Social learning strategies, network structure and the exploration-exploitation tradeoff

Daniel Barkoci1 and Mirta Galesic2

1 Max Planck Institute for Human Development, Berlin, Germany
barkoci@mpib-berlin.mpg.de
2 Santa Fe Institute, New Mexico, USA
galesic@santafe.edu

Abstract. In this paper we study how different social learning strategies composed of three cognitive building blocks (i.e., rules that guide information search, stopping search and decision making) affect population-level performance in a collective problem-solving task. We show that different social learning strategies lead to remarkably different outcomes and demonstrate how these outcomes are affected by the communication networks agents are embedded in. We argue that understanding how communication networks affect collective performance requires taking into consideration the individual strategies used by agents. To illustrate this point we show how our findings can reconcile contradictory results in the literature on network structure and collective problem-solving.

Keywords: social learning, NK model, exploration-exploitation, social networks

1 Introduction

The trade-off between exploration and exploitation lies at the heart of most adaptive problems faced by individuals, groups and organizations, who often need to decide whether to search for new, potentially better solution (e.g., a technology, social institution, or a business strategy) or keep using an existing solution that works well [1, 2, 3, 4]. The right balance between exploration - searching for superior novel solutions - and exploitation - reaping benefits of existing solutions - is thought to be essential for long-term success in such problems [5, 3, 6].

When individuals interact through social learning, this trade-off manifests itself in the balance between innovation through individual learning from the environment and the imitation of existing solutions in the population [7, 8, 9, 10, 11]. Innovation (exploration) is essential both for tracking changes in the environment and for introducing novelty in the population, while imitation (exploitation) serves the purpose of diffusing good solutions in order to increase individual and group-level performance [12].

In this paper we study how the use of different social learning strategies contributes to the population-level balance of exploration and exploitation and how this balance affects overall performance in a collective problem-solving task.
We focus on three representative classes of social learning strategies [13, 14, 11]: (1) copy-the-majority/plurality (hereafter conformity), (2) copy-the-best (hereafter best member), and (3) random copying. We formalize these strategies as algorithms composed of three cognitively plausible building blocks: rules that guide information search, stopping search and making a decision [15].

The resulting social learning strategies vary in the speed with which they are able to diffuse information about useful solutions in the population. We will refer to this speed as strategy efficiency with more efficient strategies implying faster diffusion\(^3\). More efficient strategies, by definition, lead to high levels of imitation (exploitation) in the population, while less efficient strategies promote higher levels of innovation (exploration) since they are less able to identify good solutions at any point in time. Overall, the efficiency of a strategy will determine the level of exploration and exploitation in the population, which in turn, will affect overall performance.

Our concept of strategy efficiency is related to work by Lazer and Friedman [16], who studied the role of network efficiency (i.e., the speed with which network disseminate useful information) in collective problem-solving. In a series of simulations, they varied the efficiency of the communication networks in which agents were embedded and found that efficient networks reached highest performance in the short run while in the long run inefficient networks outperformed efficient ones. Similar results were reported in a behavioral experiment by Mason et al. [17] and in an organizational model by Fang et al. [7]. However, a recent behavioral experiment by Mason and Watts came to the opposite conclusion, finding that groups connected by more efficient networks reached higher outcomes [18].

Note that the above studies either focused on a single social learning strategy (best member) or did not model the strategies underlying individual decisions. We demonstrate that in order to understand the effect of social network structure or decision strategies on collective performance, one needs to study them in conjunction. We show that network structure and decision strategies can interact in ways that are either beneficial or harmful for overall performance. This allows us to reconcile the contradictory findings of Lazer and Friedman [16] and Mason and Watts [18] by showing that both can be obtained depending on which social learning strategy agents are using in a given network.

In sum, our goal is to study how social learning strategies composed of different building blocks affect population-level performance in different social network structures and under varying levels of task complexity.

2 Model

Details of the Model are presented in ODD addendum. Here we briefly describe the task and social environment, and the strategies that agents embedded in these environments use to find good solutions.

