Research Article

COLLECTIVE INTELLIGENCE

Collective Intelligence
August-September 2022: 1–15

DOI: 10.1177/26339137221081849
journals.sagepub.com/home/col
SAGE

Partial copying and the role of diversity in social learning performance

Chelsea M Campbell
School of Informatics, Computing, and Engineering, Indiana University, USA

Eduardo J Izquierdo
School of Informatics, Computing, and Engineering, Indiana University, USA; Cognitive Science Program, Indiana University, USA

Robert L Goldstone
Cognitive Science Program, Indiana University, USA; Psychological and Brain Sciences, Indiana University, USA

Abstract

One major way that people engage in adaptive problem solving is by imitating others' solutions. Prominent simulation models have found imperfect imitation advantageous, but the interactions between copying amount and other prevalent aspects of social learning strategies have been underexplored. Here, we explore the consequences for a group when its members engage in strategies with different degrees of copying, solving search problems of varying complexity, in different network topologies that affect the solutions visible to each member. Using a computational model of collective problem solving, we demonstrate that the advantage of partial copying is robust across these conditions, arising from its ability to maintain diversity. Partial copying delays convergence generally but especially in globally connected networks, which are typically associated with diversity loss, allowing more exploration of a problem space. We show that a moderate amount of diversity maintenance is optimal and strategies can be adjusted to find that sweet spot.

Keywords

Social learning, individual learning, copying, explore/exploit, imitation, network topology

Corresponding author:
Eduardo J Izquierdo, Luddy School of Informatics, Computing, and Engineering and Program in Cognitive Science, Indiana University Bloomington, 3044 Geological Sciences Building, 1001 E 10th Street, Bloomington, IN 47405, USA. Email: edizquie@iu.edu

Significance Statement

Collaboration is a defining characteristic of human society, as is diversity of thought. We use an idealized computational model of collaboration during collective problem solving to examine the effect that partial copying has on the efficiency of a group to find solutions. Additionally, we examine how the effect of copying is modulated by the structure of the network of collaboration and how collaboration influences the diversity of solutions. Our finding that, under a broad range of conditions, collaboration preserves and perhaps even produces diversity is counter-intuitive and has applications to empirical and theoretical study of collaboration, collective intelligence, and evolutionary search. Particularly surprising is the result that highly-connected networks preserve diversity when copying is partial; in prior modeling work, these global networks had been linked to diversity loss. Crucially, we show that partial copying affects the diversity of solutions in the population of learners and that regardless of how diversity is achieved, there is a sweet spot of diversity that offers a population the optimal benefits of learning. As the complexity of a problem increases, the ideal level of diversity also increases. This research supports future work studying more strategies to influence solution diversity, but more broadly, informs how we as humans should approach collective problem solving—collaborating with others with a wide diversity of perspectives, whose ideas may differ greatly from our own.

Introduction

There are two major ways that people engage in adaptive problem solving: copying the solutions of others and trial-and-error individual learning. Individual learners most often search locally among similar solutions, retaining parts of their solution that work well while exploring other options. Making large changes is risky and could result in an under-performing solution. However, although it can result in large changes, fully copying *known better-performing* solutions guarantees improving one's standing. Perfect imitation is an easy, quick way to improve solutions, mitigating some of the risk and effort of individual learning. By approaching the problem collectively, the group benefits from the search efforts of every individual, exploring more of the problem space than any one individual would be able to. The advantages of imitation are well-documented in both empirical work (Derex and Boyd, 2016; Mason and Watts, 2012; Mason et al., 2008; Wisdom et al., 2013; Wisdom and Goldstone, 2011) and computer simulations (Barkoczi and Galesic, 2016; Csaszar and Siggelkow, 2010; Fang et al., 2010; Lazer and Friedman, 2007; Rendell et al., 2010; Posen et al., 2013; Posen and Martignoni, 2018), but, especially in empirical work, it is not always emphasized when subjects' imitation is perfect or imperfect; full or partial.

Within empirical work, imperfect imitation is often present even when it is not the focus of the study or a parameter under direct manipulation (Derex and Boyd, 2016; Mason et al., 2008; Wisdom and Goldstone, 2011; Wisdom et al., 2013). Some of these studies have shown that imitation facilitates productive innovation (Derex and Boyd, 2016; Wisdom and Goldstone, 2011). This result seems counter-intuitive; when many in the population are copying each other, one would assume solution diversity would quickly decline. One reason that diversity is instead preserved and promoted is that humans are *not* perfect imitators. We copy erroneously (Caldwell et al., 2016), adapt information from others (Derex et al., 2015), and we do not copy completely (Posen and Martignoni, 2018)—through partial, imperfect copying, we collectively innovate. Despite these findings, only a few prominent simulations have directly manipulated copying amount. These studies found that partial copying can outperform full copying (Posen et al., 2013; Posen and Martignoni, 2018) and found what amount of copying imparts the greatest benefit under different conditions (Csaszar and Siggelkow, 2010).

The ability to maintain diversity of solutions across the group over time is a common explanatory measure for the performance of social learning strategies (Derex and Boyd, 2016; Fang et al., 2010; Gomez and Lazer, 2019; Lazer and Friedman, 2007; Mason and Watts, 2012; Posen et al., 2013). Is the population focusing and settling on just a few good solutions, or are they spread out over the solution space, searching for better solutions? Diversity maintenance has been shown to be closely correlated to group fitness in genetic algorithms (Burke et al., 2004), but the ways in which different components of social learning strategies affect diversity is an active area of study.

