The Network Structure of Exploration and Exploitation

David Lazer
Allan Friedman
Harvard University

Whether as team members brainstorming or cultures experimenting with new technologies, problem solvers communicate and share ideas. This paper examines how the structure of communication networks among actors can affect system-level performance. We present an agent-based computer simulation model of information sharing in which the less successful emulate the more successful. Results suggest that when agents are dealing with a complex problem, the more efficient the network at disseminating information, the better the short-run but the lower the long-run performance of the system. The dynamic underlying this result is that an inefficient network maintains diversity in the system and is thus better for exploration than an efficient network, supporting a more thorough search for solutions in the long run. For intermediate time frames, there is an inverted-U relationship between connectedness and performance, in which both poorly and well-connected systems perform badly, and moderately connected systems perform best. This curvilinear relationship between connectivity and group performance can be seen in several diverse instances of organizational and social behavior.•

We live in a smaller world today than ever before, with more distant linkages that rapidly spread information from one corner of the globe to the other. Perhaps as a result, the last decade has witnessed a massive surge of interest in networks, both among academics (Borgatti and Foster, 2003) and in popular culture. Services to increase the efficiency with which we exploit our personal networks have proliferated, and armies of consultants have emerged to improve the efficiency of organizational networks. Implicit in this discourse is the idea that the more connected we are, the better: silos are to be eliminated, and the boundaryless organization is the wave of the future. Technologies that enable distant actors to access each other's knowledge should be utilized (Wenger, McDermott, and Snyder, 2002). Even the most hierarchical of organizations, the military, is ostensibly shifting to a more flexible, network-based system (Arquilla and Ronfeldt, 2001), with bottom-up, boundary-spanning, Web-based knowledge systems (Baum, 2005). Despite these evident secular trends, popular attention, and prescribed organizational network strategies, relatively little attention has been focused on the question of how network structures affect overall system performance and, in particular, collective problem solving—the pooling of individual abilities to solve a problem.

In contrast, there is ample research that highlights the advantage that network position can provide those who occupy them. Diffusion of innovation research (Rogers, 2003) illustrates how the well connected tend to be early adopters. Acquiring information quickly, in turn, provides advantages to individuals. Thus, for example, Granovetter's (1973) landmark research highlights the role of relationships in funneling information to individuals about employment opportunities, and Burt's (1995) research focuses on the competitive advantage that individuals or businesses gain if their network connections bridge otherwise unconnected groups. These individual-level results, however, do not necessarily translate to the

0001-8392/07/5204-0667/$3.00.

•

We gratefully acknowledge useful comments from Robert Axelrod, Beau Kilmer, Michael Macy, Robert Putnam, Adam Thomas, and Ezra Zuckerman, as well as comments from three anonymous reviewers. All faults in this paper are ours alone. The authors also acknowledge support from NSF grant #0131923. Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the NSF.

system level, because many of the individual-level results may reflect competitive advantages in zero-sum games such as finding a job. There are some examples of network research that focus on the system level, perhaps beginning with the research out of the Small Group Network Laboratory at MIT in the 1950s, where Bavelas and his colleagues examined how network structure affected the aggregation of signals (see review by Leavitt, 1962), finding, for example, that centralized networks are good at coordination for simple problems, and decentralized networks for complicated problems. This vein of research on small groups has also been followed up more recently with findings suggesting that denser ties among group members are related to success (e.g., Reagans and Zuckerman, 2001; Uzzi and Spiro, 2005; meta analysis by Balkundi and Harrison, 2006).

From a more macro perspective, Granovetter (1973), though primarily focusing on individual ties, presented a comparison of two communities, one with more casual social interactions than the other, and argued that the former is better equipped to overcome collective action problems. Putnam (1993) found a strong relationship between associational affiliations such as participation in bowling leagues or parent-teacher associations, presumably correlated with the density of societal networks, and government effectiveness. Putnam (2000) went on to document the decline in such affiliations in the U.S., cautioning about the social and economic declines that could come from the erosion of those connections. Of course, there are many processes that govern the relationship between interconnectedness and system performance, contingent on the nature of interdependence. Our first step therefore is to just take a narrow, but important, slice of this phenomenon, which we label "parallel problem solving."

NETWORK CONFIGURATION, INFORMATION DIFFUSION, AND SYSTEMIC PERFORMANCE

Parallel Problem Solving

Consider the situation of the manager of a research and development lab, who needs his or her engineers to solve some complex problem. This problem has many plausible solutions, although it is difficult at the outset to judge which approach will yield good results. One of the challenges confronting the manager is how to structure the communication among his or her engineers. Would it be wise to have high frequency meetings, so that engineers who were developing promising approaches could share their ideas with others? Or would it be better to let things just slowly and inefficiently diffuse? This scenario is an example of what we label parallel problem solving, in which a set of roughly equivalent agents are all seeking to solve the same problem, and the success of any one or subset of agents has no direct effect on other agents. In game theory terms, one might envision each agent playing a game against nature, in which an individual's performance against nature has no bearing on other players' payoffs, and players can learn from each other about what does and does not work.

This is clearly an ideal type, but there are many phenomena that approximate this ideal—whether it is professionals such as doctors (Coleman, Katz, and Menzel, 1957) and professors (Mergel, 2005) dealing with similar problems, state governments formulating public policy (Walker, 1969), brainstorming sessions (Paulus, Larey, and Ortega, 1995), or infantry in Iraq evaluating tactics on the ground (Baum, 2005). In each case, the agents in question are all wrestling with approximately the same problem and are in a position to learn from each other's actions, yet the success or failure of each is minimally affected by the performance of other agents. None of these is an exact fit to the model, of course. Rarely would agents in the real world be solving exactly the same problem or have no interdependence of payoffs. Further, there are many ways that networks might affect the success or failure of systems other than how they alter the dynamics of searching for solutions. For example, much of the literature on social capital focuses on how networks limit opportunistic behavior (Bourdieu, 1980; Coleman, 1988). Still, any problem for which agents have potentially similar means of reaching successful outcomes might be a match with this model if it is complex enough to defy quick individually obtained solutions.

A complex problem is one in which potential solutions have many dimensions that may be synergistic with respect to their impact on performance (Simon, 1962; Levinthal and March, 1981; Levinthal, 1997; Siggelkow and Levinthal, 2003). Potential solutions may be viewed as a problem space in which each solution has a performance score associated with it. Solutions involve the conjunction of multiple activities, in which the impact of one dimension on performance is contingent on the value of other dimensions. One might imagine, for example, that activities A, B, and C each actually hurt performance unless all are performed simultaneously, in which case performance improves dramatically. The presence of such synergies produces local optima, such that any incremental change results in the deterioration of performance, but some large change could produce an improvement. In this example, if someone starts out not doing any of A, B, or C, the only way to improve is to do all three simultaneously.

Actors facing complex problem spaces are often further hampered by a search process that is myopic, in that they are limited to some neighborhood of their current solution. Obviously, if agents could easily survey all possible solutions, identifying the best one would be trivial. When considering a set of options as the problem space, the myopic constraint could be seen as a limit on how many dimensions of their solution agents can manipulate at any one time to seek a superior outcome. A common assumption is that of incremental search, which sets this limit to a single dimension (e.g., March, 1991; Levinthal, 1997). In simulation models, myopic search is typically implemented through an assumption that agents can only make incremental changes to their status quo activities.

A problem space with many local optima is referred to as a "rugged" problem space, because the series of incremental changes needed to reach a better point typically requires moving through solutions with worse outcomes. Rugged problem spaces are therefore hard to search because it is easy to get stuck in a local optimum. For example, if one assumes that people can only clearly evaluate one change in their current set of activities at a time, in the preceding example, they might get stuck at not doing A, B, and C. Further, if one assumes that synergies among human activities are endemic and that people have many choices of activities, the resulting problem space that people confront will likely be quite vast and rugged. Realistically, people might anticipate the consequences of changing more than one activity with some accuracy, but the capacity to predict the consequences of changing many activities is very limited. Without the ability to forecast a wide range of solutions, problem solvers can only extrapolate the marginal improvements of small changes.

What is of primary interest in understanding parallel problem solving is not so much how individuals solve problems by themselves as much as how individuals solve problems collectively. We assume that individuals affect collective success through a network of peers, in which the activities of a particular individual offer insight to others about the configuration of the problem space. If I see someone with a different solution and with performance superior to mine, I will make that non-incremental change. In solving a complex problem, the flow of information among the individuals is likely to affect their collective performance, as each individual can look at what some of the other individuals are doing and how they are performing. But the network structure linking those individuals will determine who has access to what information and therefore how well information about the solution to the problem is disseminated in the system.

Disseminating Solutions

The assumptions underlying the idea of networked parallel problem solvers have much in common with the implicit assumptions of the diffusion and information cascade literatures. Like the organizational network structure literature above, these literatures generally assume that individuals gain information from their social environment, but they explicitly include a time dimension, something often lacking in the current organizational theory on networks (Katz et al., 2004; Balkundi and Harrison, 2006). The diffusion of innovation literature (Rogers, 2003), for example, largely focuses on the diffusion of practices that improve the well-being of agents, and the network is one of the key channels through which agents get information about what works. Although this literature looks at both individual-level (e.g., beliefs) and network-level (e.g., who talks with whom) correlates of adoption decisions, it does not examine the systemic consequences for success of how networks tie together individual cognitions, because access to information is seen as the dominant factor for beneficial innovations (Wejnert, 2002). When an innovation is of unknown quality, social contacts become more important, as imitation is the sole channel for potential improvement (Granovetter, 1978; DiMaggio and Powell, 1983). The effect of network structure varies depending on whether the change to be adopted is clearly beneficial or ambiguous (Abrahamson and Rosenkopf, 1997; Gibbons, 2004). Social diffusion models have been extensively studied, both analytically (reviewed in Mahajan, Muller, and Bass, 1990) and through simulations (Granovetter, 1978; Carley, 1991; Krackhardt, 2001).

The information cascade literature (Banerjee, 1992; Bikhchandani, Hirshleifer, and Welch, 1992) highlights how non-welfare-improving practices (fads) can spread through the system under the assumption that people observe adoption decisions but not the success of actors. Information cascades occur when an adoption sends a signal to other actors that they should adopt, leading to self-reinforcing processes, reflected in phenomena like stock bubbles and bank runs. Strang and Macy (2001) extended this essential finding to the circumstance in which actors can observe the success of other actors but not the reason for that success. The information cascade literature highlights the more general issue of information aggregation within networks, how well the system pools together unique signals. Unlike the diffusion literature, however, variation in network structure has not been deeply integrated into the information cascade research. Work at the intersection of the network and cascade literatures has also begun to examine how signals propagate through networks (e.g., Watts, 2002; Bettencourt, 2003).

