 39(1):24–56.
- Chen WR, Miller KD (2007) Situational and institutional determinants of firms' R&D search intensity. *Strategic Management J.* 28(4):369–381.
- Csaaszar FA (2018) A note on how NK landscapes work. *J. Organ. Design* 7:15.
- Csaaszar FA, Levinthal DA (2016) Mental representation and the discovery of new strategies. *Strategic Management J.* 37(10):2031–2049.
- Cyert RM, March JG (1963) *A Behavioral Theory of the Firm* (Prentice-Hall, Englewood Cliffs, NJ).
- Denrell JC, Marengo L, Romagnoli M, Zeijen A (2023) Relative versus absolute aspirations. *Acad. Management Proc.* 2023(1):12231.
- Dong JQ (2021) Technological choices under uncertainty: Does organizational aspiration matter? *Strategic Management J.* 42(5):898–916.
- Eggers JP, Park KF (2018) Incumbent adaptation to technological change: The past, present, and future of research on heterogeneous incumbent response. *Acad. Management Ann.* 12(1):357–389.
- Ethiraj S, Levinthal DA (2004) Modularity and innovation in complex systems. *Management Sci.* 50(2):159–173.
- Fiengenbaum A, Hart S, Schendel D (1996) Strategic reference point theory. *Strategic Management J.* 17(3):219–235.
- Frank RH (1985) *Choosing the Right Pond: Human Behavior and the Quest for Status* (Oxford University Press, Oxford, UK).
- Gavetti G, Levinthal D (2000) Looking forward and looking backward: Cognitive and experiential search. *Admin. Sci. Quart.* 45(1):113–137.
- Gavetti G, Greve HR, Levinthal DA, Ocasio W (2012) The behavioral theory of the firm: Assessment and prospects. *Acad. Management Ann.* 6(1):1–40.
- Gimeno J, Folta TB, Cooper AC, Woo CY (1997) Survival of the fittest? Entrepreneurial human capital and the persistence of underperforming firms. *Admin. Sci. Quart.* 42(4):750–783.
- Giustiziero G, Kaul A, Martignoni D (2022) Strategic search: Organizational adaptation with competitive positioning. *Acad. Management Proc.* 2022(1):10122.
- Greve HR (1998) Performance, aspirations, and risky organizational change. *Admin. Sci. Quart.* 43(1):58–86.
- Greve HR (2003a) Investment and the behavioral theory of the firm: Evidence from shipbuilding. *Indust. Corporate Change* 12(5):1051–1076.
- Greve HR (2003b) *Organizational Learning from Performance Feedback: A Behavioral Perspective on Innovation and Change* (Cambridge University Press, Cambridge, UK).
- Hu S, Blettner D, Bettis RA (2011) Adaptive aspirations: Performance consequences of risk preferences at extremes and alternative reference groups. *Strategic Management J.* 32(13):1426–1436.
- Iyer DN, Miller KD (2008) Performance feedback, slack, and the timing of acquisitions. *Acad. Management J.* 51(4):808–822.
- Jain A, Kogut B (2014) Memory and organizational evolvability in a neutral landscape. *Organ. Sci.* 25(2):479–493.
- Kacperczyk A, Beckman CM, Moliterno TP (2015) Disentangling risk and change: Internal and external social comparison in the mutual fund industry. *Admin. Sci. Quart.* 60(2):228–262.
- Kahneman D, Tversky A (1979) Prospect theory: An analysis of decision under risk. *Econometrica* 47(2):263–292.
- Kauffman SA (1993) *The Origins of Order: Self-Organization and Selection in Evolution* (Oxford University Press, New York).