One of the key ways to indirectly manipulate diversity and improve performance is by embedding learners in a network that determines with whom they may collaborate. Several empirical studies have found that when people interact in groups with access to only a few others' solutions, they outperform groups in which anyone can exchange solutions with anyone else (Derex and Boyd, 2016; Wisdom et al., 2013), a result that has been replicated via simulation (Derex and Boyd, 2016; Fang et al., 2010; Lazer and Friedman, 2007; Nahum et al., 2015). In groups with global

communication, where anyone can collaborate with anyone else in the group, information about good solutions spreads quickly, but this efficiency tends to cause the group to bandwagon (Goldstone et al., 2013), where individuals adopt the best solution found so far and abandon search for more optimal solutions. Limiting communication slows down information flow and, consequently, diversity loss (Derex and Boyd, 2016; Fang et al., 2010; Lazer and Friedman, 2007; Mason et al., 2008; Nahum et al., 2015), and can prevent the group from converging on a sub-optimal solution.

Additionally, the complexity and size of the search space affects the performance of globally and locally connected networks (Barkoczi et al., 2016; Goldstone et al., 2013). Although the interactions between task difficulty and network topology have been thoroughly studied (Barkoczi and Galesic, 2016; Fang et al., 2010; Goldstone et al., 2013; Lazer and Friedman, 2007; Mason et al., 2008; Nahum et al., 2015), these studies did not examine the effects that the amount of copying had in their results.

In the current work, we seek to address the following outstanding questions. First, how does partial copying interact with task difficulty and network topology to affect a group's performance and diversity maintenance? To fully characterize the behavior of each condition and the interactions between strategy elements, we study the learning choice dynamics and the diversity over time for each combination of copying condition and network topology across task difficulty. Second, we seek to better understand the relationship between diversity and performance. To do so, we define a metric that captures the overall diversity maintenance of a condition.

In what follows, we first show that partial copying is more successful than full copying regardless of network topology and that this advantage increases with problem difficulty. To understand why this is the case, we then consider the dynamics of the learning process over time and find that partial copying prolongs learning of both types (individual and social), maintains more diversity for longer, and modifies the effect of network topology on the timing and speed of convergence. In our last set of experiments, we analyze the diversity of various strategies and find that the success of a strategy depends, in large part, on how much diversity it is able to maintain overall—a moderate amount of diversity imparts the best performance. Further, we show that as problems become more difficult, successful strategies maintain more diversity. Finally, given our analysis of how strategy elements interact to affect diversity, we suggest how strategies may be adjusted to find optimal diversity and thus, performance.

Methods

Problem space

Following previous work (Barkoczi and Galesic, 2016; Csaszar and Siggelkow, 2010; Fang et al., 2010; Lazer and Friedman, 2007), we model social learning using a group of 100 agents exploring a problem space, searching for ways to improve their solutions and, in turn, group-level performance. We calculate the group performance at any given time step as the average score of the population's solutions at that time step. To create a problem space for the population to explore, we use the NK model (Kauffman and Levin, 1987; Kauffman and Weinberger, 1989), a tunably rugged landscape determined by the number of components (N) that make up solutions and the number of epistatic interactions between those components (K). We represent the solutions in the environment as bit strings of length $N = 15$ (for a total of 2^{15} possible solutions). These bit strings are the solutions which agents seek to improve and each permutation of N bits has an assigned score determined by the NK function. When $K = 0$, each individual bit independently determines a contribution to the solution's overall score: $\sum_{i=1}^{N} N_i$. The score contributions N_i are chosen randomly from a uniform distribution over $[0, 1)$. This is the simplest problem space because it contains a single optimum and the closer a solution is to this optimum the better its score is. When $K > 0$, the contribution of a specific bit also depends on the values of the K following bits. For bits at the end of the bit string, we wrap around to the beginning to continue assigning the "following" bits they are interdependent with. Each configuration of a group of interdependent bits has an associated score contribution, again randomly assigned, and the total score is the sum of the contributions from each group of bits: $\sum_{i=1}^{N} f(N_i|N_i,N_{i+1},\ldots,N_{i+k})$. See Supplemental Figure 1 for an example of score calculation. As K increases, the problem space becomes more complex and there can be many local optima obscuring the single global optimum. When K is at its highest possible value ($K = N - 1$), the problem space is effectively random. In this way, the parameter K determines the "ruggedness" of the problem space, and we vary K systematically between 0 and its highest value, $N - 1 = 14$.

We normalize scores to run between 0 and 1, with 0 corresponding to the worst possible solution and 1 corresponding to the best solution as determined by an exhaustive search of the landscape. Following previous work (Barkoczi and Galesic, 2016; Lazer and Friedman, 2007), we elevate the scores to the power of 8. In NK landscapes, there may be many solutions with scores near 1, making it hard to distinguish between global and local optima. Elevating the scores to the power of eight accentuates performance differences among solutions in the upper range of payoffs. Finally, due to the variability of different instantiations of each problem space, each condition that we report on in this paper was tested using the same set of 1000 randomly generated problem spaces.