The critical difference between the dominant approaches in the diffusion and cascade literatures and the parallel problem solving perspective we present here is in assumptions about how actors interact with the environment. The diffusion and cascade literatures generally focus on the dissemination of signals (or adoptions) through a population. Parallel problem solving is based on the assumption that agents are purposively seeking out novel signals directly from the environment at the same time that those signals are propagating through the network. Thus, at any given point in time, every agent is engaged in some potentially novel activity vis-à-vis nature and is receiving some payoff (and thus information), while existing in a social environment of information sharing. Some diffusion research does explicitly examine individual learning in a structural context (e.g., Carley and Hill, 2001; Ashworth and Carley, 2006), but this paper attempts to present an organizational learning model that builds on the emergent behavior of individual actors.

Following from March's (1991) computational work on organizational learning, the issue of parallel problem solving can be seen as a balancing of exploration and exploitation. In reviewing the myriad of definitions and applications of exploration and exploitation in recent literature, Gupta, Smith, and Shalley (2006) noted that exploration generally involves attempts to introduce new information, while exploitation leverages existing knowledge for some productive end. For the purposes of this research, we view exploration as the development of novel solutions and exploitation as the utilization of known solutions. In parallel problem solving, each agent may either choose to try a novel solution (exploration) or emulate another actor with whom he or she has a connection (exploitation). The question we examine is how the structure of the social network affects the systemic balance between exploration and exploitation. March (1991) implicitly assumed a hub-spoke communication structure, in which individuals learn from each other, always mediated through the construct of shared knowledge referred to as the organizational code. Miller, Zhao, and Calanton (2006) extended March's model by adding an interpersonal learning component on a fixed grid. Interpersonal learning amplifies the tradeoffs between exploration and exploitation, and learning too fast from either the organizational code or other actors can reduce total system knowledge. In this paper, we decentralize the organizational code, assuming that individuals learn directly from each other via their network of communication, or from the environment itself, rather than learning from an organizational hub. Our research design implemented this network of emulation in a simulation model.

RESEARCH DESIGN

Simulation through computational modeling is an especially powerful theory-building tool in the study of systems, allowing researchers to examine inductively the impact of manipulating features of systems. Analytic modeling often cannot handle the combinatorics of system dynamics. Empirical analysis, though essential, is constrained by the expense and practical challenges of studying real-world systems. Agent-based modeling, a particular type of simulation modeling, is especially well suited to the study of organizational systems (Chang and Harrington, 2006). Agent-based modeling starts with a set of simple assumptions about agents' behavior and inductively derives the emergent system-level behaviors that follow (Schelling, 1978; Macy and Willer, 2002). It has begun to be widely used in the study of human systems (e.g., Levinthal, 1997; Chang and Harrington, 2005; Harrison and Carroll, 2006) and includes a growing vein of research that explores the relationship between structure and performance (e.g., Levinthal, 1997; Kollman, Miller, and Page, 2000; Krackhardt, 2001; Rivkin and Siggelkow, 2003; Chang and Harrington, 2005).

Our objective in the simulation described here was to develop novel propositions about collective human behavior. There are four critical criteria by which to judge simulation-based research aimed at theory building. The first requirement of a model is verisimilitude. Research based on formal representations of human behavior needs to convince readers that the necessarily constrained assumptions in a model somehow capture the essence of some empirically relevant set of circumstances. For example, if one does not believe that the two-actor, two-choice set up of the prisoner's dilemma captures important features of some human relations, then the rather plentiful formal (analytic and simulation) research on the prisoner's dilemma has nothing to say about collective human behavior. Second, the model must be robust. All formal representations of human behavior require arbitrarily precise assumptions, for example, about functional relationships, and, for simulations, particular numerical values for parameters. Results should withstand unimportant changes to the model. A corollary to this point is that models need to be relatively simple. It becomes exponentially more difficult to assess robustness as the number of parameters increases. These first two criteria together represent the fundamental challenge to researchers using simulation methodology: how to achieve some reasonable level of verisimilitude in the model (which creates pressures to increase the complexity of the model) without undermining the robustness of the results. Third, it should be possible to fully replicate the results of the model based on the information provided in a publication. Further, commented code from simulations should be made available online.[1] Fourth and finally, the simulation should produce non-trivial, non-obvious results. The objective of computational modeling should not be to reproduce what could be derived from verbal exposition; rather, it is to produce novel yet convincing insight.

These criteria are especially, but not uniquely, relevant to simulation-based research. For example, the issue of model verisimilitude also applies to experimental research. If a particular experimental paradigm involving human subjects does not capture some key features of the real world, then any results based on that paradigm would be uninteresting. More generally, all social science research must do some violence to reality in order to reveal simple truths. Similarly, robustness is a serious issue with respect to quantitative empirical social science, which requires many arbitrary decisions with respect to constructing variables and specifying statistical models. A positive result that disappears when the underlying statistical model is changed in minor ways is also not robust.

Our research question was how the communication pattern among actors engaged in the broad class of collective human behaviors conceived as parallel problem solving affects their collective performance. Our objective, therefore, was to produce a formal representation of parallel problem solving that was rich enough to provide relevant insight, but simple enough to be transparent (criteria one and two above). Three key issues had to be specified in the model: (1) What does the problem space look like? (2) How do actors (individuals) make decisions? and (3) How do actors communicate?

The Problem Space

As noted above, synergies among human activities are inherent in human decision making, yielding problem spaces with many local optima. We therefore needed to produce an arbitrarily large number of statistically identical "problems" for the simulated agents to solve. There are some standard complex problems that computer scientists use to test search algorithms, such as combinatorial optimization problems like the traveling salesman problem (Lawler, 1985), but we chose to follow Levinthal (1997), Gavetti and Levinthal (2000), Rivkin (2000), Carroll and Burton (2000), Ahouse et al. (1991), and several others in using the NK model to produce numerical problem spaces. The NK problem space is named for the two parameters that are used to randomly generate problem spaces. It was originally developed by evolutionary biologist Stuart Kauffman (1995) to model epistasis, the genetic analog to synergies among human activities. Epistasis is the interaction of genetic traits to produce fitness (the tendency to survive and reproduce) that is greater (or lesser) than the sum of the contributions of the individual traits. For example, traits A, B, and C may only be helpful to fitness if they are all

1
We have posted documented code for examination and extension at http://code.google.com/p/parallelproblemsolving/.

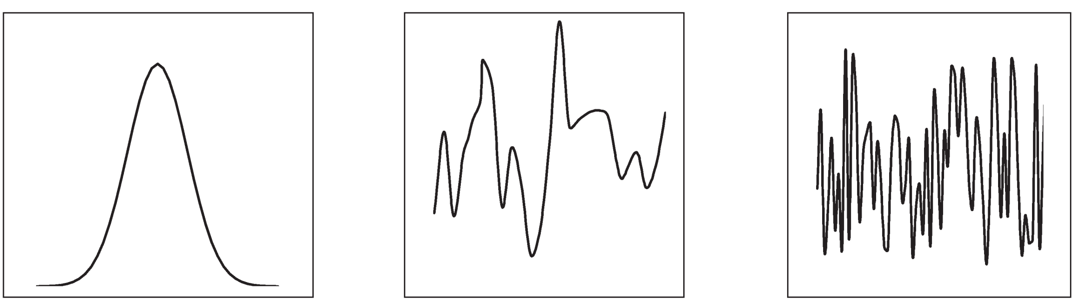
present simultaneously. In Kauffman's construction, N is the number of traits, and K is the degree of epistasis. For our purposes, N may be reinterpreted as the number of potential human activities, and each activity may be present (= 1) or absent (= 0). K is the typical amount of synergies among those activities. For example, for K = 3, the value of any given activity is contingent on the presence or absence of three other activities.

Appendix A provides a full explanation of how NK landscapes are generated. It is not possible to draw an NK landscape because it is N dimensional, but figure 1 offers a stylized representation of the impact of manipulating K. K = 0 creates a simple problem space with a single optimum. K = N – 1 creates a maximally rugged landscape, in which the performance of any given solution in the space offers no signal as to the quality of adjacent solutions because changing a single activity will change the marginal contribution of every other activity. Of greater interest is the universe of spaces in between, in which there are multiple local optima but the quality of adjacent solutions is correlated. There is no best way to find the optimal point in an unknown rugged space, but a local maximum can be easily found by searching "uphill." Moreover, there is no mechanism to prove whether or not a local maximum is the global maximum without measuring every other peak. A critical feature of any rugged space is that to move from one optimum to another incrementally, an agent must go "downhill" (do worse) before ascending to a new peak.

For most of the computational experiments below, we used a problem space that qualitatively corresponds to that in figure 1, B, which we believe captures the essence of most interesting problems that individuals and organizations in the real world face—rugged, but not chaotic. The success of an actor is defined by his or her solution. Because most randomly created solutions in the real world are bad relative to the optimum, we skewed most scores down with a monotonic function, described more fully in Appendix A.

From our perspective, the NK problem space specification has the advantage that (1) it maps reasonably intuitively to the idea that there are synergies among human activities,

Figure 1. Stylized representations of problem spaces represented by k.



A: A simple problem space, similar to K = 0.

B: A complex rugged space with local maxima and minima, similar to 0 < K < N – 1.

C: A chaotic space in which the value of every point is independent of adjacent points, similar to K = N – 1.

(2) we can produce an arbitrary number of statistically similar problem spaces from these two parameters, and (3) we can choose how large and rugged the space is through the two parameters. We have also replicated the key results reported below for the traveling salesman problem, another family of rugged problem spaces.[2]

Behavioral Rules

We assumed that at any given point in time, each actor has a solution in the NK space, so, for example, if N = 5, one solution would be in the form of the bit-string 00000, another 00001, and so on, through to 11111. Each actor also has a set of alters to which he or she is connected, forming a network. Following from the preceding discussion on human decision making and emulation, we assumed that actors are myopic, unable to directly evaluate potential solutions that deviate greatly from their status quo solution. A key exception to this myopia is that actors can see the performance level of those agents to whom they are connected. We assumed that in each round of the simulation, actors' decisions proceed in two stages. In the first stage, each actor evaluates whether anyone to whom he or she is connected has a superior solution. If so, he or she copies the most successful individual to whom he or she is connected. If no one is more successful, then the agent myopically searches for a better strategy than he or she has in the status quo strategy. To capture this myopia, we assumed that agents examine the impact of randomly changing one digit of their status quo solution, and if that potential change offers an improvement, they change their solution. Of course, people can do better than simply randomly tweaking existing strategies, but we were trying to balance the level of myopia with the magnitude of the space being searched. In real human problem solving, though there may be less myopia, there is also a vastly greater problem space to be searched. In some of the simulations reported below, we varied this simple behavioral assumption in two ways. First, we varied how often agents looked at the solutions of others. In the base model, actors look at others they communicate with every round, but in some of the computational experiments below, we assumed that agents look around less frequently, for example, every other round, every fifth round, or every tenth round. Because this slows the spread of the flow of information through a network, we label this frequency parameter "velocity," which is simply the probability that an agent will look at his or her network each round. This is related to March's (1991) "socialization rate" in an organization. March examined how rapidly individuals synchronize their views of reality with the rest of the organization, finding that slowing synchronization delays the convergence on an equilibrium and raises the long-run performance of the organization. Following from March's framework, this parameter could also be seen as capturing individual preferences for exploration over exploitation.