- Keil T, Posen HE, Workiewicz M (2023) Aspirations, beliefs and a new idea: Building on March's other model of performance feedback. *Acad. Management Rev.* 48(4):749–771.
- Kimura M (1983) *The Neutral Theory of Molecular Evolution* (Cambridge University Press, Cambridge, UK).
- Knudsen T (2008) Reference groups and variable risk strategies. *J. Econom. Behav. Organ.* 66(1):22–36.
- Knudsen T, Levinthal DA (2007) Two faces of search: Alternative generation and alternative evaluation. *Organ. Sci.* 18(1):39–54.
- Knudsen T, Levinthal DA, Puranam P (2019) A model is a model. *Strategy Sci.* 4(1):1–3.
- Lant TK, Shapira Z (2008) Managerial reasoning about aspirations and expectations. *J. Econom. Behav. Organ.* 66(1):60–73.
- Lant TK, Milliken FJ, Batra B (1992) The role of managerial learning and interpretation in strategic persistence and reorientation: An empirical exploration. *Strategic Management J.* 13(8):585–608.
- Lave CA, March JG (1975) *An Introduction to Models in the Social Sciences* (Harper & Row, New York).
- Lechler T, Kotiloglu S, Blettner D (2023) Self- vs. peer-directed search responses to organizational performance feedback. *Acad. Management Proc.* 2023(1):11326.
- Lenox MJ, Rockart SF, Lewin AY (2006) Interdependency, competition, and the distribution of firm and industry profits. *Management Sci.* 52(5):757–772.
- Levinthal DA (1997) Adaptation on rugged landscapes. *Management Sci.* 43(7):934–950.
- Levinthal DA (2021) *Evolutionary Processes and Organizational Adaptation: A Mendelian Perspective on Strategic Management* (Oxford University Press, Oxford, UK).
- Levinthal DA, March JG (1993) The myopia of learning. *Strategic Management J.* 14(S2):95–112.
- Levinthal DA, Marengo L (2016) Simulation modelling and business strategy research. Augier M, Teece DJ, eds. *The Palgrave Encyclopedia of Strategic Management* (Palgrave MacMillan, London), 1562–1566.
- Levinthal DA, Posen HE (2007) Myopia of selection: Does organizational adaptation limit the efficacy of population selection? *Admin. Sci. Quart.* 52(4):586–620.
- Li C, Csaszar FA (2021) Government as landscape designer: A behavioral view of industrial policy. *Strategy Sci.* 4(3):175–192.
- Lobo J, Miller JH, Fontana W (2004) Neutrality in technological landscapes. Working paper, Santa Fe Institute, Santa Fe, NM.
- Luce R (1959) *Individual Choice Behavior: A Theoretical Analysis* (Wiley, New York).
- March JG (1988) Variable risk preferences and adaptive aspirations. *J. Econom. Behav. Organ.* 9(1):5–24.
- March JG (1991) Exploration and exploitation in organizational learning. *Organ. Sci.* 2(1):71–87.
- March JG, Shapira Z (1987) Managerial perspectives on risk and risk taking. *Management Sci.* 33(11):1404–1418.
- March JG, Shapira Z (1992) Variable risk preferences and the focus of attention. *Psych. Rev.* 99(1):172–183.
- March JG, Simon HA (1958) *Organizations* (Wiley, New York).
- Massini S, Lewin AY, Greve HR (2005) Innovators and imitators: Organizational reference groups and adoption of organizational routines. *Res. Policy* 34(10):1550–1569.
- McNish J, Silcoff S (2015) *Losing the Signal: The Untold Story Behind the Extraordinary Rise and Spectacular Fall of BlackBerry* (Flatiron Books, New York).
- Menon A (2018) Bringing cognition into strategic interactions: Strategic mental models and open questions. *Strategic Management J.* 39(1):168–192.
- Menon A, Yao D (2024) Rationalizing outcomes: Interdependent learning in competitive markets. *Strategy Sci.* 9(2):97–117.
- Miller KD, Bromiley P (1990) Strategic risk and corporate performance: An analysis of alternative risk measures. *Acad. Management J.* 33(4):756–779.
- Moliterno TP, Beck N, Beckman CM, Meyer M (2014) Knowing your place: Social performance feedback in good times and bad times. *Organ. Sci.* 25(6):1684–1702.
- Nelson RR, Winter SG (1982) *An Evolutionary Theory of Economic Change* (Belknap Press of Harvard University Press, Cambridge, MA).
- Podolny JM (2018) An exploration of discerning search. *Strategy Sci.* 3(1):295–305.
- Posen HE, Levinthal DA (2012) Chasing a moving target: Exploitation and exploration in dynamic environments. *Management Sci.* 58(3):587–601.
- Posen HE, Keil E, Kim S, Meissner FD (2018) Renewing research on problemistic search—A review and research agenda. *Acad. Management Ann.* 12(1):208–251.
- Puranam P, Stieglitz N, Osman M, Pillutla MM (2015) Modelling bounded rationality in organizations: Progress and prospects. *Acad. Management Ann.* 9(1):337–392.
- Rahmandad H (2019) Interdependence, complementarity, and ruggedness of performance landscapes. *Strategy Sci.* 4(3):234–249.
- Shinkle GA (2012) Organizational aspirations, reference points, and goals: Building on the past and aiming for the future. *J. Management* 38(1):415–455.
- Siggelkow N (2002) Evolution toward fit. *Admin. Sci. Quart.* 47(1):125–159.
- Simon HA (1955) A behavioral model of rational choice. *Quart. J. Econom.* 69(1):99–118.
- Simon HA (1979) Rational decision making in business organizations. *Amer. Econom. Rev.* 69(4):439–513.
- Simon HA (1990) Invariants of human behavior. *Annual Rev. Psych.* 41(1):1–19.
- Simon HA (1996) *The Sciences of the Artificial* (MIT Press, Cambridge, MA).
- Tarakci M, Ateş NY, Floyd SW, Ahn Y, Wooldridge B (2018) Performance feedback and middle managers' divergent strategic behavior: The roles of social comparisons and organizational identification. *Strategic Management J.* 39(4):1139–1162.
- Thornton RA, Thompson P (2001) Learning from experience and learning from others: An exploration of learning and spillovers in wartime shipbuilding. *Amer. Econom. Rev.* 91(5):1350–1368.
- Tripsas M, Gavetti G (2000) Capabilities, cognition, and inertia: Evidence from digital imaging. *Strategic Management J.* 21(10–11):1147–1161.
- Wagner A (2005) Robustness, evolvability, and neutrality. *FEBS Lett.* 579(8):1772–1778.
- Wagner GP, Altenberg L (1996) Complex adaptations and the evolution of evolvability. *Evolution* 50(3):967–976.
- Wallace S, Etkin J (2024) Range goals as dual reference points. *Organ. Behav. Human Decision Processes* 184:104340.
- Yi S, Knudsen T, Becker MC (2016) Inertia in routines: A hidden source of organizational variation. *Organ. Sci.* 27(3):782–800.

Axel Zeijen is a senior scientist at ETH Zurich. He received his doctorate from ETH Zurich. His research interests include technology strategy and behavioral strategy. He leverages modeling and empirical methods to study firm adaptation and industry dynamics, mainly in the context of technological change.

Manuel Romagnoli is a postdoctoral researcher at Friedrich Schiller University Jena. He received his PhD from LUISS University, Rome. His research focuses on the consequences of behavioral decision making for firm performance, industry dynamics, and technological change. He employs formal models and empirical analyses to study issues related to innovation and organizational behavior.

Luigi Marengo is a professor of economics at LUISS University, Rome. He received his PhD from Science Policy Research Unit (SPRU), University of Sussex. His research interests focus on evolutionary and behavioral models of learning and adaptation, mainly in the context of organizational and technological change.