Social and individual learning

To navigate these problem spaces, agents engage in social and individual learning. For each problem, the group starts

with initially random solutions and associated scores. At each time step of the simulation, individuals observe the solution and score of one randomly chosen neighbor from their network of collaborators as determined by the network topology. If the neighbor has a better-scoring solution than the individual, the individual copies some number of bits from the neighbor's solution, depending on the copying condition. If the alternative solution is not better scoring, the individual attempts to learn on its own by "flipping" one random bit in its own solution. It keeps the change only if it improves the score of its solution, abandoning it otherwise and not learning that trial. Derex et al. (2015) found that compared to individuals who had access to others' solutions and scores, individual learners with no access to social information made more conservative changes to their solutions, as they did not know how to improve. When one *can* see another's solution and score, they know whether that solution outperforms their own and may feel more confident copying more of the solution. This asymmetry between asocial and social learning is reflected in our learning strategies—individual learners may only modify one bit, while social learners may modify more than one bit.

This social learning strategy is similar to the one used by Barkoczi and Galesic (2016), except we manipulate the number of bits that are copied when agents attempt social learning. In most previous studies, when an individual was copying from a better-performing neighbor, they adopted 100% of that neighbor's solution. We also consider cases where the individual adopts the better individual's solution only partially. In the first set of experiments, we compare partial copying, where the individual copies each bit from their better neighbor's solution with a 50% probability, to full copying, where they copy 100% of the solution. We then systematically explore the amount of copying across the full range, from 1 to 15 bits. In a follow-up experiment, we examine a condition where we again vary the number of bits copied from 1 to 15, but the rest of the bits that are not copied are set to a random bit value.

Communication network

We connect individuals through a communication network that determines who can copy from whom. In addition to manipulating the problem complexity and the amount of copying, we follow others in manipulating the structure of this collaboration network (Derex and Boyd, 2016; Fang et al., 2010; Lazer and Friedman, 2007; Mason et al., 2008; Mason and Watts, 2012). Manipulating the network of collaborations effectively alters the efficiency of information spread in that group. In this paper, we focus our in-depth analysis on the two extremes of efficiency: global groups, in which every member is connected to every other member in the group (i.e., highest efficiency of information spread), and local groups, in which individuals are geographically distributed on a 1D ring and each individual only has access to solutions from their two immediate neighbors (i.e., lowest efficiency of information spread). However, because networks of collaboration in the real world are likely to fall somewhere in between these two extremes (Watts and Strogatz, 1998), we examine the consistency of our results in three other more realistic network topology conditions. First, we examine group performance across the full spectrum of local to global connectivity from individuals being connected to two adjacent agents in the local case, to 99 in the global case (Supplemental Figure 2).

Second, we examine group performance across networks where we systematically vary the in-degree and out-degree of communication between communities. We first randomly divide the population of 100 individuals into five communities ($c = 5$) of 20 individuals each. We follow Girvan and Newman (2002) in connecting vertex pairs with probability P_i for vertices belonging to the same community and probability P_o for vertices belonging to different communities. Once P_i is defined, the remaining probability is split between the rest of the communities in the group, $P_o = (1 - P_i)/(c - 1)$. To achieve a random network in which individuals are as likely to be connected to others in their community as those outside their community, we set $P_i = P_o = 1/c = 0.2$. When $P_i = 1$, $P_o = 0$, and the communities are completely isolated from each other. We omit this condition as it is not easily compared to our other networks, which are all *connected* graphs of 100 nodes. We focus on four networks where $P_i = 0.2$, 0.8, 0.95, and 0.9875, representing a spectrum from a random network to a highly clustered network.

Lastly, we verify our conclusions further on two networks with scale-free and small-world properties (Supplemental Figure 4). The results from all of these varied network structures fall within the extremes of the local and global cases, on which we focus the majority of our analysis.

Results

Partial copying is advantageous across task difficulty, regardless of network topology

Prior studies have found that network topology affects group performance, demonstrating that groups in local networks outperform those in global networks (Derex and Boyd, 2016; Fang et al., 2010; Lazer and Friedman, 2007; Nahum et al., 2015), particularly in difficult problem spaces (Barkoczi et al., 2016; Goldstone et al., 2013; Mason et al., 2008), but these studies only considered full copying. It is unclear if the topology effects would remain across different copying conditions. To uncover the possible interactions between copying amount, network topology, and task difficulty, we study the final group performance for four conditions across two dimensions of interest: amount of copying (full or partial) and connectivity of the group (global or local) (Figure 1). We calculate the final group performance as the average score of all solutions in the group after 2500 learning trials. For the

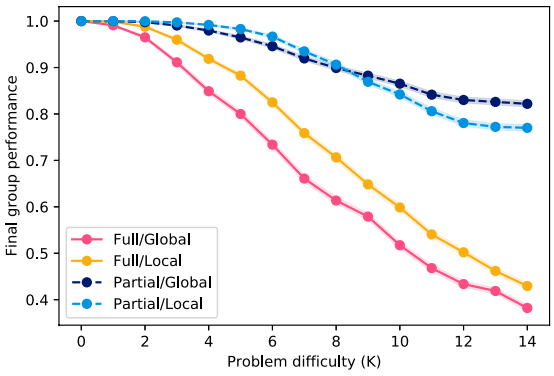

Figure 1. Final group performance across problem difficulty while varying groups across four combinations on two dimensions: topology (local or global) and copying (full or partial). Each point represents performance of the group of 100 individuals after 2500 learning steps, averaged over 1000 repetitions. Dashed lines represent partial copying conditions; solid lines represent full copying conditions. Shaded area represents standard error around the mean. Partial copying results in better group performance than full copying across group connectedness conditions and task difficulties. Although local outperforms global consistently in the full copy conditions, the effect of topology is not as pronounced in the partial copying condition. Also, the dominant topology depends on task difficulty.

