Received: 6 October 2017 | Revised: 27 February 2018 | Accepted: 26 April 2018 | Published on: 4 July 2018

DOI: 10.1002/smj.2918

WILEY STRATEGIC MANAGEMENT JOURNAL

RESEARCH ARTICLE

The power and limits of modularity: A replication and reconciliation

Christina Fang[1,2] | Ji-hyun (Jason) Kim[3]

[1]Department of Management, Leonard N. Stern School of Business, New York University, New York, New York

[2]Department of Management, Cheung Kong Graduate School of Business, Beijing, China

[3]School of Business, Yonsei University, Seoul, Korea

Correspondence
Christina Fang, Department of Management, Leonard N. Stern School of Business, New York University, 44 West 4th Street, New York, NY 10012.
Email: cfang@stern.nyu.edu

***Research Summary*:** We ask two questions: First, what are the underlying mechanisms that explain the power of modularity? Second, is the power of modularity robust in non-modular problems? We replicate and then reconcile the key results in two prior models on modularity: E&L and S-search. Our results yield several important insights. First, a significant portion of the advantage enjoyed by S-search is attributed to multi-bit mutation. Second, organization-evaluation needs to be used in combination with multi-bit mutation. Third, when the underlying problem structure becomes nonmodular, S-search outperforms E&L search, even though the advantage is reduced. More generally, organizational designers need to pay close attention to how different elements of modular search interact, and avoid making incremental adjustments.

***Managerial Summary*:** Modularity in product or organizational design is an approach that divides a system into smaller modules and attempts to augment the system level performance by experimenting with new modules. Because of its potential benefits such as parallel problem solving, adaptability in turbulent environment, or high speed in experimentation, both scholars and practitioners subscribed to the "power of modularity" thesis. Despite its popularity, there are significant number of cases where the superiority of modular design does not hold. We compare and contrast two representative prior studies that had different views on modeling organizational evolution under a modular design principle. By doing so, we are able to uncover what contributes to the superiority of modular design. Our results suggest that, when conducting experimentation under a modular design, it is important to (a) experiment multiple decision components simultaneously within a single module; and (b) allow evaluation of the changes to be made by the module-level manager not by the organization-level manager. When the manager does not know whether the modularity in organizational design fits with the modularity in the task, it is advised to do multiple experimentation in a single module at a time while allowing the organization-level manager to evaluate the changes.

KEYWORDS

innovation, interdependency, modularity, organizational search

1 | INTRODUCTION

The strategy field has long grappled with the complexity of strategizing in interdependent contexts (Bettis & Hitt, 1995; Ghemawat & Levinthal, 2008; Levinthal, 1997; Rivkin, 2000). The interaction among choices poses a challenge to firm strategists: How do organizations carry out strategic choices when these choices are conditioned by inherent interactions? Modularity, as a product and organizational design strategy, provides an answer (Baldwin & Clark, 2000; Brusoni, Marengo, Prencipe, & Valente, 2007). Although many definitions exist, modularity is generally acknowledged as a design principle that minimizes the interdependence between modules and maximizes the interdependence within each module. During the last decade, the concept of modularity has sparked vigorous interest in multiple disciplines as diverse as management (e.g., Schilling, 2000), engineering (e.g., Sosa, Eppinger, & Rowles, 2003), biology (e.g., Barabási & Oltvai, 2004), and psychology (e.g., Fodor, 1983). It has been applied to multiple units of analysis, ranging from products to production systems to organizations (Campagnolo & Camuffo, 2010).

A central theme in the literature is the "power of modularity" thesis, emphasizing the many advantages of modularity. The origin of the thesis can be traced to the seminal work by Simon (1962) on the architecture of complexity, in which he argued that complex systems evolve faster if they have hierarchical or "nearly decomposable" structures. These nearly decomposable structures are the discrete chunks, or modules, which form subassemblies or subsystems. Modularity localizes the impact of environmental disturbances within specific modules, thus increasing the survivability and adaptability of the system in a turbulent environment (Sanchez & Mahoney, 1996; Simon, 2002). In addition, modularity enables greater "division of cognitive labor" (Williamson, 2002), as groups and teams may work independently and do not have to be part of the same firm (Arora, Gambardella, & Rullani, 1997; Baldwin & Clark, 2003). Within the firm, modularity makes it possible to solve the problem in parallel: Work on modules can go on simultaneously without ongoing coordination among them (Baldwin & Clark, 2003; Marengo, Dosi, Legrenzi, & Pasquali, 2000; Parnas, 1972). Because experimentation takes place at the level of the modules, rather than at the entire product (Baldwin & Clark, 2000), more options are generated. This allows new products to be developed at a faster pace, as the integration of the final product is a matter of mix and match (Baldwin & Clark, 2000; Sanchez & Mahoney, 1996; Ulrich & Eppinger, 1995).

Despite the promise of modularity, both theoretical and empirical research have encountered some stumbling blocks. Empirically, in industries thought to be suitable for modularization, firms have not moved toward modularization in their product architecture. For instance, automobile firms are persistently integrated in their manufacturing operations (MacDuffie, 2008; Takeishi & Fujimoto, 2003). In other cases, the promise of modularity is heavily debated. For instance, Baldwin and Clark (2000, p. 258) documented how some firms (e.g., Weitek Corporation) tightly coupled the system components (to lower costs), while others stuck to a loosely coupled system (to avoid of conflicts) in developing the shared memory architecture used in computers. Baldwin and Clark concluded that "neither modular designs nor interdependent designs are *inherently* superior; the costs and benefits of each approach vary by case and over time" (p. 258). This clearly calls for a clearer understanding of both the underlying drivers of modularity and the boundary conditions. Theoretically, there does not seem to be widely agreed on notions of modularity across different disciplines, application areas or fields. A recent review (Salvador, 2007) uncovered 40 different definitions of product modularity. In addition, scholars only recently began to study the conditions under which the power of modularity does and does not hold. For example, the "mirroring hypothesis" (Colfer & Baldwin, 2016; MacCormack, Baldwin, & Rusnak, 2012) argues that the fit between organizational structure and product design is an important prerequisite for the power of modularity to materialize. Achieving the fit, however, requires a very precise understanding of the product functionalities, how they are allocated, and how the modules interact. Scholars have argued that the alignment among product architecture, knowledge, and task partitioning is not always a given fact (Brusoni & Prencipe, 2001; Chesbrough & Kusunoki, 2001; Hoetker, 2006; Takeishi, 2001). In fact, a significant number of studies show results inconsistent with the "mirroring" hypothesis (Colfer & Baldwin, 2016), suggesting that "it is possible for firms to strategically 'break the mirror' and perform well" (p. 721).

These difficulties suggest that there is a need to theoretically examine the mechanisms underlying the "power of modularity" thesis in detail. What exactly underlies the power of modularity? Does the power of modularity come from decentralized decision-making, parallel experimentation, or the ability to experiment with multiple decision choices? Which one of the mechanisms is critical, and which ones only work in combination with others? How sensitive is the power of modularity to nonmodular problems?

To study these questions, we build on two prior, representative models on modularity: Ethiraj and Levinthal (2004a) and Marengo et al. (2000). As we review in the next section, these two models each represents each tradition of using formal models to study and examine the "power of modularity" thesis. We find that the two traditions exhibit systematic differences in modeling choices, and these differences have important implications for reported results and for the design of modular search strategies. To understand how these differences influence the power of modularity, we first identify three systematic differences between the two traditions: (a) the centralized versus decentralized nature of decision-making; (b) the parallel versus sequential nature of modular search; and (c) the depth of search. To reconcile the two prior models, we vary those elements in an incremental manner to see which element(s) are responsible for explaining the performance differentials. Then, we test how the two models fare against each other on a nonmodular underlying problem structure.