\(^3\) Note that our definition of efficiency is related purely to how fast it allows information to diffuse in a population. It does not by itself imply superiority.
2.1 Environment

To design the task environment we use the NK model [19], which belongs to a family of combinatorial optimization problems. The NK model is a "tunably rugged" landscape denoted by $N$, the number of components that make up each solution, and $K$, the number of interactions between these components, which together determine a problem space where different solutions in the space have different payoffs. Consider a technology composed of different parts or an organization with different departmental configurations. Identifying a way to improve a technology or an organization’s configuration depends both on its components and on the interdependencies between the components. For example, changing one component (e.g., increasing the number of departments in an organization) is likely to also have an impact on other components of the organization, and as a result, whether this change will increase overall performance depends on whether it also has a positive effect on the other components with which it interacts.

Depending on the number of interdependencies between components $K$, the landscape can be dominated either by a single global optimum ($K = 0$; see Figure 3A in the ODD addendum) in which case the payoffs of nearby solutions are highly correlated and local search is highly effective; by multiple local optima ($0 < K < N$; see Figure 3B in ODD addendum), where payoffs of nearby solutions can have very different payoffs; or by almost completely uncorrelated landscapes where the payoffs obtained by local search become very similar to a random walk ($K = N - 1$).

We embedded agents in two different social networks, a locally ($d = 3$) and a fully connected lattice representing inefficient and efficient networks, respectively (see Figure 4 in ODD addendum and [16]).

2.2 Agents and Strategies

We simulated $N = 100$ agents. Agents searched the space of possible $N$-digit solutions in a given task and social environment by modifying single digits in their current solutions in order to improve their performance. Agents by default engaged in social learning (exploitation) and switched to innovation (exploration) if the former did not prove successful (see ODD addendum). Specifically, on each time step agents went through the following steps:

1. Implement social learning strategy composed of three building blocks: (i) Search rule: search randomly among the population (ii) Stopping rule: stop searching after looking up the solutions of $s$ other individuals. In different simulations, all agents use one of two sample sizes: either a relatively small ($s=3$) or relatively large ($s=9$) sample size\(^4\). (iii) Decision rule: in different simulations all agents either select the best performing agent (best member); select the most frequent solution (conformity)\(^5\); or select a random agent (random copying).

\(^4\) See section on Sensitivity checks for the best sample size for each strategy.

\(^5\) This implies selecting the majority/plurality solution in the sample. In case each solution is equally frequent, agents choose randomly.
(2) Observe whether the solution identified via social learning produces a higher payoff than the current solution. If yes, switch to the alternative solution; otherwise go to Step 3.

(3) Engage in exploration by modifying a single digit in the current solution and observe whether it produces a higher payoff than the current solution. If yes, switch to the alternative solution; otherwise keep the current solution.

As a benchmark, in some simulations agents did not use social learning strategies but solely an individual learning strategy consisting of Step 3 above.

This procedure is repeated for $t = 200$ time steps and the average payoff in the population is recorded for simulations involving different combinations of strategies and environments. Results reported are averaged across 1000 different NK environments.

3 Results

3.1 Performance

Figure 1 shows the average fitness level achieved by each strategy over time on a rugged landscape with both local and global maxima ($N = 15, K = 7$). Table 1 summarizes results for all other values of $K$. Notice that the results are qualitatively the same for all values of $K > 0$.

In these complex problem spaces (where $K > 0$), different components of a solution interact and, therefore, identifying good solutions is not straightforward. The best member strategies reach the highest short-run outcomes, but they quickly drive the whole population toward locally optimal solutions, from which point individual exploration is not able to find better solutions. As a result the whole population gets stuck in an inferior state. In contrast, the conformity strategies converge more slowly.

Two striking results stand out. First, the small-sample version of the conformity strategy outperforms the large-sample version. Second, the conformity strategy converges to the highest long-run outcomes, outperforming the best member strategies by a large margin. Both the best member and conformity strategies outperform two simple benchmarks, namely random copying and individual learning. This pattern of results was replicated in all complex landscapes ($N > K > 0$, see Table 1).

The intuition underlying these results is the following. The best member strategies are fast at diffusing useful information and, therefore, quickly drive the population toward locally optimal solutions. The small-sample version performs slightly better than the large-sample version, because it leads to slightly slower convergence and thereby allows the population to explore and find local optima that have higher payoffs.

A striking result is that while the large-sample version of the conformity strategy performs poorly, the small-sample version converges to the highest long-run outcomes, outperforming the best member strategies by a large margin. The reason behind this result is that higher payoff solutions identified by a few people
are more likely to sometimes appear as the most frequent solution in small samples, due to a sampling error inherent in estimating the majority option from a sample (see [20]), allowing infrequent but superior solutions to diffuse through the population. In the long run this leads to the highest outcomes because the group is able to search extensively as well as to converge on high fitness solutions over time.