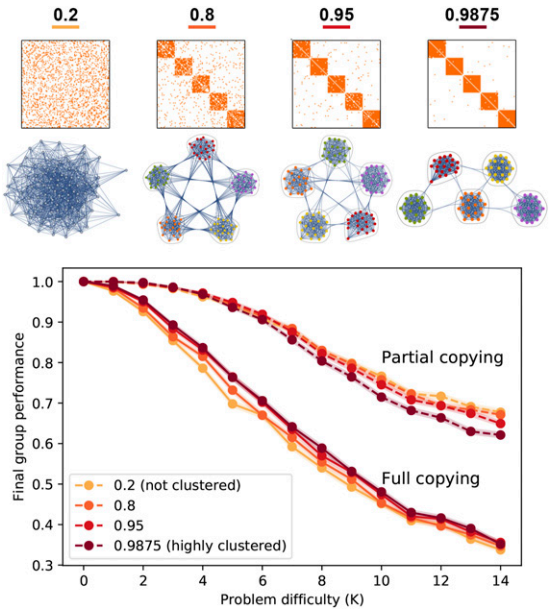

Figure 2. Final group performance across problem difficulty while varying groups across eight combinations on two dimensions: amount of clustering as determined by in-group probability (ranging from a random graph to a highly clustered graph) and copying (full or partial). We maintain a fixed number (c = 5) of equally-sized communities and vary the in-group probability from 0.2, which is a graph with no communities, to 0.9875, which is a highly clustered graph. Each point represents performance of the group of 100 individuals after 2500 learning steps, averaged over 1000 repetitions. Dashed lines represent partial copying conditions; solid lines represent full copying conditions. Shaded area represents standard error around the mean. Above the plot is an example of an adjacency matrix and network visualization for each clustering condition. As the in-group probability increases, connections within communities visually become more dense while connections between communities become more sparse. The results from Figure 1 are validated in these more realistic network topologies. Namely, the advantage of partial over full copying is robust regardless of the community structure.

easiest problem difficulty, K = 0, all conditions reach the global optimum. As K increases, all conditions suffer worse performance, but the partial copying conditions consistently outperform full copying. This advantage only increases as the problem spaces become progressively more rugged. Consistent with a prior study (Posen et al., 2013), we find that imperfect imitation improves performance outcomes. Additionally, we demonstrate that this result is robust to changes in task difficulty and network topology.

When copying is full, our results are consistent with previous findings that local groups consistently outperform global groups across task difficulty (Barkoczi et al., 2016; Derex and Boyd, 2016; Fang et al., 2010; Goldstone et al., 2013; Lazer and Friedman, 2007; Mason et al., 2008; Nahum et al., 2015). However, we also find that which of these topologies performs best depends on the problem difficulty. When copying is partial, the effect of network topology is significantly smaller and more complicated, interacting with problem difficulty. The advantage of local groups largely disappears in spaces of moderate difficulty (K between 4 and 7), while global groups have an advantage in the most difficult and random problem spaces (K > 10). As we systematically vary the group connectivity from local to global (see Supplemental Figure 2), in these most difficult problem spaces, the performance of partial copying increases gradually with connectivity. When copying is full, the advantage of local groups is lost when only a few more neighbor connections are added. The insights gained from this simulation experiment are also consistent across different population sizes and problem dimensionalities (see Supplemental Figure 3).

The advantage of partial copying is robust across more realistic network types

Because more realistic networks of communication are likely to fall somewhere in between local and global networks, we now compare partial copying to full copying across networks with varying degrees of clustering in distinct communities (Figure 2). Community structure is a common property in real-world networks, and describes networks where there are many connections within members of a community and few connections between

communities. The results for these networks are consistent with our findings from the local and global topologies. In what follows, we highlight the main insights from these additional experiments.

First, across all copying conditions and all community topology conditions, we observe that groups perform best in easier problems (low *K*s), and the performance drops gradually as problem difficulty (*K*) increases. Second, the partial copying conditions consistently outperform the full copying conditions. Third, the advantage of partial copying increases with problem difficulty. All three of these observations are consistent with the insights gained from the local and global topology experiments from the previous section.

Additionally, from the previous experiment we learned that, in the full copying condition, the local networks outperform the global networks. The results from the community networks are consistent with this insight as well: the more isolated communities outperform the more interconnected communities. We also had observed that in the partial copying condition for difficult tasks, the pattern is reversed: the global networks outperform the local ones. This again is replicated in the community structure experiments. The networks with more interconnected communities outperform the networks with more isolated communities for difficult problems in the partial copying condition.

The advantage of partial copying is again seen when we compare the performance of full and partial copying for network structures with scale-free and small-world properties (Supplemental Figure 4). Partial copying results in the highest performance for both network types. Importantly, the results from all of these additional more realistic network topologies fall within the extremes shown by the local and global networks in Figure 1. For this reason, we focus the remainder of our in-depth analysis on just those two extreme conditions.