There are three major insights we obtain from the series of experiments. First, we find that the ability to experiment with multiple decision choices is the major driver of performance differences between the two models. Second, organization-level evaluation needs to be used in combination with multi-bit search in order to attain higher performance. More generally, organizational designers are to pay close attention to how different elements of search interact, and avoid making incremental adjustments. Third, when the underlying problem structure is nonmodular, we find that S-search by Marengo et al. (2000) outperforms E&L search by Ethiraj and Levinthal (2004a), even though the advantage gets progressively reduced as the structure becomes more random. In other words, while S-search is sensitive to shifts in the underlying problem structure, E&L search is remarkably robust.

2 | LITERATURE REVIEW

Consider a firm trying to decide whether to move toward modularity in design. Under which conditions would it expect to derive the benefits frequently attributed to modularity? Should it allow decentralized decision-making at the module level or should it retain centralized decision-making authority? Should it allocate its attention sequentially or in parallel? How should it brace itself against possible misalignment between its representation and the problem at hand (which may never be known with perfect accuracy)?

We build on a stream of work that uses computer simulations and other formal methods that explore these questions. At the core of this line of inquiry is the numerical research paradigm based on the NK model (Kauffman, 1993; Levinthal, 1997), where a parameter K is used to formally capture the varying degree of interdependencies in a problem. If K is high, interactivity among decisions is high. The now classic model of Levinthal (1997) was the first to import NK models from evolutionary biology (Kauffman, 1993) to the study of management and organizations. Even though this article does not address at all issues of modularity, it provides a common foundation for two distinct "schools" that later emerged. These two "schools" differ notably in their modeling assumptions and choices: One might be called the American School (involving scholars like Ethiraj, Levinthal, Rivkin, Siggelkow, and others) and the other, the European School (involving scholars like Brusoni, Dosi, Frenken, Marengo, Prencipe, Valente, and others). Even though each tradition aims to model adaptation in social organizations vis-à-vis biological systems, they differ from each other in important ways.

In Table 1, we outline representative research from these two schools, and review their differences along three key modeling dimensions: (a) whether decisions are made at the organization (versus module) level; (b) whether attention is allocated sequentially (versus in parallel); and (c) whether to experiment with multiple decisions at any given time period. In addition, we also briefly outline their high-level research questions, respectively.

The first dimension where prior models differ is the level of evaluation whether any new initiative is evaluated at the organization (i.e., centralized) level or at the unit/module level (i.e., decentralized). Prior studies have generally assumed either one level or the other (except Rivkin and Siggelkow [2003] and Siggelkow and Levinthal [2003], which modeled both possibilities). Modularity allows a more fine-grained unit of selection (from the organization to the modules), leading to rapid search and quicker, more flexible decision-making (Mintzberg, 1979). Consistent with this, research in the American School tends to model module-level decentralized decision. In other words, if a change yields an improvement for the module, the decision-maker adopts it. Thus, each module makes its decisions independent of the interest of other modules. In contrast, research in the European School has generally assumed that new ideas generated at the lower-level modules are evaluated by a central decision-maker at the organization level. This organization-level evaluation ensures that the organization pays sufficient attention to the need for coordination (Helfat & Eisenhardt, 2004; Khandwalla, 1973). In other words, whether or not to adopt changes is determined at the organization level. Only if the new alternative yields improvement at the organization level, the organization adopts the new combination. The decision-making authority is highly centralized. In reality, firms differ dramatically with regard to the level of evaluation. As documented in Baldwin and Clark (2000), both Microsoft's "synchronize and stabilize" process for software development and Intel's "copy exactly" strategy for designing fabrication plants allow only evaluation at the organization- or system-level. In contrast, Sun Microsystem uses off-the-shelf components to make greater use of module-level evaluation (p. 272).

TABLE 1 Review of prior modeling efforts

Authors	Research question	Evaluation	Attention	Degree of search (i.e., # of flips per module)	Search mechanisms
Local search tradition (i.e., "American School")					
Rivkin and Siggelkow (2003)	How do a CEO's management style, ability, and incentive structures interact with different interaction patterns?	Module & Organization	Parallel (or simultaneous)	Mostly one-bit flip; multiple flips are possible depending on the level of ALTSUB	1. ALTSUB local alternatives are considered at the module level (local defined as in Levinthal, 1997). 2. Each manager ranks alternatives & send up P proposals. 3. CEO chooses ALTCEO random alternatives out of P and selects the best (active CEO).
Siggelkow and Levinthal (2003)	How do different organizational structures—centralized, decentralized, or a combination of the two—fare under changing environment?	Module & Organization	Parallel (or simultaneous)	One-bit flip	1. In each period, a single variable is changed in its variable sets of interest (either module or entire configuration). 2. If there is performance improvement either in the module or in the entire configuration, adopt the change.
Ethiraj and Levinthal (2004a)	What are the performance implications of over- and undermodularity?	Module	Parallel (or simultaneous)	One-bit flip (in Local Search)	1. In each period, the actors within each module search. 2. Within each module, a single decision choice is selected at random, and actors within each module evaluate the efficacy of flipping at the module. 3. If better performance, accept the change.
Ethiraj and Levinthal (2004b)	How does complexity affect the usefulness of boundedly rational design efforts?	Module	Parallel (or simultaneous)	One-bit flip (in "first-order adaptation")	Same as Ethiraj and Levinthal (2004a); labeled as "first-order adaptation."
Ethiraj, Levinthal, and Roy (2008)	What are the trade-offs between innovation benefits and imitation deterrence across different modular structures?	Module	Parallel (or simultaneous)	One-bit flip (in "first-order adaptation")	Same as Ethiraj and Levinthal (2004a).
Baumann and Siggelkow (2013)	How much of the entire complex system should be taken into consideration during a search process?	Search domain (subset of the entire system)	Only a single search domain exists	Mostly one-bit flip; multiple flips are possible depending on the level of Search Radius	1. In each period, a new alternative involving only a local change is made within the search domain. 2. The new alternative is evaluated with respect to the contribution within the search domain. 3. If better performance, accept the new alternative.
S-search tradition (i.e., "European School")					
Frenken et al. (1999)	How does the evolution of social organizations differ	Organization		Possibly multiple flips	1. Choose one module randomly. 2. Mutate multiple (possibly all) bits.

TABLE 1 (Continued)

Authors	Research question	Evaluation	Attention	Degree of search (i.e., # of flips per module)	Search mechanisms
	from the biological model of mutation and selection?		Sequential (i.e., one module at a time)		3. Test the fitness of the new (entire) string. 4. Accept the new string if higher.
Marengo et al. (2000)	What is the extent to which a complex problem can be decomposed without affecting the possibility of finding good solutions?	Organization	Sequential (i.e., one module at a time)	Possibly multiple flips	1. Choose one module randomly. 2. Mutate Z bits ($1 \leq Z \leq$ Module size). 3. Test the fitness of the new (entire) string. 4. Accept the new string if higher.
Marengo and Dosi (2005)	How do degrees of decentralization, including nearly decomposable structures, affect the possibility of generating optimal solutions?	Organization	Sequential (i.e., one module at a time)	Possibly multiple flips	Same as Marengo et al. (2000).
Brusoni et al. (2007)	How does the trade-off between speed of search and possibility of lock-in change in a complex and volatile environment?	Organization	Sequential (i.e., one module at a time)	Possibly multiple flips	Same as Marengo et al. (2000).