Finally, random copying also engages in high levels of exploitation in the beginning and drives agents to several locally optimal solutions. However, since this decision rule is not biased towards any criteria related to success (e.g. best option, most frequent option) it is not able to drive the population to good solutions.

Taken together, these results indicate that strategies vary in their efficiency, leading to different patterns of explorative and exploitative behavior over time. Efficient strategies such as the best member strategy leads to high levels of exploitation and drives the population toward local optima. Less efficient strategies such as the conformity strategies promote higher levels of exploration and enables the population to find higher-payoff solutions. Since these strategies differ in the patterns of exploration and exploitation they promote, it is important to study how this difference plays out in different network structures which are also known to affect exploration and exploitation [16, 18]. We next turn to the question of how social learning strategies and network structures jointly affect performance.
3.2 Network and Strategy Efficiency

In this section we study the connection between strategy and network efficiency. As mentioned in section 1, two recent studies produced contradictory findings about the performance of efficient and inefficient networks [16, 18]. [16] used simulations to model social learning using the best member strategy with a sample size of two (each agent could communicate with two other neighbors), whereas [18] used behavioral experiments and did not model social learning strategies. The authors differ in how they defined network efficiency but agree that clustering (i.e., the number of local cliques in the network) and average path length (number of steps required to get from one agent to any other agent) are important factors. The former study found that inefficient networks outperformed efficient networks, whereas the latter came to the opposite conclusion.

Here we show that both results can be obtained depending on the social learning strategies that agents use. The underlying explanation is that strategy and network efficiency have similar effects on group-level performance: therefore, their interaction can counteract the individual effect of each. For example, an inefficient strategy in an inefficient network can increase the level of inefficiency to a level where it is no longer beneficial. Similarly, an efficient strategy in an efficient network can result in too much efficiency. Therefore, the concepts of strategy and network efficiency must be studied together.

We compare the performances of the best member and conformity strategies (an efficient and inefficient strategy, respectively) in a fully connected network (as above) with a locally connected lattice ($N = 100, d = 3$; see ODD addendum).
for graphical representations). Since the locally connected lattice only has a degree of 3 we restrict attention to the small-sample strategies, which have also performed best in our main study. Figure 2 shows strategy performance on these two different networks over time.

These results show that both effects can be obtained depending on the underlying social learning strategy. When individuals rely on the best member strategy, inefficient networks outperform efficient ones. When individuals rely on the conformity strategy, inefficient networks are better. As a result we are able to reconcile the contradictory findings from the literature and highlight the need to study the interaction between networks and individual-level strategies.

3.3 Sensitivity Analyses

Noisy social learning

We introduced copying error into the process of social learning. Following [16] we implemented copying error by assuming that whenever agents copy the solutions of others, each digit has a small chance of being incorrectly copied. Whenever this error occurs, the digit from the copied solution is replaced by the respective digit from the agent’s original solution. We varied the error rate from 0.2 to 0.8 with increments of 0.2. Implementing copying error this way ensures that when the solution of an individual differs only slightly from the solution to be copied, the likelihood of making an error in copying is smaller than when the two solutions are very different.

Copying error has several interesting effects. First, it turns out that copying error can have a positive effect on overall outcomes. All strategies perform better with small amounts of copying error than when there is no copying error. Copying error has the largest advantage on the best member strategies. Unsurprisingly all strategies perform poorly when copying error is high (0.8), and in this situation the best member strategy becomes the best performing strategy. Our implementation of copying error is analogous to assuming that agents are able to make long-jumps (i.e., try out solutions that are further away from their current solutions). As a result, copying error is a source of novelty which can sometimes lead to configurations that prove better than the original solutions (see also [10]).

Changing task environments. We study two task environments: a slightly turbulent environment where change occurs only once halfway through the simulation, and a highly turbulent environment where change occurs on every 20th time step. Following [5], we model environmental change by redrawing the fitness contribution of a randomly selected digit in the space. Our results remain the same in slightly turbulent environments, however, when the environment is highly turbulent the best member strategy initially performs better, because it is best strategy in the short run. Given that the landscape changes frequently, there is insufficient time for the long-run advantage of the conformity strategy.