Second, we varied how accurately agents copy the solutions of others. A key mechanism for adaptation in systems is the mixing of strategies, in which taking a piece of A's strategy and mixing it with B's strategy could result in something that is better than both A's and B's strategies. In biological

2
Results are available upon request.

systems this is called crossover, reflecting the fact that offspring receive genetic material from both parents. Genetic programming also uses crossover, which can greatly improve search (Mitchell, 1996). We label this parameter the error rate, because it is implemented as the probability of accurately copying another actor's strategy.

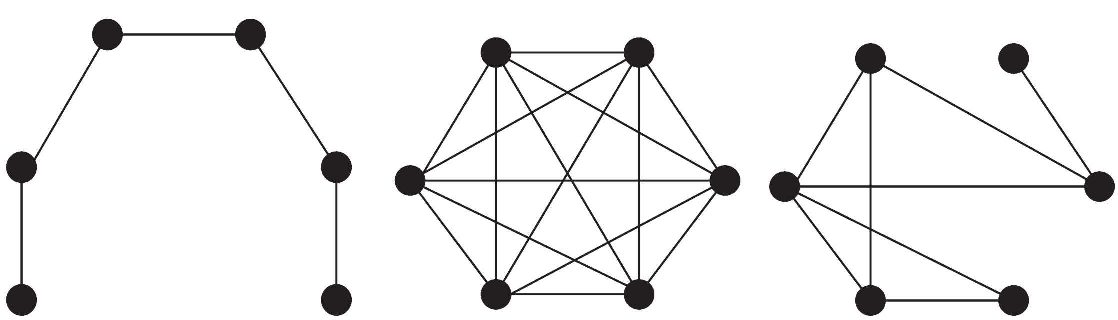
In short, actors will mimic other successful actors, and when there is no one to mimic, they will attempt to adapt. New successful adaptations will subsequently be copied by others, and so on. If no actor (or subset of actors) is isolated, then all will eventually converge on the same solution. These behavioral rules qualitatively capture the two key elements of human decision making that we discussed in the preceding section: (1) search capacity is very limited compared with the magnitude and complexity of real-world problems, and (2) emulating successful others provides a key way of finding short cuts in the problem space.

Network Configuration

The central question we examined is the impact on the performance of the system of changing the network structure, given the aforementioned search behaviors of the constituent actors. The actors are the nodes of these networks. We limited our initial focus to networks with a single component (a set of nodes containing some path between every pair of nodes). In the simulations reported, we examined four archetypical networks: a linear network, a totally connected network, a variety of random networks, and a variety of small-world networks. We assumed in all of these networks that communication is two way. Figure 2 represents the first three types of networks graphically.

A linear network, shown in figure 2, A, is simply a set of nodes in which each node, except for two, communicates with two other nodes, and the nodes and their relationships are arrayed linearly. A linear network has the minimum number of links possible in a single-component network and produces the maximum degree of separation between an average pair of nodes. A totally connected network (figure 2, B) is one in which every node communicates with every other node. A random network (Erdos and Renyi, 1959) is defined by a stochastic generation process. Each node in the network has a probability p of being connected to each other node. When p is 1, the random network is identical to the full net-

Figure 2. Graphic representation of the three network types.

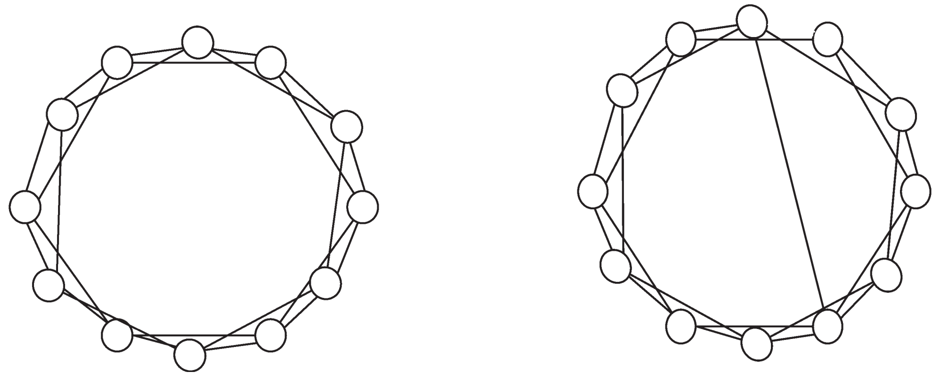


A: Linear network B: Totally connected network C: Random network (p = 0.5) work; when p is 0, there are no network ties. In the simulations, we examined the impact of varying p to test the effect of increasing network ties. Figure 2, C offers an illustration of a randomly generated network.

To examine the importance of average path length in a network, we used Watts and Strogatz's (1998) small-world model. A small-world network is one that is highly clustered (friends of friends tend to talk to each other) but still characterized by short path distances between any two actors. Figure 3 shows how a lattice network structure can be rewired to shorten path distances. By rewiring a defined lattice in which actors are arrayed in a circle, and each actor talks to his or her immediate two neighbors on either side, in this case (see figure 3, A), the number of links is held constant, and thus density is held constant. A lattice tends not to be a small-world network because, relative to the size of the system, there may be large path distances between two nodes. Watts and Strogatz's (1998) key insight was that a fairly modest random rewiring of a lattice can dramatically lower typical path distances. As figure 3, B illustrates, switching a local link to a long-distance link can significantly lower the average path distance. More generally, as the probability of rewiring increases, the average path distance between two nodes will drop rapidly, making the network into a small-world network.

There are many descriptive tools available to describe and compare networks (cf. Wasserman and Faust, 1994). This paper focuses on two: density and average path length. Density is simply the proportion of potential ties that actually exist. Average path length is the average number of steps it takes to go from one node to another.

Figure 3. Graphic representation of rewiring a highly clustered lattice structure to shorten path distance.



A: A regular lattice with each node connected to its 4 closest neighbors, with an average path length between any two nodes of 1.909.

B: The lattice with one link rewired to connect to a distant node, shortening the average path length to 1.833. In a 100-node lattice, a single rewired link can reduce average distance by over 20%.

Data Generation and Analysis

The core model is thus fairly simple—three parameters plus variation in the network structure—but still captures some of the essential features of collective problem solving. For each simulation, we assumed a population of 100 actors who interact in discrete time steps. The initial solutions of the actors are randomly generated, and those actors are randomly placed in networks with preexisting configurations. For the NK model, we assumed N = 20, and K = 5, except where noted below. When N = 20, there are 1,048,576 possible solutions for each space. The interlinking factor of 5 produces a space that is moderately rugged, with a few hundred local optima but in which the quality of proximate solutions is highly correlated. Reducing the space size or complexity reduces search time but does not alter the key findings reported below.

For each simulation, the population was run to the point that it converged on a single solution. For each network and set of parameter values, the simulation was run on the same 1000 NK spaces, and we report the average performance across the set of simulations, measured by the average performance of actors as a ratio of the global optimum (see Appendix A for details).

Our method of inference was experimental, in that we manipulated various features of the model and examined their impact on systemic outcomes (in this case, average performance). In the analyses summarized below, we varied the configuration and processes of the network and tested the robustness of these findings. For each set of simulations with population = 100, N = 20, and K = 5, we used the same random seed of 1000 starting points of solutions on the same NK spaces. These simulations may thus be viewed as 1000 "contests"—given exactly the same problem and the same population starting point, which way of linking actors will result in the discovery of the best solutions? We compared the performance over time of each of the network archetypes summarized above and explored how varying the actors' patterns of copying interacts with the network structure to determine group performance. Appendix B offers a pairwise comparison of the performance of a sample of the networks we examine below.

RESULTS

Linear vs. Totally Connected Networks

Linear and totally connected networks are extreme opposite archetypical networks. A single-component network cannot have a greater average path length than a linear network, nor a smaller average path length than a totally connected network. Figure 4 plots the average performance over 1000 simulations of linear networks versus totally connected networks over time.

Figure 4 demonstrates a clear pattern: the totally connected network finds a good solution quickly and outperforms the linear network in the short run, but in the long run, the linear network performs significantly better.[3] That is, having fewer structural opportunities for communication improves long-term social outcomes. The reason for this pattern is captured in figure 5, which plots the average number of unique solutions in the system over time in the linear and totally connected networks. As the figure shows, the number of unique solutions held by agents plummeted in the totally connected networks, while the linear network maintained diversity longer.

3
We also examined the hub-spoke network (99 agents connected to a single hub agent) and "caveman network" (Watts, 1999), with communities of five totally connected nodes separated by five intervening nodes. The hub-spoke network (average degree of separation of 2) performed just slightly worse than the totally connected network in the short run and slightly better in the long run. The caveman network (average degree of separation of 17.98) performed slightly better than the linear network in the short run and slightly worse in the long run. Results are available from the authors upon request.

Figure 4. Average performance over 1000 simulations of linear versus totally connected networks.