A clear tortoise–hare social learning pattern for full copying; a more complicated pattern for partial copying

A common pitfall for social learning strategies is premature convergence upon local optima. A strategy may lead a group to improve rapidly at first, only to get stuck on a sub-optimal peak and be unable to improve for the rest of the duration of the search. Does partial copying succeed over full copying because it is able to avoid this and keep improving longer before converging? To answer this, we plot the average score of solutions over time for each group for one intermediate problem difficulty, $K = 6$ (Figure 3). In line with our hypothesis, the two partial copying groups converge much later and more gradually than the full copying groups, for which convergence happens abruptly. When copying a solution on a local peak, a *full* copy is guaranteed to trap the copier on that same peak, while a *partial* copy can put the copier in a new area of the problem space instead, allowing for further exploration of solutions and delaying convergence.

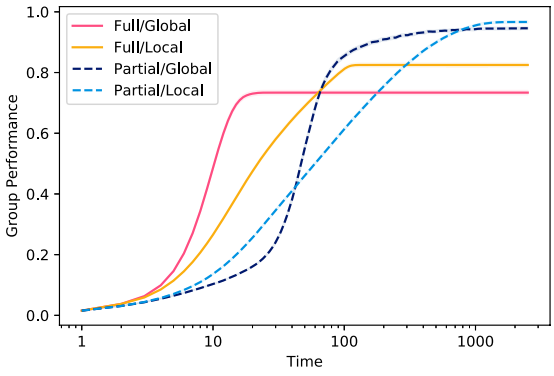

Figure 3. Group performance over time. Each trace represents group performance over time, averaged over 1000 repetitions, for all combinations of full or partial copying, and global or local topology conditions. Shaded area is within perceptible thickness and represents standard error around the mean. Time shown on a log scale because most of the learning occurs in the initial period. For simplicity, we only visualize an intermediate problem difficulty ($K = 6$). Full copying shows the expected tortoise-hare social learning pattern: global topologies improve fast but get stuck early, while local topologies improve slower but also delay getting stuck, and thereby outperform the global condition. The pattern for partial copying is not as simple as the classic tortoise-hare: the local group in this case starts improving faster than the global, but the global changes from slower to faster improvement dynamically during search.

Groups with the same network topology share characteristics in their patterns of improvement, despite differences in convergence due to copying amount. In the full copying condition, our results reflect the tortoise-hare pattern typical of local and global groups adapting in rugged fitness spaces (Nahum et al., 2015). In reference to the fable of the tortoise and the hare, the local group (tortoise) improves slowly but steadily, eventually overtaking the global group (hare), which improves rapidly initially but stops short on a sub-optimal solution. In fact, *both* local groups are characterized by steady improvement, due to the inefficiency of local networks, although the partial copying group improves more slowly. In contrast, the global groups' patterns are both marked by a period of rapid, significant improvement, but the timing of this spike is delayed for the partial copying group. For local and global groups, partial copying suppresses or delays improvement compared to full copying in the same topology. Because of this, when copying is partial, the local group is faster initially (opposite to the full copy case). It gets overtaken once the global group picks up speed, but eventually improves past the global group. Delayed or suppressed *copying rates* could explain the slower initial improvement among partial copiers, and prolonged learning generally could explain

why they improve past full copiers in the end. To test our idea, we consider the dynamics of the learning choices in the population over time.

Partial copiers take more opportunities to learn than full copiers and spend the initial transient exploring the problem space

To explain why partial copying outperforms full copying, regardless of network topology, we examine the learning choices in the populations over time. Recall that agents have three options on each learning trial: they copy if a randomly chosen neighbor has a better solution than they do, attempt individual learning if not, and abandon individual learning if it does not generate a better solution than what they had originally. In the context of the explore-exploit trade-off, individual learning is more exploratory while copying others is exploitation of their good solutions. In Figure 4, we examine the proportion of the population that is copying (Figure 4(a)) and learning (Figure 4(b)) at each time step. The full copying groups cease both types of learning sooner than the partial copying groups, which take more opportunities to improve their solutions, corresponding to continued improvement late in the simulation.

Partial copying increases the proportion that is copying later in the simulation and modifies the timing and extent of peak copying (Figure 4(a)). In global networks, compared to the full copying condition which has an early, significant peak in copying, the partial condition reaches a lower peak much later. In local networks, compared to the full copying condition which has an early but small peak, the partial condition reaches a *higher* proportion of copying only slightly later. In both of these cases, partial copying delays the peak of copying, regardless of the network topology, although the delay is more significant for global groups. The effect of partial copying on the *height* of the copying peak *is* modified by network topology, however. Partial copying reduces the height of peak copying in global networks but increases it in local networks.

Additionally, the partial copying groups have higher proportions of individual learning throughout the simulation (Figure 4(b)). The full copying conditions start at a lower proportion of learning and decrease steadily. Both partial copying conditions start at a higher proportion and decrease more slowly, especially in the global condition, which loses learners the slowest initially. However, the proportion learning in global/partial plummets suddenly around the same time that copying peaks in this condition. Only partial/global shows this distinct learning pattern. The reason for this and for the delayed peaks in copying proportion imparted by partial copying have to do with the solution diversity in the group, which we consider next.