Second, prior work has also differed in how they incorporate the allocation of attention in organizations. One of the key benefits from modularity is greater division of labor. Search can be carried in parallel, as experimentation takes place at the level of the modules, rather than at the entire product or an organization (Baldwin & Clark, 2000; Simon, 1962). Embracing this idea, research in the American School models multiple experiments carried out in parallel: Multiple modules are activated in a single period and are allowed to experiment (see the column labeled "Attention" in Table 1). On the other hand, it is possible that organizations cannot allocate their limited attention to multiple, conflicting objectives simultaneously (Cyert & March, 1963). March (1994, p. 10) noted: "targets are considered sequentially. A satisficing search process is serial rather than parallel; things are considered one at a time—one target, one alternative, one problem."[1] This is the premise of the European School, in which only one module is allowed to experiment at any given point in time, and the organization attends to the module-level innovations in a sequential, rather than parallel, manner.

This distinction in attention allocation is aptly illustrated in the case of Oticon, a Danish hearing aids producer. In 1987, faced with a significant loss in market share, the newly appointed CEO Lars Kolind implemented organizational changes to migrate from an integrated organizational structure toward a modularized one (Puranam, Alexy, & Reitzig, 2014). For instance, even employees who were not members of the top management group were supposed to "work on at least two strategic initiatives at any given point of time" (Lovas & Ghoshal, 2000, p. 888). In short, attention allocation moves from sequential to parallel.

Finally, these prior models differ substantially on whether to experiment with multiple decisions in any given time. Scholars in the American School favor local search where only alternatives in the immediate neighborhood (i.e., one-bit/gene mutations) are examined (see the column labeled "Degree of Search" in Table 1). This is consistent with the underlying assumption that "decision makers generally act as though they assume a solution will be found in the neighborhood of a symptom" (March, 1994, p. 28). A typical implementation of this local search idea is that in each period, each module searches in parallel with other modules by generating one-bit or one-gene mutation from the status quo. A single decision is selected at random, and actors within each module evaluate the efficacy of changing that decision in that module. No organization-level evaluation is carried out as the decision-making power is decentralized. In contrast, scholars in the European School prefer more distant search (what is called "S-search" in Brusoni et al., 2007[2]). In each period, a module is randomly chosen to experiment by generating a Z-bit ($1 \leq Z \leq$ module size) mutation from the status quo, and if the resulting organizational (not module) fitness of this new string is higher than before, the new mutation is accepted. The underlying belief behind the Z-bit mutation is that social organizations "are not necessarily limited to one-gene-mutation algorithms which characterizes biological systems." (Frenken, Marengo, & Valente, 1999, p. 146).

To summarize, this prior modeling literature has made a rich and diverse set of modeling assumptions. The divergence in modeling choices makes it difficult to (a) delineate the extent to which the findings from different research are in fact consistent; and (b) consolidate the research findings and build a collective, cumulative body of knowledge. As a result, it is often difficult to pinpoint the sources of possible differences in results as well as root causes underlying the power of modularity thesis. As we previously reviewed, the two distinct schools in this tradition differ in important modeling dimensions, and these differences have interesting consequences for reported results and for the

[1]Indeed, as argued by Gavetti, Levinthal, and Ocasio (2007), a neo-Carnegie perspective calls for integrating "Simon's (1962) views on decomposability with March's perspective on sequential attention" (p. 532).

[2]While S-search had been used by European scholars previously, it was first coined by Brusoni et al.: "We can consider a search strategy that divides the N components of the configuration into modules, each containing a given number of components, say S" (2007, p. 127).

design of search strategy in general. Our goal in this article is first to explicate these differences, and second to explore how these differences may have influenced the results and their interpretations. In the section that follows, we first replicate two representative models (one from each distinguished school), and compare and contrast their modeling choices in order to better understand the crucial drivers of the power of modularity. Building on these two established models allows us to examine in detail each mechanism underlying the "power of modularity." It enables us to say more precisely, within the confines of the same framework, which mechanisms are critical to the power of modularity and which others are not.

3 | MODEL

We build on and extend two prior, representative models, namely, Ethiraj and Levinthal's (2004a) local search (E&L search hereafter) and Marengo et al.'s (2000) search strategy (S-search hereafter as it has been labeled such in Brusoni et al. [2007]); *S* represents the number of "components" (i.e., decision variables) in a module (p. 127), which are both based on variants of an NK model. Before introducing the formal components of the model, we thematically describe the parts that are common to both. We operationalize a module-based search as a search strategy guided by the module structure of an organization's representation of the problem structure—a pattern of underlying interdependency among decision variables. Corresponding to each possible configuration of decision variables is a set of payoffs, determined by the payoff generation rule similar to the family of NK models. The module structure is a representation of the underlying problem (may or may not be accurate), and is often exemplified as the organizational structure. Given an underlying problem structure with corresponding payoffs, the organization then carries out a search according to its modular representation. New alternatives are generated according to various search strategies. These alternatives are then evaluated either at the organization or module level. Consistent with prior works on modular search, either the organization or the module adopts the alternative when it yields performance improvement. Otherwise, they reject it. At the end of this process, a new cycle of search begins, whereby new alternatives are generated, evaluated, and incorporated, until the performance of the organization reaches an equilibrium. We start with a nearly modular structure, and later, introduce nonmodular underlying structures.

3.1 | Problem structure

Building on prior works, notably Simon (1962) and Ethiraj and Levinthal (2004a, 2004b), we represent a firm's decision problem as a vector of *N* decision variables, each of which takes on a binary value (i.e., 0 or 1). As in Ethiraj and Levinthal, we model the nature of interdependencies in the problem by an "interaction matrix," Figure 1 illustrates three possible interaction matrixes: (a) nearly modular, (b) chained, and (c) random. Following the convention of Ethiraj and Levinthal, in each interaction matrix, an alphanumeric notation denotes a decision variable whereas an alphabetic notation represents a module. In Figure 1, there are three modules (a, b, and c), and each module consists of five decision variables, resulting in a total of 15 decision variables (a1–a5, b1–b5, c1–c5). Thus, the total number of decision variables (denoted by *N*) is 15 and the size of module is 5. A ● symbol indicates an interdependency relationship among decision variables. Reading down each column, ● indicates that the column decision variable affects the row decision variable. In the nearly modular case (Figure 1a), there is maximum interdependency among decision variables within a given module

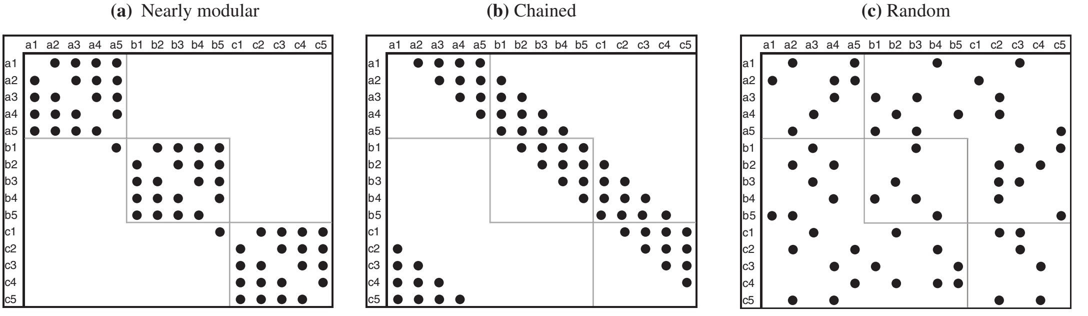


FIGURE 1 Interaction matrix of decision variables. (a) Nearly modular (b) Chained and (c) Random

(i.e., each variable affects every other variable). However, there is minimum but nonzero dependency across modules. Figure 1b represents a nonmodular chained structure, which corresponds to the typical interdependency pattern in an NK model where $N = 15$ and $K = 4$ (e.g., Levinthal, 1997). Last, another nonmodular case is characterized by random interactions, which are created by distributing ●'s randomly (Figure 1c). When we compare performances of different search strategies across different underlying structures, we control for the number of interdependencies (i.e., number of ●'s). In this way, we ensure that differences in results are driven by different interdependency patterns, rather than by different degrees of interdependencies.