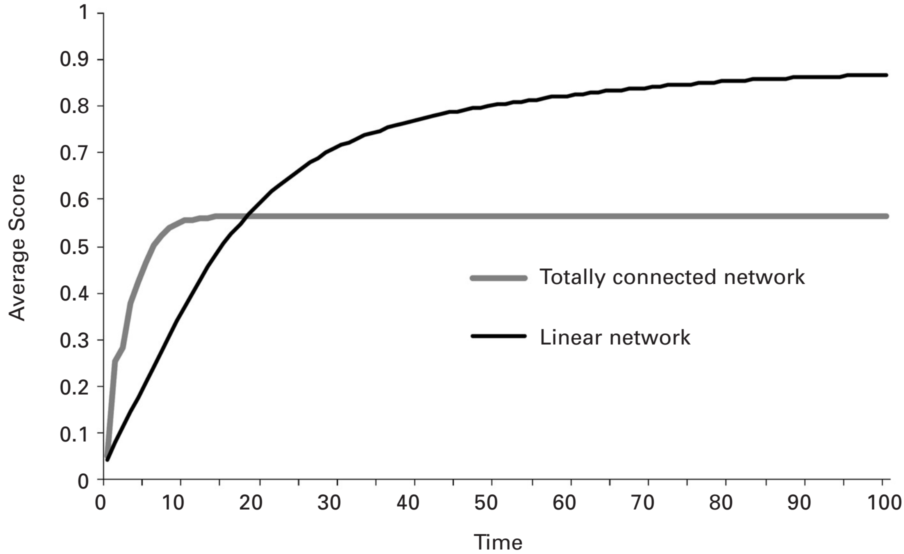
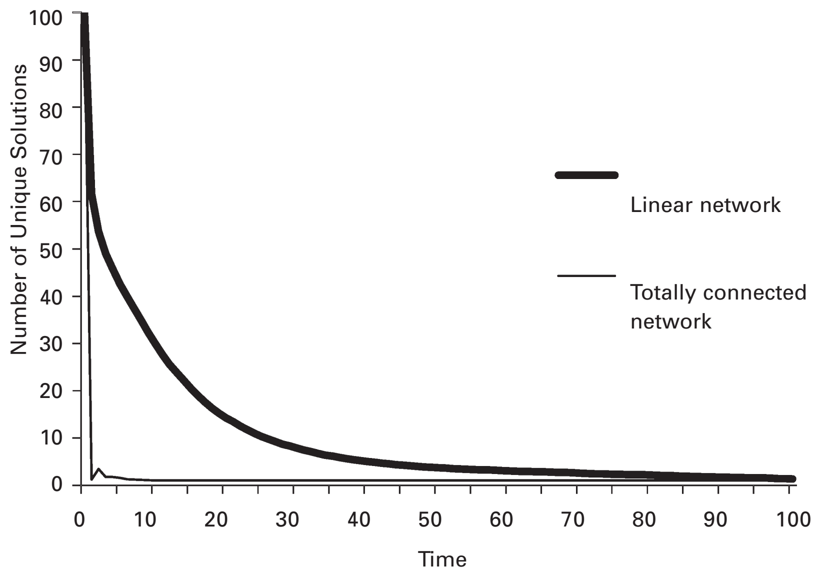


Figure 5. Average number of unique solutions in the system over time in linear and totally connected networks.



The totally connected network quickly drives out diversity, with only one or two unique strategies after the first round because 99 actors converge on the strategy of the best-performing actor. After this convergence, the actors explore (because their performance is identical) and then again converge, and so on, resulting in the "bouncing" observed in figure 5. The system can, at best, only find the best local optimum that is reachable by climbing uphill from the best strategy that exists in the population when it is initially placed in the problem space.

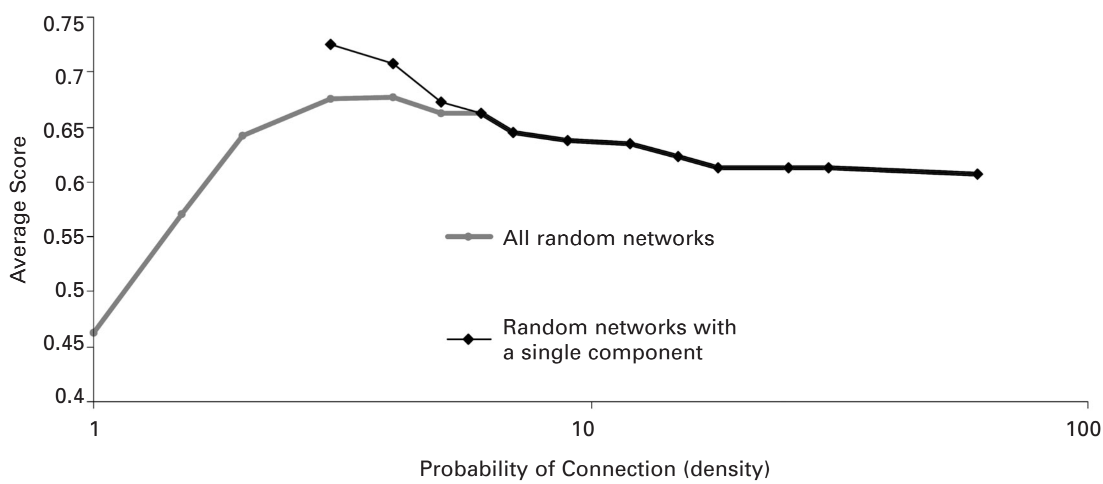
The linear network eliminates diversity far more slowly, allowing exploration around a number of the better strategies that exist in the initial population set. In the short run, this diversity means that most actors' solutions are worse than the best solution, but a better "best" solution will typically be found in the long run. This structural protection of diversity is significant: by time t = 16, on average, there are still approximately 13 solutions left, and the score is roughly equivalent to the average long-term performance of the totally connected network. That diversity allows the linear network to improve substantially before converging on a shared solution, whereas the totally connected network fails to improve as a group.

Random Networks

Random networks offer a continuum between full and minimal connectivity to examine the relationship between connectedness and short- and long-run performance. We generated a series of sets of 1000 random networks based on a density parameter that varied from 1 percent to 50 percent, in which density is the probability that an actor has a tie to any other given actor. In some of the cases, the low-density networks produced by this algorithm contained multiple components. That is, when the probability of links between actors is low enough, networks can contain multiple smaller clusters with no ties between them. When this happens, there is no path between some pairs of actors. Figure 6 plots the average long-run performance of two sets of random networks against network density. The darker line in figure 6 shows only the networks with a single component, while the lighter line indicates the results for all random networks generated.

The results for all the random networks demonstrate a distinct inverted-U shape—very low densities and very high den-

Figure 6. Average performance across network densities of all random networks and random networks with a single component.



sities are low performing, and the optimum performance is at in-between levels of density. The smaller subset of single-component networks reveals the underlying process. At very low densities, the network breaks up into multiple components. For a given structure, smaller populations perform worse than larger populations because they start from a smaller set of possible solutions.[4] The very low densities

4
The impact of population size is modest relative to the impact of network configuration. The results suggest that it is more important how the agents are organized than how many agents are working on the problem (results are available upon request).

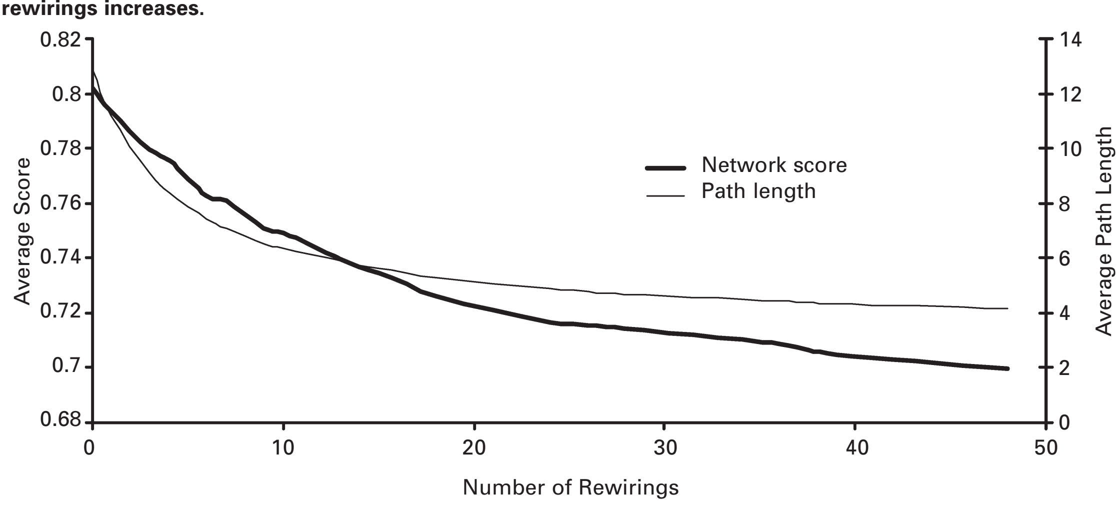
thus are made up of a number of small and thus poorly performing subgroups. But if one only looks at the random networks that are a single component in figure 6, the inverse-U shape disappears, because the less dense networks are better at preserving diversity. This suggests that as long as all actors are connected through some path, the fewer connections the better in the long run.

Small-world Networks

In a random network, increased density corresponds with improved diffusion of information, but the density of a network does not necessarily correspond with its effectiveness at diffusing information. A sparse network may be much more effective at diffusing information than a denser network, depending on its architecture. To separate density from the effectiveness of information diffusion, we examined a series of small-world networks. By varying the probability of rewiring the network, the density of the graph is held constant (p = .04), but the average path length declines as the number of rewired potential shortcuts increases. To test the impact of average path distance, we examined the average long-run performance of a variety of small-world networks, varying the probability of rewiring, such that path distance declines with increased rewiring. Figure 7 presents the average path length and long-run score as a function of the number of shortcuts. We implemented a variation on the standard small-world model (Watts and Strogatz, 1998), to eliminate the effects of isolated components: if a network link selected to be rewired would create an isolated component, that link is passed over, and another is selected. The impact of this selection rule was minimal: with 10 shortcuts, there were no forced exceptions in 1000 simulations, and with 50 shortcuts, less than 5 percent of the networks generated would have created an isolated component. The results discussed here also hold for Newman and Watts' (1999) small-world model, which precludes the potential of fragmented networks by only adding new links without deleting old links.

Strikingly, and consistent with the preceding results, the average long-run scores decline monotonically with the

Figure 7. The average population performance and the average network path length as the number of rewirings increases.