---

6 We chose a network with a degree of 3 to be as close to the original networks studied by [16] and [18] as possible.
to show up. However, over time, agents using the conformity strategy will also get closer to better solutions, and will need to modify fewer digits after environmental change to identify good solutions. As a result, over time, the small sample conformity strategy takes over the best member strategy.

Best sample size. We focus on identifying the best sample size for the best member and conformity strategies, sample size has no effect on random copying. In case of the best member strategy the best sample size turns out to be $s = 2$, however, the difference between different sample sizes is relatively small\(^7\). Therefore we choose to keep sample size of $s = 3$ in the main text to make it directly comparable to the conformity strategy that also uses a sample size of $s = 3$\(^8\). For conformity the best sample size is $s = 3$.

4 Discussion

We studied how social learning strategies composed of different cognitive building blocks affect population-level performance in different social network structures and under varying levels of task complexity.

We introduced the concept of strategy efficiency and related it to the concept of network efficiency \([16, 18]\). We decomposed strategies into three cognitively plausible building blocks - that is, rules that guide search, stopping search, and making a decision - and studied how these rules affect strategy efficiency and how they interact with the structure of the task environment in two different social networks. Our results indicate that collective performance depends both on the network structure agents are embedded in as well as the social learning strategies they use.

Our main result is that best member strategies reach the best performance in the short run, but a small amount of conformity (achieved by relying on small samples) ensured the highest long-run outcomes whenever task environments were complex.

These findings enabled us to reconcile contradictory findings from the literature showing that both inefficient and efficient networks can be beneficial for the same task, depending on the social learning strategies used by individuals \([7, 16, 17, 18]\).

Our study has broad implications for organizational learning, technological innovation and the diffusion of innovations. Most studies of exploration and exploitation in organizations focus on how to design the external environment to make firms more adaptive \([21, 7, 16, 22]\). Our study highlights that it is also important to consider the social learning strategies used by agents and organizations. In addition our study on the relationship between strategies and social

\(^7\) Note that the best member strategy with a sample size of 1 would correspond to the random copying strategy.

\(^8\) Note that the difference between the performance of the best member strategy with sample size 2 and sample size 3 is very small, so using a sample size of $s = 2$ in the main study would not change any of the results. Therefore, to allow for direct comparison with the conformity strategy we stick to a sample size of $s = 3$. 
networks shows that changing the social environment without paying attention to the individual-level strategies might not produce the desired effect.

Research on technological innovation has highlighted the combinatorial nature of innovation with most new inventions being recombinations of existing innovations [23, 24]. Much of this research has focused on how innovation occurs, whereas there has been very little attention devoted to the co-evolution of innovation and imitation. Our study identifies situations where imitation can both help and hinder the development of technological innovation.

Several open questions remain to be addressed. In line with previous studies we focused on the $NK$ landscape as a form of a tunably rugged landscape. The extent to which our results (and other results from the literature) would apply to other landscape problems is a question for future research. We also assumed for the sake of clarity that populations rely on a single social learning strategy. Future research should address the dynamics of exploration and exploitation in a population using multiple strategies at the same time. Our model could also be tested empirically. There are only a handful of studies on how people behave in combinatorial optimization problems that have a rugged structure and we know very little about how these results translate to other problems [25, 26, 27, 28].

Acknowledgments

We thank John H. Miller, Scott E. Page, participants of the 21st Graduate Workshop in Computational Social Science Modeling and Complexity Workshop and the Max Planck Institute for Human Development for financial support.
Table 1. Strategy performance at the final time step ($t = 200$) in different environments.