Given an interaction matrix, the performance of a firm (or a module if decisions are made at the module level) is calculated as the average of each decision variable's contribution. Contribution of a variable is calculated in the standard NK way: Each variable's contribution is a function of the focal variable itself and all other K variables that influence it. How these K other variables are chosen depends on the interaction matrix as in Figure 1. Each possible combination of variables has a one-to-one mapping onto a random number drawn from the uniform distribution [0,1] identically and independently. The resulting problem structure is fixed within a simulation run, but varies across different simulation runs.

3.2 | Search strategies

Our implementation of the two search strategies is as follows. First, to replicate E&L search, in each period, a randomly chosen decision variable in each module is changed (i.e., "flipped" from 0 to 1 or from 1 to 0). Thus, the number of variables changed in a period is equal to the number of modules. Consistent with prior work (Ethiraj & Levinthal, 2004a, 2004b; Rivkin & Siggelkow, 2003), one decision-maker is assigned to each module and authorized to evaluate the performance of the new configuration. Within each module, if the new configuration yields a higher performance than the previous level, he or she adopts the change. If not, the new configuration is discarded and the module retains the current configuration. Note that (a) all the modules are activated for adaptation, (b) a single variable is flipped in each module, and (c) evaluation is made at the module-level.

Second, to replicate S-search, one of the modules is randomly chosen per period. An integer Z is then randomly chosen between 1 and the number of variables within a module (i.e., size of a module). Since all modules are of the same size, the number of variables in a module is a constant. Next, a new configuration is obtained by changing Z variables randomly. If this change results in a higher performance for the entire problem, the decision-maker accepts the new change; if not, he or she abandons it. Unlike in E&L search, a single module is activated per period, multiple variables can be flipped within a module, and evaluation is made at the organization-level.

4 | RESULTS

Figure 2 presents the average performance of the two search strategies over time in a log scale. For all the results hereafter, we report average values over 500 independent runs. In keeping with the two prior models, we first simulate an interaction matrix that is nearly modular (e.g., Figure 1a) with total number of decisions set to 15 and the module size set to 5. In order to compare the performance of the two search strategies, we need to ensure that any performance difference is not attributable to more or less extensive search per period (i.e., the total number of decision variables experimented with). With total number of variables set at 15, the number of modules 3, and the number of decision per module becomes 5 (15 divided by 3). This parameter setup controls for the total number of decisions experimented between the two models within each period. The number of decisions E&L search experiments with in each period is exactly equal to 3 (each module experiments with only one decision at a time). In the case of S-search, since only one module is activated in a single period, and a random number Z is generated between 1 and the size of the module (5), the average number of decisions experimented with in each period is $(5 + 1)/2 = 3$. Thus, both E&L and S-search vary, on average, the same number of decisions.[3]

As seen in Figure 2, while the performance of E&L search improves quickly and outperforms S-search in the first 20 periods, it stabilizes at 0.67. In contrast, S-search continues to improve and reaches a much higher equilibrium performance of 0.78. The 95% confidence intervals of each search strategy indicate that the difference in the equilibrium performance levels is statistically significant.

Which aspects of module-based search drive these differences in performance? Next, we reconcile the two models and delve more deeply into the mechanisms that may account for such differences. In particular, we examine (a) the sources of the performance difference, and (b) whether the performance of the two search strategies is robust to alternative problem structures.

4.1 | Reconciliation of the two prior models

What explains the superior performance of S-search as compared to the E&L search? As we noted earlier, the two differ in three crucial model details—(a) organization versus module evaluation; (b) sequential versus parallel attention to modules, and (c) single-bit versus multi-bit mutations. These differences in modeling choices reflect different assumptions about organizational behavior: for instance, whether decision-making is centralized and sequential or decentralized and in parallel. Our goal is to first understand which modeling choices are accountable for producing the performance differences. Our strategy is to take E&L search as the baseline and modify model components incrementally to arrive at S-search. Since the two models differ in three major aspects, there are altogether six different possible configurations in between E&L and S-search: three one-step changes (e.g., E&L search with organization-level evaluation), and three two-step changes (e.g., E&L search with organization-level evaluation AND single-module activation). We explore the performance of each one of these six configurations in separate experiments (Experiments 2 to 7). Table 2 summarizes the combinations of different modeling choices corresponding to each one of the experiments.

[3]In results not reported here, we confirmed that our implementation of E&L search exactly replicates the "5 modules" case in Figure 4 in Ethiraj and Levinthal (2004a, p. 168) under their parameter setting (i.e., N = 30 with 5 modules).

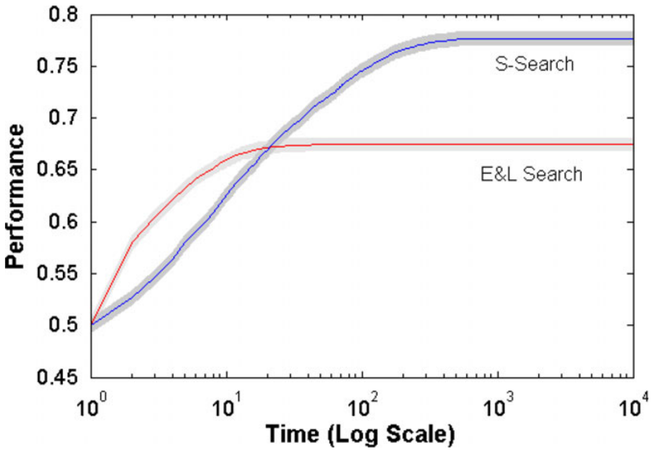


FIGURE 2 Longitudinal performance of the two search strategies. *Note.* Results are based on an average of 500 independent runs and on a nearly modular problem structure (number of decision variables = 15; module size = 5, number of modules = 3); shaded error bars represent 95% confidence intervals.