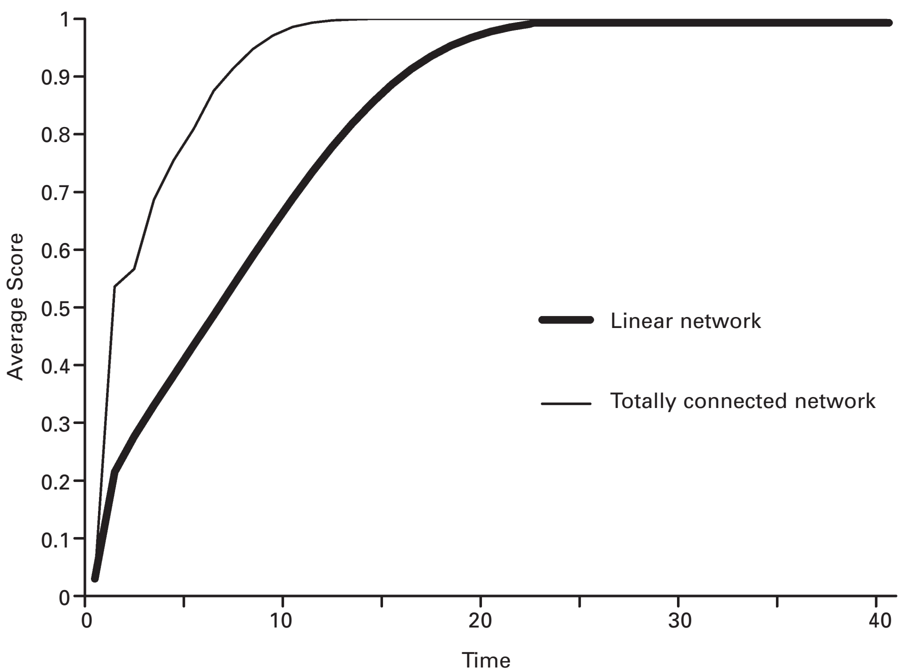
increase in the probability of rewiring. As we increase the number of shortcuts, the density and number of links remain constant: only the average path length decreases.

Ruggedness of the Problem Space

One critical question is whether the results above are contingent on the configuration of the problem space. To examine this, we manipulated the K parameter of the NK space, setting it to 0 (i.e., assuming no synergies and thus producing a space with a single optimum). Figure 8 summarizes the relative average performances of the linear and totally connected networks.

Figure 8 illustrates that the preceding results are contingent on the ruggedness of the problem space. Facing a simple problem, both the linear and totally connected networks find the global maximum in the long run, but the totally connected network gets there much faster. The essential dynamic in the totally connected network is the same as in the more complex world: all actors are pulled up to the strategy of the highest performing actor at the beginning of the simulation. The key difference, of course, is that in the simple world, the system does not get caught in a local optimum. Notably, even a small degree of ruggedness confounds the totally connected network in the long run, which is, on average, beaten by the linear network when K = 1.

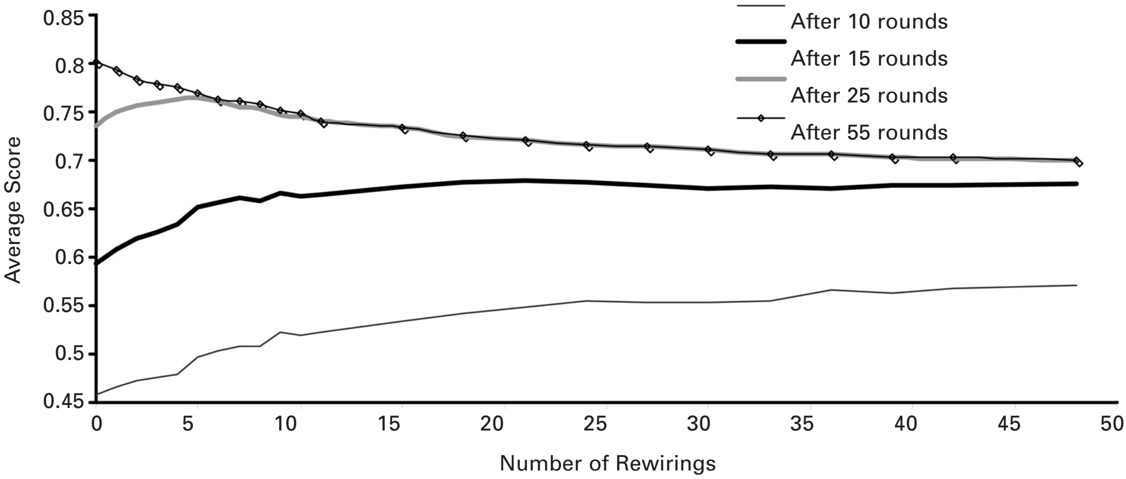
Figure 8. Relative average performance for linear and totally connected networks solving a simple problem.



Time Scale

The well-connected networks consistently improve and plateau rapidly, whereas the poorly connected networks with a single component improve more slowly but reach a higher equilibrium solution. To illustrate this point, in figure 9 we plot system performance against the number of shortcuts for

Figure 9. Rewiring small-world networks for different time durations.



small-world networks for different points in time: t = 10, 15, 25, and 55. These time scales capture a range in which the artificial worlds are far from convergence (t = 10) or close to convergence (t = 55).

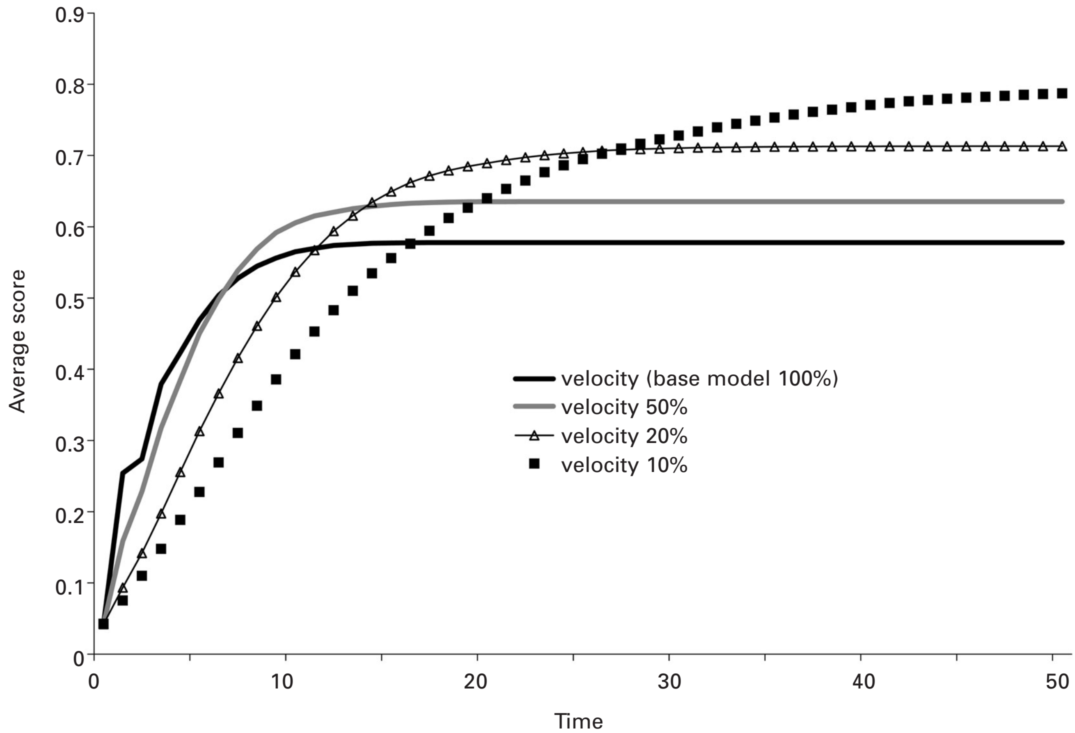
Figure 9 shows that the optimally efficient network depends on the relevant time scale. Given a very short time scale, the more additional ties and the smaller the world, the better. Given a very long time scale, the "bigger" and the more degrees of separation in the world, the better, because it forces actors to explore fully where their initial solution can take them. In the time scale in between, probably where most of real life occurs, the relationship between connectedness and performance is an inverse U, such that the peak of that U depends on the time scale of the problem. The longer the time scale, the less connected the optimal network. For example, the peak of the t = 25 world is to the left of the t = 15 world. This relationship with time can also be seen in figure 4 above, comparing linear and totally connected networks. For the initial periods, the fully connected network dominates, but for t > 18, the linear network becomes superior.

Velocity of Information

The results above suggest that the configuration of the network has a strong effect on the balance between exploration and exploitation, but the system's performance also depends on the speed with which potential solutions diffuse. We modeled this velocity of information diffusion as an asynchronous communication process in which every actor looks around with the same average frequency, but the actual timing for each actor is stochastic.

Figure 10 compares different velocities of information through a totally connected network. Reducing velocity has a clear positive impact on long-run performance for the network at the expense of short-run performance. Reducing the flow of information preserves diversity, forcing actors to achieve improvements based on their initial solutions. The effect is quite dramatic for the full network, because more actors have time to explore in the initial stages, increasing the chance of finding a better solution than the initial highest

Figure 10. Impact of information velocity on average performance over time for a totally connected network.



point. The linear network demonstrated a similar, if not quite as dramatic, improvement.

Strategy Mixing or Error in Copying

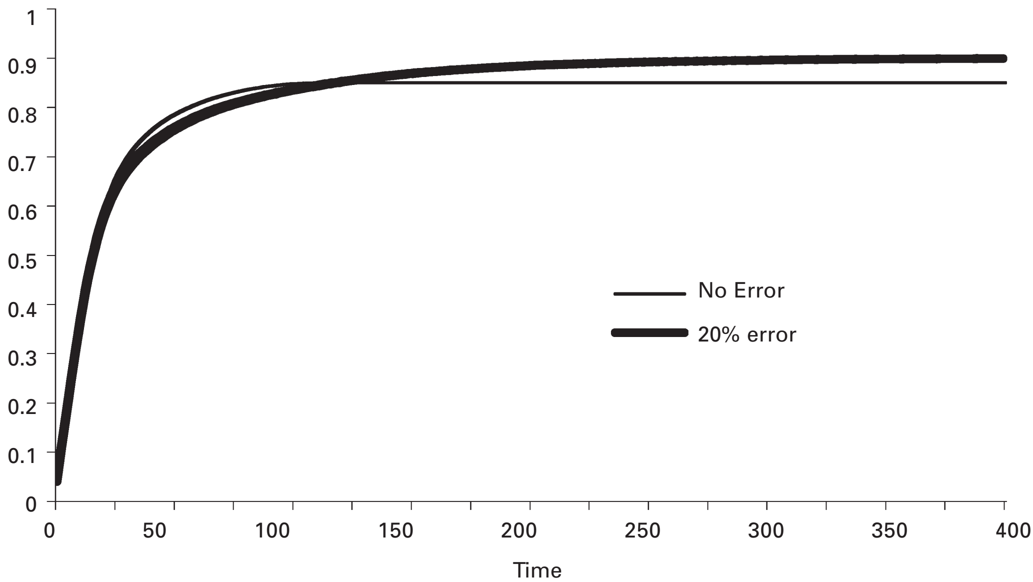
In the above simulations, we assumed that emulation, when it occurred, was complete and perfectly accurate. Here we examine the impact of errors in copying, which, as noted above, produces a mixing of the strategies of the copier and the copied. We implemented error in copying with parameter e affecting dissimilar bits. Recall that each actor's solution is represented as a bit-string, with each bit denoting a dimension of the problem. Imitation in the basic model involves perfect copying of all the bits in the superior solution. With the introduction of error into the copying process, when an actor copies another, it will replace each bit in its solution with the corresponding bit with a probability $(1 - e)$. The rest of the time, the bit will remain in its original state, with probability e, meaning that an error has occurred. The likelihood of no error occurring, then, is $(1 - e)^d$, where d is the number of bits (distance) that are dissimilar. Note that when two actors have very similar solutions, the effective error rate is lower, because there are fewer dissimilar bits. Consistent with the above models, we assumed that actors revert to their status quo solution if an error-induced state is inferior.