<table>
<thead>
<tr>
<th>Strategy</th>
<th>$N = 15$, $K = 0$</th>
<th>$N = 15$, $K = 1$</th>
<th>$N = 15$, $K = 3$</th>
<th>$N = 15$, $K = 5$</th>
<th>$N = 15$, $K = 7$</th>
<th>$N = 15$, $K = 9$</th>
<th>$N = 15$, $K = 11$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Best member</td>
<td>mean</td>
<td>$1$</td>
<td>$0.934$</td>
<td>$0.783$</td>
<td>$0.651$</td>
<td>$0.542$</td>
<td>$0.480$</td>
</tr>
<tr>
<td></td>
<td>min</td>
<td>$1$</td>
<td>$0.453$</td>
<td>$0.250$</td>
<td>$0.185$</td>
<td>$0.130$</td>
<td>$0.157$</td>
</tr>
<tr>
<td>$s = 3$</td>
<td>max</td>
<td>$1$</td>
<td>$1$</td>
<td>$1$</td>
<td>$1$</td>
<td>$1$</td>
<td>$1$</td>
</tr>
<tr>
<td></td>
<td>st.dev</td>
<td>$0$</td>
<td>$0.111$</td>
<td>$0.199$</td>
<td>$0.219$</td>
<td>$0.201$</td>
<td>$0.190$</td>
</tr>
<tr>
<td>Best member</td>
<td>mean</td>
<td>$1$</td>
<td>$0.919$</td>
<td>$0.758$</td>
<td>$0.617$</td>
<td>$0.517$</td>
<td>$0.450$</td>
</tr>
<tr>
<td>$s = 9$</td>
<td>min</td>
<td>$1$</td>
<td>$0.362$</td>
<td>$0.229$</td>
<td>$0.137$</td>
<td>$0.078$</td>
<td>$0.128$</td>
</tr>
<tr>
<td></td>
<td>max</td>
<td>$1$</td>
<td>$1$</td>
<td>$1$</td>
<td>$1$</td>
<td>$1$</td>
<td>$1$</td>
</tr>
<tr>
<td></td>
<td>st.dev</td>
<td>$0$</td>
<td>$0.125$</td>
<td>$0.203$</td>
<td>$0.222$</td>
<td>$0.211$</td>
<td>$0.188$</td>
</tr>
<tr>
<td>Conformity</td>
<td>mean</td>
<td>$1$</td>
<td>$0.996$</td>
<td>$0.981$</td>
<td>$0.944$</td>
<td>$0.884$</td>
<td>$0.811$</td>
</tr>
<tr>
<td>$s = 3$</td>
<td>min</td>
<td>$1$</td>
<td>$0.913$</td>
<td>$0.659$</td>
<td>$0.579$</td>
<td>$0.479$</td>
<td>$0.307$</td>
</tr>
<tr>
<td></td>
<td>max</td>
<td>$1$</td>
<td>$1$</td>
<td>$1$</td>
<td>$1$</td>
<td>$1$</td>
<td>$1$</td>
</tr>
<tr>
<td></td>
<td>st.dev</td>
<td>$0$</td>
<td>$0.003$</td>
<td>$0.026$</td>
<td>$0.059$</td>
<td>$0.104$</td>
<td>$0.141$</td>
</tr>
<tr>
<td>Conformity</td>
<td>mean</td>
<td>$1$</td>
<td>$0.998$</td>
<td>$0.973$</td>
<td>$0.925$</td>
<td>$0.854$</td>
<td>$0.798$</td>
</tr>
<tr>
<td>$s = 9$</td>
<td>min</td>
<td>$1$</td>
<td>$0.751$</td>
<td>$0.530$</td>
<td>$0.463$</td>
<td>$0.294$</td>
<td>$0.263$</td>
</tr>
<tr>
<td></td>
<td>max</td>
<td>$1$</td>
<td>$1$</td>
<td>$1$</td>
<td>$1$</td>
<td>$1$</td>
<td>$1$</td>
</tr>
<tr>
<td></td>
<td>st.dev</td>
<td>$0$</td>
<td>$0.016$</td>
<td>$0.068$</td>
<td>$0.119$</td>
<td>$0.156$</td>
<td>$0.174$</td>
</tr>
<tr>
<td>Random copying</td>
<td>mean</td>
<td>$1$</td>
<td>$0.816$</td>
<td>$0.593$</td>
<td>$0.470$</td>
<td>$0.399$</td>
<td>$0.342$</td>
</tr>
<tr>
<td></td>
<td>min</td>
<td>$1$</td>
<td>$0.165$</td>
<td>$0.069$</td>
<td>$0.071$</td>
<td>$0.075$</td>
<td>$0.053$</td>
</tr>
<tr>
<td></td>
<td>max</td>
<td>$1$</td>
<td>$1$</td>
<td>$1$</td>
<td>$1$</td>
<td>$1$</td>
<td>$1$</td>
</tr>
<tr>
<td></td>
<td>st.dev</td>
<td>$0$</td>
<td>$0.201$</td>
<td>$0.244$</td>
<td>$0.227$</td>
<td>$0.197$</td>
<td>$0.177$</td>
</tr>
<tr>
<td>Individual learning</td>
<td>mean</td>
<td>$1$</td>
<td>$0.815$</td>
<td>$0.588$</td>
<td>$0.463$</td>
<td>$0.383$</td>
<td>$0.329$</td>
</tr>
<tr>
<td></td>
<td>min</td>
<td>$1$</td>
<td>$0.483$</td>
<td>$0.349$</td>
<td>$0.234$</td>
<td>$0.159$</td>
<td>$0.158$</td>
</tr>
<tr>
<td></td>
<td>max</td>
<td>$1$</td>
<td>$1$</td>
<td>$0.811$</td>
<td>$0.657$</td>
<td>$0.556$</td>
<td>$0.481$</td>
</tr>
<tr>
<td></td>
<td>st.dev</td>
<td>$0$</td>
<td>$0.1$</td>
<td>$0.078$</td>
<td>$0.071$</td>
<td>$0.062$</td>
<td>$0.055$</td>
</tr>
</tbody>
</table>
Bibliography