For readers interested in how we implemented each experiment, we also provide in the last column of Table 2 a visual illustration of each case.[4]

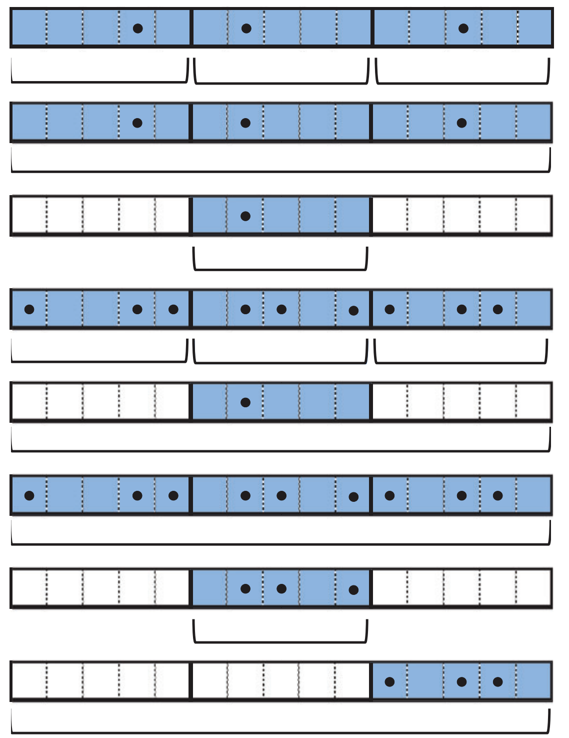
Figure 3 presents the performances of this complete series of experiments. On the vertical axis is the equilibrium performance with 95% confidence intervals. On the horizontal axis, we mark eight different configurations of search, starting with E&L search (Experiment 1) and ending with S-search (Experiment 8). Between these two are the six additional experiments, each of which varies one or two of the three model components: level of evaluation, number of activated modules per period, and number of flips in a module. In particular, Experiments 2 to 4 correspond to E&L with one-step changes (i.e., organization-level evaluation, OR single-module activation, OR multi-bit mutations, respectively). Experiments 5 to 7 correspond to E&L with two-step changes, that is, organization evaluations AND single-module activations, organization evaluation AND multi-bit mutations, and single-module activations AND multi-bit mutations. When all three model components of E&L search change, we arrive at S-search (organization evaluation AND single-module activation AND multi-bit mutations). In the following, we go through these additional experiments in sequence.

Experiment 2 implements the case where we change the module-level evaluation in E&L search to organization-level evaluation, and doing so, decreases the performance from the baseline significantly from 0.674 to 0.645. This performance decline is significant at 95% confidence level. Experiment 3 implements the case where only a single module is activated in E&L search (versus the baseline where all modules search in parallel), and this does not have significant impact on baseline performance. Most notably, the biggest lift in performance is associated with Experiment 4 where we introduce multi-bit mutations to E&L search. Performance goes up significantly from 0.674 to 0.775, almost reaching the performance level of S-search. Next, we examine the performance of two-step changes from the baseline E&L search. When we change evaluation (to organization level) and

[4]Note that in Experiments 3 and 5, a single decision is flipped in each period, whereas in Experiments 4 and 6, nine decisions are flipped on average. Throughout our analysis in this section, we have kept the underlying size and the structure of the task constant across different experiments, while allowing the number of decisions experimented with each period to vary. Our goal is to make comparisons across experiments as meaningful as possible (by changing the various model components incrementally and in combination). This is due to the fact that there is an inherent conflict between keeping constant the size of the problem and keeping constant the number of decisions experimented with each period. One cannot control both.

TABLE 2 Differences in model components in prior models and current experiments

	Level of decision-making	Activated modules	Degree of search	Visual illustration (: Activated; : Not activated; : Flipped; : Evaluation span)
Experiment 1 (E&L)	Module	Multiple	One-bit	
Experiment 2	**Organization**	Multiple	One-bit	
Experiment 3	Module	**Single**	One-bit	
Experiment 4	Module	Multiple	**Multi-bit**	
Experiment 5	**Organization**	**Single**	One-bit	
Experiment 6	**Organization**	Multiple	**Multi-bit**	
Experiment 7	Module	**Single**	**Multi-bit**	
Experiment 8 (S-Search)	**Organization**	**Single**	**Multi-bit**	



Note. From Experiments 2 to 8, components different from Ethiraj and Levinthal (2004a) are in bold and underlined. To graphically illustrate how we have implemented the models, we include one string of decisions for each experiment. A string represents a problem consisting of three modules, each of which contains five variables. Darkened modules are those that have been activated for search. Dots indicate decision/variables that have been experimented with. Thus, multi-bit mutations are represented by multiple dots that are contained within a single module, whereas single-bit mutations are represented by having only one dot within any given module. The square brackets indicate evaluation span: Organization-level evaluation is indicated with the widest possible bracket covering the entire string, whereas in module-level evaluation, square brackets only cover an individual module.

allocation of attention (to single-module activation) as in Experiment 5, performance remains the same as that in the baseline case (around 0.68). Furthermore, if we change evaluation (to organization level) and the extent of search (to multi-bit mutations) as in Experiment 6, performance does improve significantly to 0.76. However, the biggest jump in performance (to around 0.78) is attained in Experiment 7, where we allow both multi-bit mutations AND single-module activation.

To understand what modeling choices (alone or in combination) are driving the performance differences between E&L search and-S search, we compare pairs of experiments that only differ in one modeling choice. That is, we pick a pair of experiments that match on two out of the three model details (organization versus module evaluation, sequential versus parallel attention to modules, and single-bit versus multi-bit mutations). Note that because all three dimensions are varied systematically, the total number of decisions experimented on average varies across these experiments, ranging from 1 (Experiments 3 and 5) to 3 (Experiments 1, 2, 7, 8) to 9 (Experiments 4 and 6). Two modeling choices jointly determine the total number of decisions: number of activated modules and the degree of search (see Footnote 4).

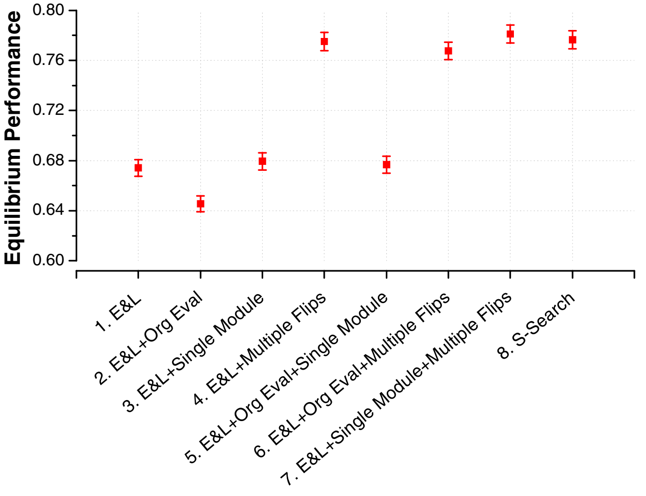


FIGURE 3 Decomposing performance differences between E&L search and S-search. *Note*. Error bars represent 95% confidence intervals

4.1.1 | Effect of organization-level evaluation

Four experiments are based on organization-level evaluation: Experiments 2, 5, 6, and 8 (See Table 2). We compare each one of these with a matching experiment that differs only in the evaluation span. By comparing the performance levels of these pairs, we examine the marginal effect of changing evaluation spans from module level to organization level. For instance, as compared to Experiment 2, Experiment 1 has the exact same setup except that it has module-level evaluation. This matching process produces four pairs of experiments that differ only in the level of performance evaluation: Experiments 1 versus 2 (total experimented variables = 3), Experiments 3 versus 5 (total experimented variables = 1), Experiments 4 versus 6 (total experimented variables = 9) and Experiments 7 and 8 (total experimented variables = 3).

Organization-level evaluation seems to have a negative or neutral impact on performance. From Experiments 1 to 2, equilibrium performance drops significantly (statistically at 95% level), whereas there is no noticeable performance change in the other three cases. The reason for this significant decline is the difficulty of implementing adaptive decisions. Under organization-level evaluation, a change that results in performance improvement can be implemented only when the overall performance changes from other modules are positive. Under module-level evaluation, the implementation decision of a module is independent of other modules. Consider a decision that improves the performance of the focal module. Since the problem structure is nearly modular and cross-module interdependencies are minimum, improvement to a module's performance tends to improve the performance of the organization. Under module-level evaluation, such a change will be implemented, independent of the changes by other modules. This allows the firm to search the landscape more extensively, and in the long run, improves its likelihood of discovering higher performance. However, when evaluation is done at the organization level, this positive performance hit may be offset by negative performance changes as a result of concurrent changes in other modules. A decision beneficial to the focal module may not be implemented if the net change (when combined with other cumulative and concurrent changes) is negative at the organization level. Evaluating performance at the organization level, therefore, imposes a more stringent criterion for implementing decisions, thwarting local adaptations that are potentially fruitful.