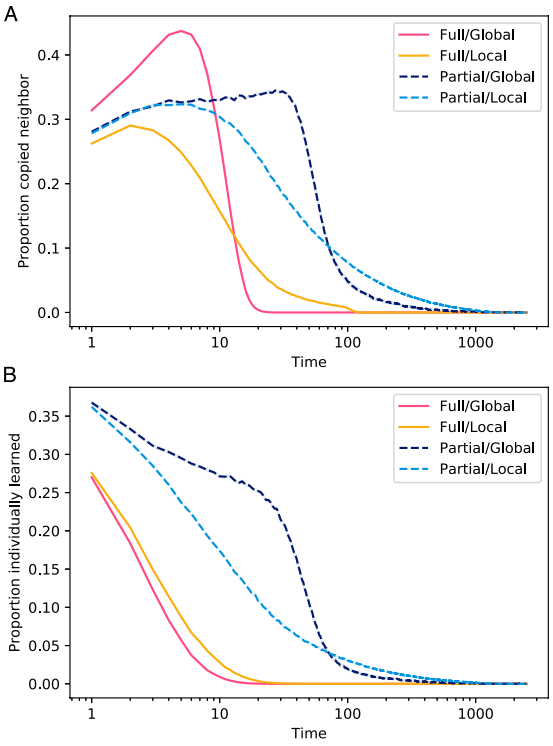

Figure 4. Learning choice dynamics (K = 6). (a) Proportion of individuals in the population that copied a neighbor. Note that in the partial conditions, it is not guaranteed that copying will improve an individual's score. (b) Proportion of individuals in the population that learned individually. Each trace represents group performance over time, averaged over 1000 repetitions, for all combinations of full or partial copying, and global or local topology conditions. The uncertainty is within the perceivable thickness. The proportions in the figures do not sum to 1; aside from copying or individually learning, individuals could also abandon learning that trial if neither copying nor individual learning resulted in an improvement. The learning choices of the group offer insight into their rates of improvement over time, leading to their eventual final performance.

How much individuals copy has a larger effect on diversity than network topology

The performance of many social learning strategies can be at least partly explained by their ability to maintain solution diversity, as this prevents premature convergence on sub-optimal peaks (Derex and Boyd, 2016; Fang et al., 2010; Lazer and Friedman, 2007; Mason et al., 2008; Nahum et al., 2015). Strategies that keep the population spread out over many solutions allow for greater search of the problem space, increasing the likelihood of finding strong solutions before convergence. Because partial copiers spend the initial transient exploring solutions, we expect the diversity maintenance of the partial copying strategies to be high. This is supported by Figure 5, which shows the diversity, measured as the proportion of unique solutions in the group, over time for each condition (see Supplemental Figure 5 for

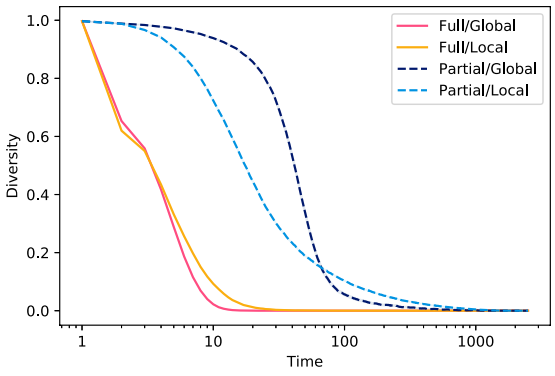

Figure 5. Diversity over time ($K = 6$). Each trace represents the proportion of unique solutions in the population over time, averaged over 1000 repetitions, for full and partial, global and local conditions. The uncertainty is within the perceivable thickness. The partial conditions maintain more population diversity than the full copying conditions overall, regardless of the network topology.

results using Hamming distance to measure diversity). Partial copying groups maintain higher levels of diversity than full copying groups, regardless of their network topology. Consequently, they explore more of the problem space and increase the chance of finding rare, high-scoring solutions to exploit. The ability of partial copying to maintain diversity is also what delays group copying (Figure 4(a)); when diversity is high, average performance is low, and few are lucky enough to have a better neighbor to copy from until diversity wanes.

The partial copying conditions lose diversity the slowest initially, regardless of network topology, because of the two-way interaction between partial copying and diversity. Partial copying is risky because it does not guarantee improvement. Especially combining two dissimilar solutions carries acute risk (Csaszar and Siggelkow, 2010) but can produce a solution in an entirely new part of the problem space, increasing solution diversity. It is more likely that any two solutions will be dissimilar when diversity is high, as at the beginning of the simulation. In this way, diversity and partial copying influence each other; partial copying generates more solution diversity when diversity is higher, which in turn keeps diversity high. The inverse is also true: as the population exploits solutions and diversity naturally diminishes, partial copying *aids* convergence, and learning declines. The solutions being combined are more similar when diversity is low, so partial copying makes limited changes; the new solution will not be far from the better solution that was copied from. This is an idea first discussed conceptually by Goldstone et al. (2013). This hypothesis is supported by the accelerating rate of diversity loss in the partial groups in both network topologies as the simulation progresses. Diversity loss slows again at the end as individuals gradually move from clustering around local optima to converging on the best solution found so far by the population.