Error in copying vastly expands the potential space examined by actors because, in principle, any solution in the entire space between two strategies might be sampled during the copying process. On a large problem space, attempting to copy a distant point would result in exploring a large number of intermediary points. High levels of error will therefore increase the number of search opportunities.

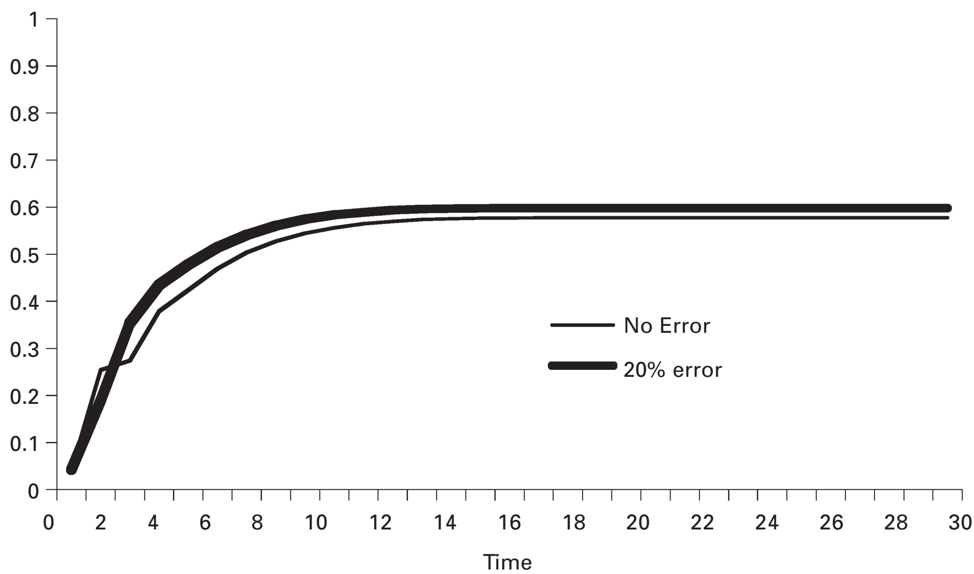
Results of imperfect exploration are shown in figure 11 for the linear and totally connected networks, with error-free models included as points of comparison. Unsurprisingly, error in copying reduces performance in the short run, because propagation of the most successful strategies in the beginning takes longer. In the totally connected network depicted in figure 11, this only lasts for one turn, but higher levels of error increase the lag. The high-error systems outperform the perfect fidelity systems in the long run for both the networks. In fact, the long-run performance of any network presented here is increasing in e. Error rates in copying, in short, alter the balance between exploration and exploita-

Figure 11. Results of errors in exploration in linear and totally connected networks.

A: Error on Linear Networks



B: Error on Totally Connected Networks



tion in the system, increasing the amount of experimentation but reducing the rate with which successful strategies spread. Increasing the error rate magnifies this effect.

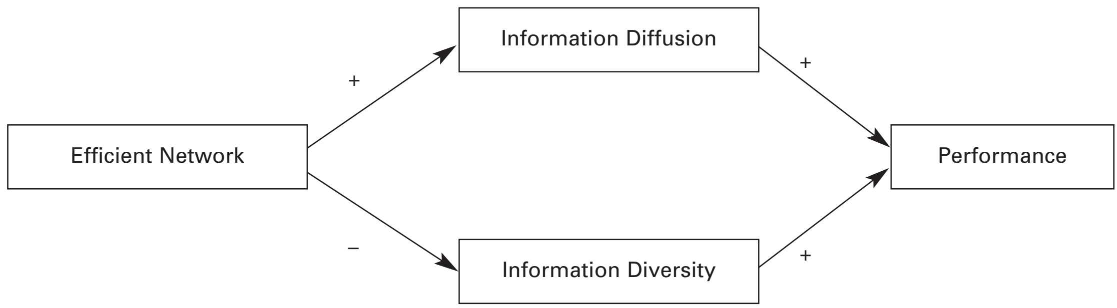
DISCUSSION AND CONCLUSION

The preceding analyses provided multiple layers of reinforcing results pointing to a single core insight: the more efficient the network at disseminating information—as determined by the network configuration, the velocity of information over existing ties, and the fidelity of transmission—the better the system performs in the long run and the worse in the short run. More generally, these results highlight a tradeoff between maintaining the diversity necessary for obtaining high performance in a system and the rapid dissemination of high-performing strategies. Diversity has been found to be beneficial to system performance in a variety of settings (Page, 2007), including small-group performance (Nemeth, 1985), democratic deliberation (Sunstein, 2003), entrepreneurial systems (Florida, 2002), and elite decision making (George, 1972; Janis, 1972). In parallel with these findings is a body of research in the various literatures on knowledge sharing (Wenger, McDermott, and Snyder, 2002) and diffusion (Rogers, 2003) that highlights the performance benefits to the system of spreading information about high-quality strategies. But the research on institutional isomorphism (DiMaggio and Powell, 1983) and social influence (Friedkin, 1998; Lazer, 2001) highlights a necessary corollary to this mimetic process—that imitation reduces the diversity present in a system, which, as discussed, leads to lower performance.

Our analyses thus capture both positive and negative pathways through which the network configuration affects performance. An efficient network positively affects information diffusion, which facilitates the spread of effective strategies, but negatively affects information diversity, which is also positively related to performance. These relationships are depicted in figure 12.

The results of our simulation have implications for empirical work on the relationship between systemic networks and performance. The current literature does not include many of these, and most studies that fit into this category have examined the relationship between team performance and network structure (for a review, see Katz et al., 2004). These studies show a mixed relationship between the density of connectedness and performance. The simulations described above focus on larger populations but, as noted, our results hold for smaller groups as well. One meta-analysis (Balkundi and Harrison, 2006) showed a modest positive relationship, but other recent studies have demonstrated curvilinear relationships (e.g., Leenders, van Engelen, and Kratzer, 2003; Oh, Chung, and LaBianca, 2004; Uzzi and Spiro, 2005) or no relationship at all (Sparrowe et al., 2001). Our results offer potential insight into these apparently contradictory results, where different time horizons (figure 9) reflected a variety of functional relationships between performance and small worldness. Given a short time horizon, reducing average path length, making the world smaller, will unambiguously

Figure 12. The trade-off between information diffusion and diversity in an efficient network.



improve performance. Given a very long time horizon, a shorter path length will unambiguously hurt performance, and for time horizons in between, the relationship is curvilinear. These results suggest that the functional relationship is contingent on the time scale of the task, as well as its complexity, factors not included in any of the studies that we have identified.

Further, of course, what we offered here was a single process model, examining how network structure affects the search of a problem space. Although this is a virtue in the theory-building process, in the real world, many processes may be operating at one time that are affected by the network's configuration. The network configuration of a team, for example, might plausibly affect the coordination of activities, team solidarity, and opportunistic behavior. What might be an optimal configuration for one group in one situation might be dysfunctional for groups in other situations, depending on what process matters in which scenario. It is plausible, for example, that dense ties among teammates, though negatively associated with collective creativity, is important for team solidarity. Results in field studies focusing on overall performance will thus be contingent on the importance of each of these processes for performance. For example, in Balkundi and Harrison's (2006) meta-analysis, 11 of the 37 teams they included involved military teams, the success of which (one would guess) puts a premium on solidarity, coordination, and quick problem solving, all of which are likely positively associated with the density of ties inside the team. Moreover, a complete inquiry must separate density (the number of ties) from average path length.

In contrast, the applications by Uzzi and Spiro (2005) to Broadway productions and by Leenders, van Engelen, and Kratzer (2003) to teams working on product development likely put a premium on creativity inside the team, at some moderate time scale. This is the type of task for which we would predict (and they found) a curvilinear relationship between connectedness and performance. In a related vein, work on brainstorming groups has found that groups working together produce fewer and worse ideas than equivalent isolated individuals exploring the same problem (Lamm and Trommsdorff, 1973; Diehl and Stroebe, 1987). The dominant explanation is that individuals who might discover good ideas are "blocked" by the dominant discussion themes (Diehl and Stroebe, 1991).

Diamond (1999) offered a much more macro example of the phenomena described by our model. He posed the question of why, over the last several centuries, European-based civilization triumphed over civilizations from other continents. One of the key processes that Diamond focused on is the development and diffusion of societal innovations. He argued that there may be a curvilinear relationship between the speed of information diffusion and the technological performance of regions, depending on their "geographic connectedness." As he concluded, "Technology's course over the last 1000 years in China, Europe and possibly the Indian subcontinent exemplifies the net effects of high, moderate, and low connectedness, respectively" (p. 416). If we take Diamond's geographic connectedness to represent ties between the different communities in those regions, our results provide insight into the processes involved.

Limitations and Future Directions

As with any purely theory-based research, our results are only as strong as the premises of our approach. For example, if individuals in a system, rather than simply changing behavior incrementally, tried something radically new on a regular basis, then the above results might not hold. There are thus important potential extensions in both empirical testing and modeling.

Empirical testing. The results of this study posit a particular causal chain. Each part of this chain is testable in the laboratory and in the field. Thus, for example, one can imagine providing a group in a controlled setting with a complex problem to solve and manipulating communication along the lines proposed above: the network structure, information diffusion speed, and signal fidelity. Mason, Jones, and Goldstone (2005) have moved in this direction, showing that totally connected groups can quickly find the optimal solution for a single-peaked problem, but not for a problem with a tri-peaked distribution. More complex problems, such as the traveling salesman problem, are feasible for laboratory participants, and their complexity can be manipulated. In the field, ideally one would begin by testing these findings in settings that put a premium on creativity, measuring both time pressure for decision making and heterogeneity in the group. The prediction would be that systems that are more efficient at disseminating information will quickly converge on a single solution and perform well in the short run but poorly in the long run.