4.1.2 | Effect of single-module activation

Four pairs of experiments differ only in whether it is sequential (i.e., single-module activation) or parallel search (i.e., multiple module activation). In comparing the first two pairs (Experiments 1 versus 3, Experiments 4 versus 7) which share module-level evaluation, we find no performance difference. Comparing the remaining two pairs (Experiments 2 versus 5, Experiments 6 versus 8) suggests that there is a small performance boost from Experiments 2 to 5: The equilibrium performance increases from 0.65 to 0.68. From Experiments 6 to 8, it is from 0.77 to 0.78. In sum, sequential search seems to offer neutral or at best marginal performance benefits as compared to the parallel search baseline in E&L. This is perhaps not surprising. Because there is minimum interdependency across modules, working on a single module (versus many modules) may slow down the adaptation process in the short run, but in the long run, there is no performance penalty (cf. Helfat & Eisenhardt, 2004). Furthermore, coupling organization-level evaluation with single-module activation seems to counter the difficulty of implementability (e.g., from Experiments 2 to 5 and Experiments 6 to 8). To further validate this, we implemented this experiment under a perfectly modular setting where there is no intermodular interdependency. Results not reported here confirm that the two performance levels (each corresponding to parallel and sequential nature of search) under module-level evaluation (i.e., Experiments 1 versus 3, Experiments 4 versus 7) are almost identical. In this perfectly modular setup, the performance-boosting effect of sequential activation with organization-level evaluation becomes more pronounced.

4.1.3 | Effect of multi-bit search

To understand the effect of multiple flips, we again examine pairs of experiments (1 & 4, 2 & 6, 3 & 7, 5 & 8) that differ only in the degree of search (i.e., whether search is single-bit or multi-bit). In each pair, the first experiment is based on a one-bit flip and the second on multi-bit search. In all four pairs, performance improves as the number of flips changes from one-bit to multi-bit.

This result provides an important clue why S-search performs much better than E&L search. So far, we have seen that neither the level of evaluation nor the nature of search boosts performance sufficiently to bridge the gap between E&L and S-search. Why does performance improve so dramatically when multi-bit search is incorporated? Because the problem structure is nearly modular and most interdependencies are contained within modules, we can treat each module as a small "rugged landscape" or subproblem. With multi-bit search in S-search, the number of variables experimented with is randomly determined. As a result, a module may experiment by chance with many variables simultaneously. Thus, multi-bit search makes it possible to simultaneously experiment with multiple decisions across different subproblems, thus overcoming possible local peaks. Organizations may experiment with and adopt alternatives far removed from the organization's current mode of operation (March & Simon, 1958; Nelson & Winter, 1982), similar to occasional "long-jumps" (Levinthal, 1997) or "reorientation" (Tushman & Romanelli, 1985).

This beneficial effect of multi-bit search is so strong that it can overcome the difficulties created by lack of implementability we mentioned before. Recall that there is a significant performance decline from Experiments 1 to 2, yet negligible performance change from Experiments 4 to 6. Experiment 2 is associated with organization-level evaluation, whereas Experiment 6 is associated with both organization-level evaluation and multi-bit search.

4.1.4 | Summary

In a nutshell, our comparisons of the two foundational models yield a couple of insights. First, a significant portion of the performance advantage enjoyed by S-search can be attributed to a single-model component—multi-bit mutation. As seen from Figure 3, there are two distinct levels of performance—a high level around 0.76 that is associated with Experiments 4, 6, 7, and 8, and a lower level around 0.68 that is associated with Experiments 1, 2, 3, and 5. The common component in the former group is multi-bit search—that is, all four experiments incorporate multi-bit search, though in varying combination with other components. This shows unambiguously the power of expansive search: Organizations can realize performance gains if they allow multiple decisions to be experimented within a module at the same time.

Second, another element of the S-search strategy, organization evaluation, has a more nuanced impact on performance. Even though it is part of a higher performing search strategy (i.e., S-search), it alone does not yield performance benefits. As seen in Experiment 2, performance in fact declines when we introduced organization evaluation to the basic set up of E&L. The reason, as we mentioned before, is the difficulty in implementing adaptive decisions that benefit local modules to the potential detriment of others. Mandating that only decisions that lead to organization-wide improvements can be adopted may be unnecessarily cautious and overly exploitative.[5] This deficiency, however, is more than compensated for when organization evaluation is used in combination with multi-bit search. As seen in both Experiments 6 and 8, performance is much higher at around 0.76 when both are adopted (versus the baseline of E&L around 0.68). In other words, sufficient perturbations or explorations are generated whenever we allow possible within module changes that occur simultaneously, even if these changes are evaluated at the organization level.

Overall, our series of experiments clearly reveal that organizational designers need to pay close attention to how different elements of search interact and work in combination with one another. Adopting a single element or component in an incremental fashion may not be productive.

4.2 | Performance comparisons under nonmodular problem structures

So far, module-based search is guided by a modular representation structure (e.g., organizational structure), which matches near perfectly with the underlying problem structure. It is somewhat a weak test of the power of modularity because a modular search strategy is expected to do well when the underlying problem structure is in fact modular. What happens when the structure is not modular, but more or less random? In reality, the organization does not know, a priori, the exact pattern of interactions among the interdependent decisions (e.g., whether the problem is modular or not). According to the "mirroring hypothesis" (Colfer & Baldwin, 2016), any misalignment between the pattern of interdependency of the technical architecture and that of the organizational structure may create performance shortfalls.

Our analysis in this section explores this idea—whether the power of modularity is indeed sensitive to the underlying problem. We introduce two additional nonmodular problem structures: chained and random interaction structures as seen in Figure 1b,c, respectively. In both cases, each variable's contribution to payoff depends on four other decision variables, consistent with the baseline. Figure 4 reports the results under these two nonmodular problem structures. The bar charts in Figure 4 represent the cross-sectional view of longitudinal performance graphs like in Figure 2 at t = 10,000 after which there are no further performance changes.

There are several notable findings from Figure 4. First, even when problems are nonmodular, S-search consistently outperforms E&L. In other words, the benefits conferred by having multi-bit

[5]As a side note, this potential implementability challenge is a deficiency only in a model where modules cannot adopt decisions that are lower performing, which is the case here. If for some reasons (either by design or by mistake), individual modules are capable of making changes that are worse than the status quo, then organization evaluation could also be adaptive—in that it ensures that individual experiments that go astray do not result in organization-wide disasters.

mutations and single-module activations are robust to the changes in the underlying problem structure. This implies further that even if organizational designers or strategists do not have a clue of the true problem structure, they can still improve performance potentially by incorporating elements of S-search.