How much individuals copy modifies effect on diversity: global networks lose diversity when copying is full, but maintain it when copying is partial

It is significant that partial/global maintains the most diversity of the conditions in the initial transient, although it eventually falls below partial/local at the end of the simulation (Figure 5). In previous works, global networks were associated with diversity *loss* (Derex and Boyd, 2016; Fang et al., 2010; Lazer and Friedman, 2007; Mason et al., 2008; Nahum et al., 2015), as we see in the full/global condition. Why does partial copying cause global networks to hold on to diversity even more than the local networks initially?

As we have discussed, partial copying is diversifying when solution diversity is high, and conforming when it is low. The network structure adds another explanatory layer to this theory. In local networks, diversity decreases gradually; individuals repeatedly copy from the same few others, and solutions in a local area start to resemble each other. However, reaching conformity throughout the social network takes time due to local networks' inefficiency of information spread. This is why partial/local is the slowest and last condition to converge, retaining more diversity than partial/global at the very end of the simulation.

The situation is different in global networks. When copying is partial, global groups maintain more solution diversity than local groups. Because an individual can copy from any other individual, the solutions produced via partial copying in a global network are more diverse than in a local network, where similar solutions are more often combined. This is surprising; many theories (Derex and Boyd, 2016; Fang et al., 2010; Lazer and Friedman, 2007; Mason et al., 2008; Nahum et al., 2015) would suggest that global networks produce *less* diversity, because all solutions end up emulating the best solution in the group. Although partial copying suppresses this tendency initially, once diversity decreases, partial copying becomes more effective in propagating good solutions and the group benefits from the efficiency of its global network, converging faster than partial/local.

Recombining solutions results in good, robust performance across copying amount, but is outperformed by copying with noise for higher copying amounts

Where most simulation studies have modeled copying as taking 100% of a neighbor's solution (Barkoczi and Galesic, 2016; Fang et al., 2010; Lazer and Friedman, 2007; Rendell et al., 2010), humans are likely to vary how much they copy, effectively recombining two partial solutions (Derex et al., 2015; Posen and Martignoni, 2018). One prior study investigated how much individuals should copy to maximize performance, but did not manipulate network topology

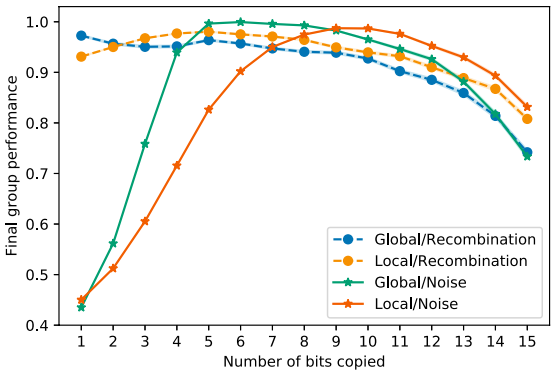

Figure 6. Comparing performance of recombining solutions (standard partial copying) and copying with noise ($K = 6$). Each point represents final group performance after 2500 learning steps, averaged over 1000 repetitions, for local and global conditions, and for copying conditions where individuals copy a number of bits from 1 to 15 and, in the noise condition only, replace the rest of the solution with random bits. Dashed lines denote recombination while solid lines with stars denote copying with noise. Shaded area represents standard error around the mean and is within perceivable thickness. Recombination gives the best performance for low copying amounts; copying with noise gives the best performance for moderate to high copying amounts.

(Csaszar and Siggelkow, 2010). We have demonstrated that strategies where individuals copy 50% of a neighbor's solution outperform strategies where the individual is forced to copy entire solutions, but how much copying is best for a group and how does this optimal amount change with network connectivity? Further, we suggest that the success of partial copying is due to recombination of useful parts of two solutions. However, could it simply be due to the noise added by recombining solutions? To answer these questions, we systematically vary the number of bits that an individual copies for local and global groups and compare two copying conditions: recombining solutions and copying with noise (Figure 6). What we have been calling partial copying amounts to recombining part of a better solution with one's own solution. Here we compare this original strategy of recombining solutions with one where individuals copy with noise—they combine part of a better solution with random bits.

To find what amount of copying imparts the greatest benefit, we consider only the recombination condition. In the global groups, copying a single bit imparts the greatest benefit and each additional bit copied leads to a slight decrease in performance. Copying a single bit introduces a mixture of good existing solutions, but as more bits are copied, there is a destruction of combinations of bits that previously worked well. In the local groups, there is a similar effect, except that copying a few more bits is useful. This is because of the high similarity of solutions in the local region of the social network; recombining solutions does not cause as much disruption as in the global condition.

However, does the improved performance come from recombination or simply from added noise? To answer this question, we compare the performance during recombination and during copying with noise. Recombining solutions has more consistent performance across copying amount than copying with noise, which performs poorly at low copying amounts. When only a few bits are copied (1–4 in the global case; 1–7 in the local case), recombination is the best strategy because the solutions produced in the noise condition are too random. This is the first example of a case where too *much* diversity can hurt performance. However, when more bits are copied from the better neighbor (5–13 in the global case; 8–15 in the local case), discarding the current solution for random bits is the better strategy.

It is also interesting to note that network topology has a noticeable effect on the role of recombination in relation to mere added noise. In local groups, recombination outperforms noise for a broader range of bits copied than in global groups. This suggests that recombination plays a more important role in local groups than it does in global groups. The point at which adding noise becomes optimal is at fewer bits for global (5 bits) than for local (8 bits). The performance of recombining solutions is stable in this range, but the performance of copying with noise in the global condition increases much more steeply with each additional bit copied than in the local condition.