Extending the model. We purposely kept the model to the minimum complexity necessary to capture some essential aspects of parallel problem solving in human systems. There is therefore substantial potential for extending the model. It would be useful to add dimensions to the model that incorporate other factors that we know affect system performance, such as exploring performance with environmental turbulence (Siggelkow and Rivkin, 2005). While we focused on a population's success at solving a generalized problem, one might imagine a set of exogenous shocks altering the problem landscape, driving a constant cycle of adaptation. Would one find that rapidly converging systems perform better, because they converge on adequate solutions before the environment changes, or do systems that maintain some long-run diversity do better because they contain strategies that might be better adapted to future environments? Alternatively, one could assume that different actors optimize on slightly different landscapes, reducing (but not eliminating) what actors can learn from one another. Does that reduced possibility of learning diminish systemic performance, or does it maintain a higher level of diversity in the system, yielding benefits in the long run?

One might also build in assumptions of specialization by actors. Collective problem solving, especially in organizational settings, often involves specialization. That is, there is often a shared conception of the decomposability of a problem into parts that are relatively independent (Simon, 1962). Search thus often involves a set of actors solving different, but interconnected problems. A direction in which one could take this modeling paradigm would be to create problem spaces that reflect this decomposability, with actors specializing in different sections of the problem space. This approach is closer to the model of organizational structure and decomposability first presented by Levinthal (1997). Cooperative action involving specialization would also create the potential for generative information, as each specialist actor or subunit would explore a subset of the problem. The process of sharing information would bring together these results, and the recombination would create new global strategies different from solutions found through a more centralized optimizing process. Although the model described above allows for some active synthesis through error, a more in-depth exploration may be necessary.

It would also be worthwhile to examine a wider array of network structures. Are there, for example, systems that do very well in the short and long run? One could also manipulate the emulation process. Rather than assuming that actors only emulate those who are most successful, one could assume other emulation patterns, such as that based on structural equivalence (Burt, 1987). Finally, it is plausible that modest improvements in individual search would yield vast improvements in system performance. Alternatively, changes in systemic rules—such as rewarding exploration or restricting copying—might have significant effects on outcomes. Along these lines, there are a variety of assumptions that could be made about actors' behavior to incorporate various economic theories of individual choice. For example, an assumption that actors are competing would reduce the utility of sharing a solution, while assumptions about the presence of some property rights might increase the incentives for exploration over exploitation.

Although we would not make managerial recommendations without empirical validation, the practical implications of this vein of research are substantial. Remarkably, the highest performing network in the long run was the slowest, most error prone, and had the longest average path length. More generally, our results highlight that given a long time horizon, efficient communication can actually be counterproductive to systemic performance, whereas given a short time horizon, more communication generally contributes to better outcomes. Our results thus suggest a potential dark side to the rush to connectedness in today's world: it will limit our pathways to developing new and better ways of doing things, even as the demand for new solutions is increasing.

REFERENCES

Abrahamson, E., and L. Rosenkopf
1997 "Social network effects on the extent of innovation diffusion: A computer simulation." Organization Science, 8: 289–309.

Ahouse, J., E. Bruderer, A. Gelover-Santiago, N. Konno, D. Lazer, and S. Veretnik
1991 "Reality kisses the neck of speculation: A report from the NKC Workgroup." In L. Nadel and D. Stein (eds.), 1991 Lectures in Complex Systems: 331–352. Reading, MA: Addison-Wesley.

Arquilla, J., and D. Ronfeldt
2001 Networks and Netwars: The Future of Terror, Crime, and Militancy. Santa Monica, CA: RAND.

Ashworth, M., and K. Carley
2006 "Who you know vs. what you know: The impact of social position and knowledge on team performance." Journal of Mathematical Sociology, 30: 43–75.

Balkundi, P., and D. A. Harrison
2006 "Ties, leaders, and time in teams: Strong inference about the effects of network structure on team viability and performance." Academy of Management Journal, 49: 49–68.

Banerjee, A. V.
1992 "A simple model of herd behavior." Quarterly Journal of Economics, 107: 797–817.

Baum, D.
2005 "Battle lessons: What the generals don't know." New Yorker, January 17: 42–48.

Bettencourt, L. M. A.
2003 "Tipping the balances of a small world" Condensed Matter, abstract cond-mat/0304321 available at http://arxiv.org/pdf/cond-mat/0304321.

Bikhchandani, S., D. Hirshleifer, and I. Welch
1992 "A theory of fads, fashions, custom, and cultural change as informational cascades." Journal of Political Economy, 100: 992–1026.

Borgatti, S. P., and P. C. Foster
2003 "The network paradigm in organizational research: A review and typology." Journal of Management, 29: 991–1013.

Bourdieu, P.
1980 "Le capital social: notes provisoires." Actes de la Recherche en Sciences Sociales, 31: 2–3.

Burt, R. S.
1987 "Social contagion and innovation: Cohesion versus structural equivalence." American Journal of Sociology, 92: 1287–1335.
1995 Structural Holes: The Social Structure of Competition. Cambridge, MA: Belknap Press.

Carley, K. M.
1991 "Theory of group stability" American Sociological Review, 56: 331–354.

Carley, K. M., and V. Hill
2001 "Structural change and learning within organizations." In A. Lomi and E. Larsen (eds.), Dynamics of Organizations: Computational Modeling and Organization Theories: 63–92. Cambridge, MA: MIT Press.

Carroll, T., and R. M. Burton
2000 "Organizations and complexity: Searching for the edge of chaos." Computational and Mathematical Organization Theory, 6: 319–337.

Chang, M.-H., and J. E. Harrington
2005 "Discovery and diffusion of knowledge in an endogenous social network." American Journal of Sociology, 110: 937–976.
2006 "Agent-based models of organizations." In K. L. Judd and L. Tesfatsion (eds.), Handbook of Computational Economics II: Agent-Based Computational Economics: 1273–1338. New York: Elsevier/North-Holland.

Coleman, J., E. Katz, and H. Menzel
1957 "The diffusion of an innovation among physicians." Sociometry, 20: 253–270.

Coleman, J. S.
1988 "Social capital in the creation of human capital." American Journal of Sociology, 94: S95–S120.

Diamond, J.
1999 Guns, Germs, and Steel. New York: W.W. Norton.

Diehl, M., and W. Stroebe
1987 "Productivity loss in brainstorming groups: Toward the solution of a riddle." Journal of Personality and Social Psychology, 53: 497–509.
1991 "Productivity loss in idea-generation groups: Tracking down the blocking effect." Journal of Personality and Social Psychology, 61: 392–403.

DiMaggio, P. J., and W. W. Powell
1983 "The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields." American Sociological Review, 48: 147–160.

Erdos, P., and A. Renyi
1959 "On random graphs." Publicationes Mathematicae, 6: 290–297.

Florida, R.
2002 The Rise of the Creative Class: And How It's Transforming Work, Leisure, Community and Everyday Life. New York: Basic Books.

Friedkin, N. E.
1998 A Structural Theory of Social Influence. Cambridge: Cambridge University Press.

Gavetti G., and D. Levinthal
2000 "Looking forward and looking backward: Cognitive and experiential search." Administrative Science Quarterly, 45: 113–137.

George, A. L.
1972 "The case for multiple advocacy in making foreign policy." American Political Science Review, 66: 751–785.

Gibbons, D. E.
2004 "Network structure and innovation ambiguity effects on diffusion in dynamic organizational fields." Academy of Management Journal, 47: 938–951.

Granovetter, M. S.
1973 "The strength of weak ties." American Journal of Sociology, 78: 1360–1380.
1978 "Threshold models of collective behavior." Journal of Sociology, 83: 1420–1443.

Gupta, A. K., K. G. Smith, and C. E. Shalley
2006 "The interplay between exploration and exploitation." Academy of Management Journal, 49: 693–706.

Harrison, J. R., and G. R. Carroll
2006 Culture and Demography in Organizations. Princeton, NJ: Princeton University Press.

Janis, I. L.
1972 Victims of Groupthink. Boston: Houghton Mifflin.

Katz, N., D. Lazer, H. Arrow, and N. Contractor
2004 "The network perspective on teams." Small Group Research, 35: 307–332.

Kauffman, S.
1995 At Home in the Universe: The Search for Laws of Self-Organization and Complexity. New York: Oxford University Press.

Kollman, K., J. H. Miller, and S. E. Page
2000 "Decentralization and the search for policy solutions." Journal of Law, Economics, and Organization, 16: 102–128.

Krackhardt, D.
2001 "Viscosity models and the diffusion of controversial innovations." In A. Lomi and E. Larsen (eds.), Dynamics of Organizations: Computational Modeling and Organization Theories: 243–268. Cambridge, MA: MIT Press.

Lamm, H., and G. Trommsdorff
1973 "Group versus individual performance on tasks requiring ideational proficiency (brainstorming): A review." European Journal of Social Psychology, 4: 361–388.

Lawler, E. L.
1985 The Traveling Salesman Problem: A Guided Tour of Combinatorial Optimization. New York: Wiley.

Lazer, D.
2001 "The co-evolution of individual and network." Journal of Mathematical Sociology, 26: 69–108.

Leavitt, H. J.
1962 "Unhuman organizations." Harvard Business Review, July–August: 90–98.

Leenders, R. Th. A. J., J. M. L. van Engelen, and J. Kratzer
2003 "Virtuality, communication, and new product team creativity: A social network perspective." Journal of Engineering and Technology Management, 20: 69–92.

Levinthal, D. A.
1997 "Adaptation on rugged landscapes." Management Science, 43: 934–950.

Levinthal, D. A., and J. G. March
1981 "A model of adaptive organizational search." Journal of Economic Behavior and Organization, 2: 307–333.

Macy, M., and R. Willer
2002 "From factors to actors: Computational sociology and agent-based modeling." Annual Review of Sociology, 28: 143–166.

Mahajan, V., E. Muller, and F. Bass
1990 "New product diffusion models in marketing: A review and directions for research." Journal of Marketing, 54: 1–26.

March, J. G.
1991 "Exploration and exploitation in organizational learning." Organization Science, 2: 71–87.

Mason, W. A., A. Jones, and R. L. Goldstone
2005 "Propagation of innovations in networked groups." Proceedings of the Twenty-seventh Annual Conference of the Cognitive Science Society: 1419–1424. Hillsdale, NJ: Lawrence Erlbaum Associates.

Mergel, I.
2005 "The influence of multiplex network ties on the adoption of eLearning practices—A social network analysis." Dissertation thesis, No. 3026, University of St. Gallen, Switzerland.