Second, while S-search continues to dominate E&L, its superiority is gradually reduced when the problem becomes more random. Why does this misalignment cause a performance decline in the case of S-search? As we previously mention, S-search activates only one module in any given period, thus ignoring possible cross-module interdependencies. It examines only a small part of the problem surface in any given time period. When the underlying structure is nearly or perfectly modular, ignoring cross-module interdependencies poses virtually no penalty. Furthermore, although evaluation is nominally at the organization level, search is effectively evaluated at the module level because only one module is activated in a period and there are no cross-module interdependencies. Thus, any decision beneficial to the module will be implemented. This combination of (a) ignoring cross-module interdependencies, and (b) implementing beneficial decisions at the module level helps the S-search to attain the highest performance when the underlying problem is modular. However, when the underlying problem becomes less modular and there are many more cross-module interdependencies, changes made and evaluated at the module-level can influence other modules in a negative way. Coordination among modules becomes critical: Decisions that are beneficial to one module should not be implemented if it would result in disruptions in other modules. Although evaluation is centralized, S-search does not realize the potential benefit arising from coordination at the organization level. This is because when interdependencies are spread out (e.g., in the random case), effective adaptation requires the organization to work on multiple parts simultaneously. The S-search instead focuses on a single part of the problem. As a result, its performance systematically declines as a problem becomes increasingly nonmodular.

Similar reasoning also helps us understand why E&L search is not sensitive to or even slightly performance-enhancing under the change in the problem structure. As the underlying problem becomes less modular, it becomes important to work on multiple parts of the problem. E&L search does precisely that, making single flips in multiple modules. Working on multiple parts while interdependencies are spread out allows long jumps (Levinthal, 1997). However, this increase is offset by the lack of coordination. Since whether or not to adopt changes is decided at the module level, long jumps are conducted without coordination: Decision-makers implement changes that benefit the modules only, causing unanticipated disruptions to other modules. As a result of these two countervailing forces, performance improvement from E&L search is limited under nonmodular problem structures.

Our results offer some supporting evidence of the mirroring hypothesis depending on the type of modular search implemented: Consistent with the mirroring hypothesis, S-search experiences a decline in performance as the problem becomes nonmodular while E&L search is robust to or even slightly enhanced. According to Colfer and Baldwin (2016), mirroring works because of "information hiding," where "each module in a technical system is informationally isolated from other modules within a framework of system design rules" (p. 711). Such isolation of information by design, however, results in a misalignment between technical interdependencies and organizational structure as the problem becomes increasingly nonmodular. In S-search, the search process focusses on a single part of the problem one at a time, whereas E&L search works on multiple parts of the problem. The latter therefore is able to attenuate the performance penalty associated with changes in the underlying problem.

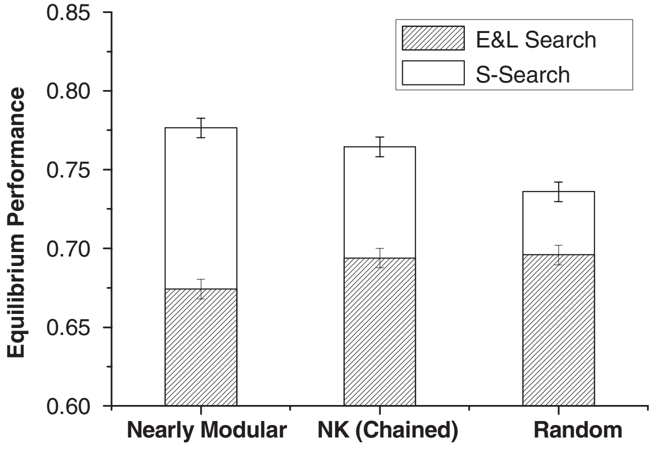


FIGURE 4 Two search strategies under alternative problem structures. *Note*. Equilibrium performances are cross-sectional performance levels at $t = 10{,}000$ after which there is no performance change. Error bars represent 95% confidence intervals

5 | CONCLUSION

We start with two simple questions: (a) What mechanisms underlie the power of modularity? and (b) Is the power of modularity robust to nonmodular problems? We first replicate and then reconcile the key results in two prior models on modularity: E&L and S-search. Our efforts to replicate and reconcile two streams of work have yielded several important insights. First, a significant portion of the performance advantage enjoyed by S-search can be attributed to a single-model component—multi-bit mutation. This implies that organizations can realize performance gains if they allow multiple decisions to be experimented within a module at the same time. Second, organization-level evaluation needs to be used in combination with multi-bit search in order to attain higher performance. More generally, organizational designers need to pay close attention to how different elements of search interact, and avoid making incremental adjustments. Third, when the underlying problem structure is nonmodular, we find that S-search continues to outperform E&L search, even though the advantage gets progressively reduced as the problem structure becomes more random. In other words, while S-search is sensitive to massive shifts in the underlying problem structure, E&L search is remarkably robust.

Our work expands significantly existing theoretical ideas about the power and limits of modularity, and clarifies the theoretical lens through which to examine the empirical phenomenon of modularity. We operationalize modularity as module-based search guided by representation (modular or not) at the organization level. By replicating two prior models that differ on important dimensions, we are able to explain, within the confines of our model, what assumptions or mechanisms are critical to the power of the modularity thesis. For instance, allowing multiple changes turns out to be the single most important factor in accounting for the differences between the two models. By reconciling these two models, we understand better the root causes of different predictions.

Second, our work provides some tentative evidence for the mirroring hypothesis (Colfer & Baldwin, 2016; MacCormack et al., 2012). While results based on the S-search are consistent with the hypothesis, results based on E&L are not. One possible boundary condition is whether attention is paid to only a small (versus multiple) part(s) of the entire problem at a time. For instance, S-search only works on a single part of the problem, and as a result, its performance gradually declines as the problem becomes less modular. As we elaborated earlier, increasing cross-module interdependencies requires a search strategy to attend to multiple parts of the problem. A module-based search that allows concurrent changes on multiple parts would be robust to changes in the problem structure (e.g., E&L search). Thus, whether the mirroring hypothesis holds depends on how module-based search is implemented.

Third, while we have focused on the mechanisms underlying two prior models, these modeling differences have clear organizational analogues and can be seen as design choices within the control of managers. By replicating and reconciling two archetypes of design choices, we have also learned something about how modular systems may be optimally managed.

To conclude, we have built on two prior models to understand the source of power of modularity and the robustness of module-based search to changes in the underlying problem. Future work, we hope, will continue to explore the boundary conditions under which the "power of modularity" thesis operates and to uncover general properties that characterize effective search strategies in interdependent problems.

ACKNOWLEDGEMENTS

For the helpful comments, the authors thank Editor Richard Bettis, several anonymous reviewers, and Jeho Lee. Ji-hyun Kim is grateful to Yonsei School of Business and YVIP (Yonsei Venture, Innovation, and Startup Program) for research funding.

REFERENCES

Arora, A., Gambardella, A., & Rullani, E. (1997). Division of labour and the locus of inventive activity. *Journal of Management & Governance*, *1*(1), 123–140.

Baldwin, C. Y., & Clark, K. B. (2000). *Design rules: The power of modularity*. Cambridge, MA: MIT Press.

Baldwin, C. Y., & Clark, K. B. (2003). *The value, costs and organizational consequences of modularity* (Harvard Business School Working Paper). Retrieved from http://www.people.hbs.edu/cbaldwin/DR1/DR1Overview.pdf

Barabási, A. L., & Oltvai, Z. N. (2004). Network biology: Understanding the cell's functional organization. *Nature Reviews Genetics*, *5*(2), 101–113.

Baumann, O., & Siggelkow, N. (2013). Dealing with complexity: Integrated vs. chunky search processes. *Organization Science*, *24*(1), 116–132.

Bettis, R. A., & Hitt, M. A. (1995). The new competitive landscape. *Strategic Management Journal*, *16*(S1), 7–19.