ODD addendum

Purpose

The model aims to help understand 1) how social learning strategies varying in efficiency lead to different patterns of explorative and exploitative behavior in the population and 2) how these patterns of behavior interact with the structure of social environments and the complexity of task environments to produce different levels of performance.

Entities, state variables, and scales

The model includes agents using different social learning strategies, embedded in different task and social environments.

Agents/Individuals: All simulations involve a population of 100 agents. Each is randomly assigned an initial task solution consisting of $N$ digits (see section Environment for more details about tasks). In each simulation, all agents using the same rules to explore and/or exploit solutions are present in their environments. There are seven types of agents. Six populations can use social learning strategies, composed of combinations of one search rule, two stopping rules and three decision rules, as follows:

(i) Search rule: all agents search randomly among the population of other agents.

(ii) Stopping rule: agents stop searching after looking up the solutions of $s$ other individuals. We focused on two sample sizes: agents stop after collecting either a relatively small ($s = 3$) or a relatively large ($s = 9$) sample size.

(iii) Decision rule: agents either select the best performing agent (best member); select the most frequent solution (conformity; in case each solution is equally frequent, choose randomly); or select a random agent (random copying).

An additional population of agents use only individual learning, that is explore possible solutions by randomly changing one digit in their current solution, but never copy other agents (see section Process overview and scheduling for more details about exploitation and exploration).

Environment: Agents are embedded in two layers of environmental structure: task environment and social environment.

(i) Task environment is modeled as a tunably rugged $N K$ landscape, where $N$ denotes the number of components of the system and $K$ represents the number of interactions between the components. A value of $K = 0$ produces single peak environment with only one optimum, while $K = N - 1$ produces a completely rugged landscape with no correlation between adjacent solutions, making local exploration ineffective. Intermediate values of $K$ produce landscapes with both local and global optima, with some correlation between nearby solutions. Each solution in the environment is an $N$-length vector composed of binary strings:
that is, each element of the vector can take on two values, 0 and 1, leading to a total of $2^N$ possible solutions in the problem space. Each solution has a payoff that is calculated as the average of the payoff contributions of each element and the other elements with which they interact. The payoff contribution of each element is a random number drawn from a uniform distribution between 0 and 1. In the case of $K = 0$, a simple average of the $N$ elements is taken: $1/N \times \sum_{i=1}^{N} N_i$, whereas with $K > 1$, individual payoff contributions are determined by values of the other $K - 1$ elements, that is, $f(N_i \mid N_i, N_{i+1}, ..., N_K)$, where $f()$ is the payoff function and the total payoff is $1/N \times \sum_{i=1}^{N} f(N_i \mid N_i, N_{i+1}, ..., N_K)$. In other words, when $K = 0$, changing any single element of the solution will affect only the contribution of that element, whereas when $K > 0$, changing a single element will change the payoff contribution of the $K - 1$ other elements. When $K = 0$, exploration of solutions through the modification of single components can prove effective, but as $K$ increases, local exploration becomes less and less effective [5]. Figure 3 displays a simplified graphical illustration of environments that vary in ruggedness. Panel A shows a simple environment where only one unique optimum exists and it is possible to reach this optimum by gradually modifying digits in one’s solution. In contrast, Panel B shows a situation where several local optima exist, which mean that agents can get stuck in a local optima and be unable to reach higher payoffs via local search.