Miller, K., M. Zhao, and R. J. Calanton
2006 "Adding interpersonal learning and tacit knowledge to March's exploration-exploitation model." Academy of Management Journal, 49: 709–722.

Mitchell, M.
1996 An Introduction to Genetic Algorithms. Cambridge, MA: MIT Press.

Nemeth, C. J.
1985 "Dissent, group process, and creativity." In E. J. Lawler (ed.), Advances in Group Processes, 2: 57–75. Greenwich, CT: JAI Press.

Newman, M. E. J., and D. Watts
1999 "Renormalization group analysis of the small-world network model." Physics Letters A, 263: 341–346.

Oh, H., M. H. Chung, and G. Labianca
2004 "Group social capital and group effectiveness: The role of informal socializing ties." Academy of Management Journal, 47: 860–875.

Page, S. E.
2007 The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies. Princeton, NJ: Princeton University Press.

Paulus, P. B., T. S. Larey, and A. H. Ortega
1995 "Performance and perceptions of brainstormers in an organizational setting." Basic and Applied Social Psychology, 17: 249–265.

Putnam, R.
1993 Making Democracy Work. Princeton, NJ: Princeton University Press.
2000 Bowling Alone. New York: Simon & Schuster.

Reagans, R. E., and E. W. Zuckerman
2001 "Networks, diversity, and performance: The social capital of corporate R&D teams." Organization Science, 12: 502–517.

Rivkin, J. W.
2000 "Imitation of complex strategies." Management Science, 46: 824–844.

Rivkin, J. W., and N. Siggelkow
2003 "Balancing search and stability: Interdependencies among elements of organizational design." Management Science, 49: 290–312.

Rogers, E. M.
2003 Diffusion of Innovations, 5th ed. New York: Free Press.

Schelling, T .S.
1978 Micromotives and Macrobehavior. New York: W.W. Norton.

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

Siggelkow, N., and J. W. Rivkin
2005 "Speed and search: Designing organizations for turbulence and complexity." Organization Science, 16: 101–124.

Simon, H. A.
1962 "The architecture of complexity." Proceedings of the American Philosophical Society, 106: 467–482.

Sparrowe, R. T., R. C. Liden, S. J. Wayne, and M. L. Kraimer
2001 "Social networks and the performance of individuals and groups." Academy of Management Journal, 44: 316–325.

Strang, D., and M. Macy
2001 "In search of excellence: Fads, success stories, and adaptive emulation." American Journal of Sociology, 107: 147–182.

Sunstein, C.
2003 Why Societies Need Dissent. Cambridge, MA: Harvard University Press.

Uzzi, B., and J. Spiro
2005 "Collaboration and creativity: The small world problem." American Journal of Sociology, 111: 447–504.

Walker, J. L.
1969 "The diffusion of innovations among the American states." American Political Science Review, 63: 880–889.

Wasserman, S., and K. Faust
1994 Social Network Analysis: Methods and Applications. Cambridge: Cambridge University Press.

Watts, D.
1999 "Networks, dynamics, and the small-world phenomenon." American Journal of Sociology, 105: 493–527.
2002 "A simple model of global cascades on random networks." Proceedings of the National Academy of Sciences USA, 99: 5766–5771.

Watts, D., and S. Strogatts
1998 "Collective dynamics of 'small-world' networks." Nature, 393: 440–442.

Wejnert, B.
2002 "Integrating models of diffusion of innovation: A conceptual framework." Annual Review of Sociology, 28: 297–326.

Wenger, E., R. McDermott, and W. M. Snyder
2002 Cultivating Communities of Practice. Cambridge, MA: Harvard Business School Press.

APPENDIX A: The NK Space

The scores in the simulation were calculated using an NK model for complex problem spaces. Problem spaces are multidimensional mathematical spaces, in which each point is assigned a value. An NK space is represented as an N bit binary space, $S = \{0,1\}^N$. A point in this space is therefore a binary string of length N. The value of point x is a function mapping predetermined draws from the uniform distribution to each of the N bits in the string:

$$\text{Score}(x) = \frac{1}{N}\sum_{i=1}^{N} v_i(b_i)$$

When K = 0, the function $v_i(b_i)$ is simply one of two set draws from the uniform distribution. When there is no interdependence, experimentation by changing a single bit b_i will only affect the contribution of b_i, allowing a random walk seeking improvements to determine the optimal string x*. In more complex spaces, when K > 0,

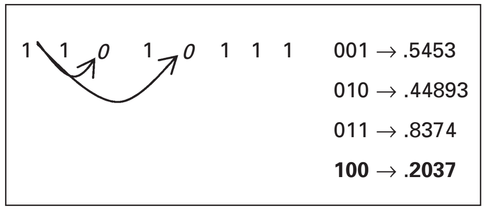
$$v_i(b) = v_i(b_i, b_i^1, \ldots, b_i^K)$$

meaning that the value of each bit is dependent on that bit and on K other bits. The set of $(b_i^1, \ldots, b_i^K)$ is selected randomly for each bit b_i but is constant across the space. In the example below (figure A.1), $v_1(1) = v_1(b_1, b_3, b_5) = v_1(1,0,0)$. Each bit in any point x will yield a value; values are summed and then normalized to produce the value for point x. There are 2^N points in an NK space.

In this paper, we used 1000 NK spaces to test the simulations. The results examine how the mean score of all actors' NK strings change over time. Because different NK spaces will have different structures, and some will have a larger global maximum (high score) than others, we normalize the score against the global maximum for that space with the ratio r_{NK}.

A normalized NK space would produce a distribution of scores similar to the normal curve in figure A.2. The variance is decreasing in problem space complexity (K). This is unsatisfactory for several reasons. If points in the space represent the myriad possible ways to solve a problem, it is unlikely that they would all cluster around the same moderate utility level: most haphaz-

Figure A.1. Computing a sample NK score.*

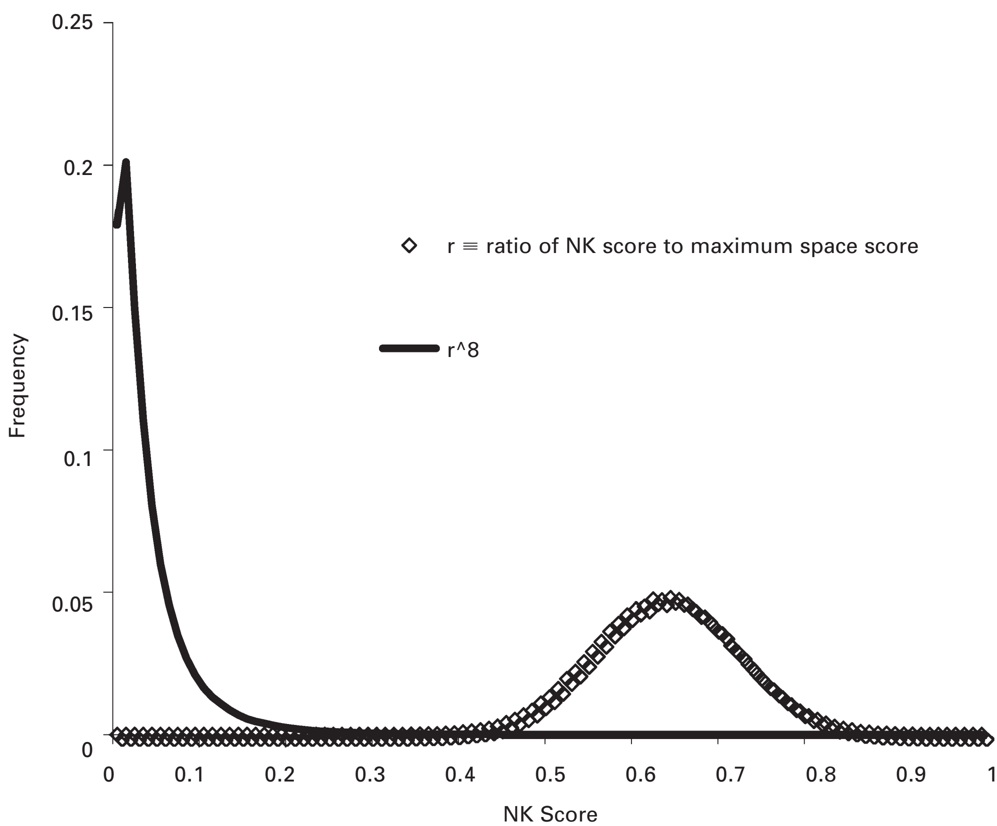


* Each digit contributes to the final score, and each digit's contribution is dependent on K other digits. Here, the first digit adds .2037 to the sum of all digits, which will be averaged over the N bits to produce this NK score.

ard solutions are quite bad, with a much smaller subset having any potential and a very few being "quality" solutions. We used a simple exponential function as a monotonic transformation, r_{NK}^8, to elongate the distribution, classifying most solutions as "bad."

This arbitrary transformation does "bias" the data by making changes in the average outcome easier to detect. We believe that this is acceptable, because rank order is preserved and it reflects the underlying model of problem solving. Another way to gauge results immune to clustering inherent in the NK model would be to look at the percentage of actors who find the global maximum: this yields results similar to those presented above but does not capture "almost as good" solutions, which are a critical part of understanding parallel problem solving.

Figure A.2. Density curves of the value distribution for a sample NK space (n = 20, k = 5).*



* We used the monotonic transformation r^8 to skew normalized scores to demonstrate the relative value of higher scores.

APPENDIX B: Pairwise Comparison of Network Models

The data presented above summarize the average score of a range of network models on the same set of 1000 NK spaces with a set of fixed starting points on each space. Because the only variation was across network structure, we can compare each of the 1000 initial configurations with the results of other models. Note that, for the small world, the random network, and the velocity models, the use of a random seed in the network generation or model execution produces non-deterministic results. Table B.1 summarizes comparisons across networks. Reading across each row indicates the number of times the row network achieved higher performance than the column network. Consistent with the above discussion, the best-performing network was the linear network, followed by the lattice (a "circular" network in which each actor talks to his or her four closest neighbors), small world with 10 extra ties, a sparse (p = .07) random network, and a totally connected network. Diagonal elements do not sum to 1000 because of ties. Ties occur most frequently when both models reach the global maximum.

Table B.1

Number of Times That the Row Network Has a Higher Score Than the Column Network on the Same Set of Problem Spaces and Starting Points

	Full	Line	Lattice	Small world (10 added)	Random (p = .07)
Full	—	26	50	88	138
Line	845	—	468	605	787
Lattice	762	216	—	478	684
Small world	632	132	222	—	557
Random	320	54	83	136	—