Brusoni, S., Marengo, L., Prencipe, A., & Valente, M. (2007). The value and costs of modularity: A problem-solving perspective. *European Management Review*, *4*(2), 121–132.

Brusoni, S., & Prencipe, A. (2001). Unpacking the black box of modularity: Technologies, products, and organizations. *Industrial and Corporate Change*, *10*(1), 179–205.

Campagnolo, D., & Camuffo, A. (2010). The concept of modularity in management studies: A literature review. *International Journal of Management Reviews*, *12*(3), 259–283.

Chesbrough, H., & Kusunoki, K. (2001). The modularity trap: Innovation, technology phase-shifts, and the resulting limits of virtual organizations. In I. Nonaka & D. Teece (Eds.), *Managing industrial knowledge: Creation, transfer and utilization* (pp. 202–230). Thousand Oaks, CA: Sage.

Colfer, L. J., & Baldwin, C. Y. (2016). The mirroring hypothesis: theory, evidence, and exceptions. *Industrial and Corporate Change*, *25*(5), 709–738.

Cyert, R. M., & March, J. G. (1963). *A behavioral theory of the firm*. Englewood Cliff, NJ: Prentice Hall.

Ethiraj, S. K., & Levinthal, D. A. (2004a). Modularity and innovation in complex systems. *Management Science*, *50*(2), 159–173.

Ethiraj, S. K., & Levinthal, D. A. (2004b). Bounded rationality and the search for organizational architecture: An evolutionary perspective on the design of organizations and their evolvability. *Administrative Science Quarterly*, *49*(3), 404–437.

Ethiraj, S. K., Levinthal, D. A., & Roy, R. R. (2008). The dual role of modularity: Innovation and imitation. *Management Science*, *54*(5), 939–955.

Fodor, J. A. (1983). *The modularity of mind*. Cambridge, MA: The MIT Press.

Frenken, K., Marengo, L., & Valente, M. (1999). Interdependencies, nearly-decomposability and adaptation. In T. Brenner (Ed.), *Computational techniques for modelling learning in economics* (pp. 145–165). Boston, MA: Springer.

Gavetti, G., Levinthal, D. A., & Ocasio, W. (2007). Neo-Carnegie School's past, present and reconstructing for the future. *Organization Science*, *18*(3), 523–536.

Ghemawat, P., & Levinthal, D. A. (2008). Choice structures and business strategy. *Management Science*, *54*, 1638–1651.

Helfat, C. E., & Eisenhardt, K. M. (2004). Inter-temporal economies of scope, organizational modularity, and the dynamics of diversification. *Strategic Management Journal*, *25*(13), 1217–1232.

Hoetker, G. (2006). Do modular products lead to modular organizations? *Strategic Management Journal*, *27*(6), 501–518.

Kauffman, S. (1993). *The origins of order: Self-Organization and selection in evolution*. New York, NY: Oxford University Press.

Khandwalla, P. N. (1973). Viable and effective organizational designs of firms. *Academy of Management Journal*, *16*(3), 481–495.

Levinthal, D. A. (1997). Adaptation on rugged landscapes. *Management Science*, *43*(7), 934–950.

Lovas, B., & Ghoshal, S. (2000). Strategy as guided evolution. *Strategic Management Journal*, *21*(9), 875–896.

MacCormack, A., Baldwin, C., & Rusnak, J. (2012). Exploring the duality between product and organizational architectures: A test of the "mirroring" hypothesis. *Research Policy*, *41*(8), 1309–1324.

MacDuffie, J. P. (2008). *Technological and organizational barriers to modularity: Persistent integrality in the global automotive industry*. Mimeo. Philadelphia: The Wharton School, University of Pennsylvania.

March, J. G. (1994). *A primer on decision making: How decisions happen*. New York, NY: Free Press.

March, J. G., & Simon, H. A. (1958). *Organizations*. New York, NY: Wiley.

Marengo, L., & Dosi, G. (2005). Division of labor, organizational coordination and market mechanisms in collective problem-solving. *Journal of Economic Behavior and Organization*, *58*, 303–326.

Marengo, L., Dosi, G., Legrenzi, P., & Pasquali, C. (2000). The structure of problem-solving knowledge and the structure of organizations. *Industrial and Corporate Change*, *9*(4), 757–788.

Mintzberg, H. (1979). *The structuring of organizations*. Englewood Cliffs, NJ: Prentice-Hall.

Nelson, R. R., & Winter, S. G. (1982). *An evolutionary theory of economic change*. Boston, MA: Belknap.

Parnas, D. L. (1972). On the criteria to be used in decomposing systems into modules. *Communications of the ACM*, *15*(12), 1053–1058.

Puranam, P., Alexy, O., & Reitzig, M. (2014). What's "new" about new forms of organizing? *Academy of Management Review*, *39*(2), 162–180.

Rivkin, J. W. (2000). Imitation of complex strategies. *Management Science*, *46*(6), 824–844.

Rivkin, J. W., & Siggelkow, N. (2003). Balancing search and stability: Interdependencies among elements organizational design. *Management Science*, *49*(3), 290–311.

Salvador, F. (2007). Toward a product system modularity construct: Literature review and reconceptualization. *IEEE Transactions on Engineering Management*, *54*(2), 219–240.

Sanchez, R., & Mahoney, J. T. (1996). Modularity, flexibility and knowledge management in product and organization design. *Strategic Management Journal*, *17*(S2), 63–76.

Schilling, M. A. (2000). Toward a general modular systems theory and its application to interfirm product modularity. *Academy of Management Review*, *25*(2), 312–334.

Siggelkow, N., & Levinthal, D. A. (2003). Temporarily divide to conquer: Centralized, decentralized, and reintegrated organizational approaches to exploration and adaptation. *Organization Science*, *14*(6), 650–669.

Simon, H. A. (1962). The architecture of complexity. *Proceedings of the American Philosophical Society*, *106*(6), 467–482.

Simon, H. A. (2002). Near decomposability and the speed of evolution. *Industrial and Corporate Change*, *11*(3), 587–599.

Sosa, M. E., Eppinger, S. D., & Rowles, C. M. (2003). Identifying modular and integrative systems and their impact on design team interactions. *Journal of Mechanical Design*, *125*(2), 240–253.

Takeishi, A. (2001). Bridging inter- and intra-firm boundaries: Management of supplier involvement in automobile product development. *Strategic Management Journal*, *22*(5), 403–433.

Takeishi, A., & Fujimoto, T. (2003). Modularization in the car industry: Interlinked multiple hierarchies of product, production and supplier systems. In A. Prencipe, A. Davies, & M. Hobday (Eds.), *The business of systems integration* (pp. 254–278). Oxford, England: Oxford University Press.

Tushman, M. L., & Romanelli, E. (1985). Organizational evolution: A metamorphosis model of convergence and reorientation. *Research in Organizational Behavior*, *7*, 171–222.

Ulrich, K., & Eppinger, S. (1995). *Product design and development*. New York, NY: Irwin/McGraw-Hill.

Williamson, O. E. (2002). Empirical microeconomics: Another perspective. In M. Augier & J. G. March (Eds.), *The economics of choice, change and organization: Essays in memory of Richard M. Cyert* (pp. 419–449). Cheltenham, U.K.: Edward Elgar.

How to cite this article: Fang C, Kim J-h (Jason) The power and limits of modularity: A replication and reconciliation. *Strat Mgmt J*. 2018;39:2547–2565. https://doi.org/10.1002/smj.2918