**Fig. 3. Simplified illustration of the two environments studied.** A: Simple environment with a single global optimum. B: Complex environment with multiple local optima and a global optimum. In the simple environment solutions one-digit apart from each other have very similar payoffs, therefore, modifying single digits in a solution will eventually lead to the global optimum. In the complex environment payoffs of nearby solutions can be very different, therefore, search by single digit modification can lead to local optima from which it is impossible to improve and, as a result, to find the global optimum.
We explore landscapes with $N = 15$ and $K = [0, 3, 5, 7, 9, 11]$. Our choices for values of $N$ and $K$ are representative of the literature. Following several authors, we normalize the payoffs of different solutions by dividing them by the maximum obtainable payoff on a landscape $P_{Norm} = P_i / \max(P)$ [16, 22]. The distribution of normalized payoffs tends to follow a normal distribution with decreasing variance as $K$ increases. This implies that most solutions tend to cluster around very similar payoff values. Following [16] we use a monotonic transformation $(P_{Norm})^8$ to widen the distribution, making most solutions "mediocre" and only a few solutions "very good". Note that this assumption does not change any of the results.

(ii) Social environment: Since our main focus is on the properties of different strategies, the majority of our simulations is based on a fully connected lattice with $N = 100$ nodes. We compare the fully connected lattice with a locally connected lattice with 100 nodes and a degree of $d = 3$ (see section on network versus strategy efficiency). This allows us to compare more and less efficient networks as in previous studies [16, 18].

Process overview and scheduling

On each time step agents simultaneously interact with the environment in order to avoid possible sequence effects [29]. On the first time step agents start out with a randomly assigned solution of $N$ digits and on each subsequent time step go through three steps in the following order (unless they are exclusively individual learners, in which case they only engage in Step 3):

1. Implement social learning by means of their search rule, stopping rule, and decision rule (see section Agents for more details).

2. Observe whether the solution identified via social learning produces a higher fitness score than the current solution. If yes, replace current solution with this newly identified solution; otherwise go to Step 3.

3. Engage in exploration via individual learning, by modifying a single digit in the current solution and observe whether it produces a higher payoff than the current solution. If yes, replace current solution with this newly identified solution; otherwise keep the current solution.

This procedure is repeated for $t = 200$ time steps and the average payoff in the population is recorded for different strategies and environments after each time step. Results reported in the paper are averaged across 1000 different $NK$ environments. At the end of each time step, we record payoff of each agent, percentage of agents who exploited other agents solutions through social learning, percentage of agents who keep using their existing solutions, and number of unique solutions in the population.

Design concepts

Basic principles. See Section 1 of the main paper for theoretical rationale underlying the models design.
Emergence. Different combinations of social learning strategies building blocks, task environments, and social environments are expected to produce different patterns of exploration and exploitation behavior, which will in turn affect strategy performance.

Adaptation. Individuals aim to find the solution in the task landscape that has the highest payoff. To do that, they exploit better solutions developed by other agents, or explore the task environment individually by changing one digit at a time. However, building blocks of their learning strategies are fixed.

Objectives. Individuals try to maximize their payoffs. See section Entities, State variables, and Scales for description of how payoffs are calculated for each solution.

Learning. Individuals learn new solutions over time, either by exploiting socially or by exploring individually.

Sensing. Agents can only learn from individuals they are connected to, which depends on the underlying structure of social environments.

Interaction. Agents observe solutions and payoffs that other agents they are connected with had on a previous time step.

Stochasticity. To avoid idiosyncrasies of specific task environments, we repeat the simulation on 1000 different NK environments and present averaged results (see also Table 1 for standard deviation, minimum and maximum value).

Collectives. Agents interact with their neighbors; the number of neighbors is $d = 3$ when the lattice is locally connected, and equal to the size of the whole rest of the network ($d = N = 99$) when the lattice is fully connected.

Observation. At the end of each time step, we record payoff of each agent, percentage of agents who exploited other agents solutions through social learning, percentage of agents who keep using their existing solutions, and number of unique solutions in the population.

Initialization

Each simulation starts with 100 agents. All agents in a given simulation use the same search, stopping and decision rules; and are embedded in a particular task and social environment (see section Entities, State variables, and States). Each agent is assigned a task solution drawn from the underlying NK task environment.

Input data

The model does not use input from external sources such as data files or other models to represent processes that change over time.

Submodels

There are no submodels within the model.
Fig. 4. Illustration of the two social network studied. A: Fully connected lattice where all agents are connected to all other agents B: Locally connected lattice where agents are connected to $d = 3$ local neighbors.