Copying fewer bits and copying with noise both generate diversity

We have begun to see that diversity may be a key explanatory component of the performance of different strategies. Before fully investigating the relationship between diversity and performance, we first examine how our conditions interact to affect diversity. Figure 7 shows how the network structure, amount of copying, and copying strategy (i.e., recombining solutions or adding noise) influence the average diversity of the population. We calculate the average diversity for each condition as a negative log weighted average of the diversity across all time steps, so that measures in the initial transient have more impact on the average (see Supplemental Note 1). The results reveal two main patterns. First, diversity decreases monotonically with increasing copying amount. Second, noise is associated with higher diversity in general, which was expected since copying with noise generates highly diverse solutions.

An intermediary level of diversity maintenance produces the best group performance

Although diversity maintenance is instrumental in the success of partial copying, we found that performance suffers when

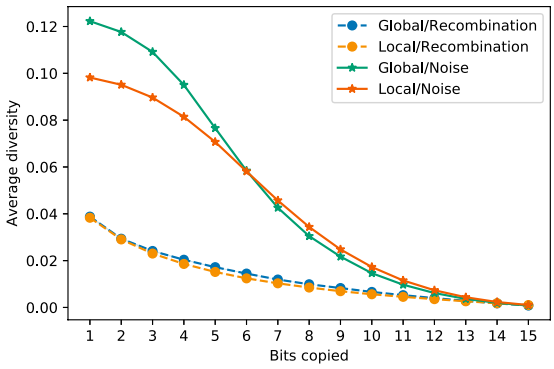

Figure 7. Average diversity associated with number of bits copied (K = 6). Each point represents the negative log weighted average of the diversity over time, averaged over 1000 repetitions, for global and local, recombination and noise conditions. The uncertainty is within the perceivable thickness. Copying fewer bits and copying with noise are both associated with high diversity, while copying more fully decreases diversity.

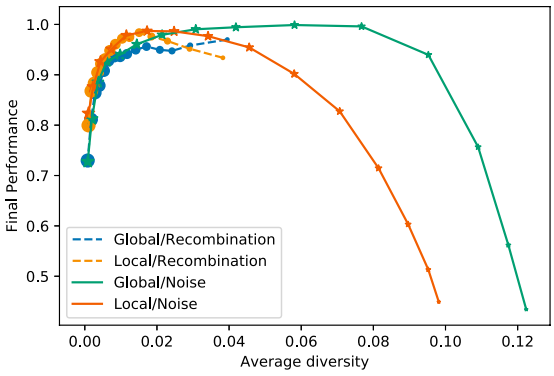

Figure 8. Final performance associated with differing amounts of diversity (K = 6). Each point represents the final performance and diversity associated with a certain number of bits copied, averaged over 1000 repetitions, for global and local, recombination and noise conditions. Point size represents copying amount such that larger points denote more bits copied and smaller points denote fewer bits copied. The left-most, largest points represent the full copying conditions. Each successive point to the right visually decreases in size and corresponds to copying 1 fewer bits, such that each condition has 15 points covering the full range from copying all 15 bits to copying only 1. There is a clear relationship between diversity and performance, such that both too little and too much diversity lead to poor performance. A moderate amount of diversity imparts the best performance outcomes.

there is too much diversity, as when solutions are combined with noise. Conversely, in Figure 5, the full copying conditions lost diversity immediately and also had worse performance. There seems to be a correlation between performance and diversity, but how exactly are they related? To study this relationship, we plot the average diversity for local and global, recombination and copying with noise conditions, through the full range of copying 1–15 bits, against each condition's final performance (Figure 8) (see Supplemental Figure 6 for Hamming distance results). Thus, average diversity is not independently manipulated, but rather varies as a function of the number of bits copied. The most striking result is that all conditions show an inverted U-shape. A moderate amount of diversity is most correlated with high performance. Too little diversity typically indicates that the population converged on a sub-optimal solution. Full copying tends to create this situation and, indeed, the under-performing, low-diversity points to the left capture conditions where copying is relatively full. Contrarily, too much diversity indicates that good solutions were never propagated or exploited, a problem the noise conditions face. In general, we conclude that the average diversity of a population is a good predictor of their collective performance, and this relationship still holds for other values of K. As K increases, the curve shifts to the right (see Supplemental Figure 7), suggesting that more diversity is needed to succeed in more complex problems. We investigate this claim further in the next section.

More difficult problems necessitate greater diversity maintenance

We test our hypothesis from the last section by investigating if the diversity associated with the best performance increases with task difficulty (see Supplemental Figure 8 for Hamming distance results). To find the optimal diversity, we first use a Gaussian filter to smooth the curved relationship between performance and diversity (as in Figure 8) for several values of K (see Supplemental Figure 9 for a visualization of the Gaussian smoothing). We then find the maximum final performance of the smoothed data for each condition, and finally the associated diversity. In Figure 9, for each of the four combinations of network topology and copying strategy and for K = 2, 4, 6, and 8, we plot this optimal diversity. The results confirm that there is a monotonic relation between increasing K and optimal diversity for all four conditions. To do well in more difficult problems, populations must maintain greater overall solution diversity.

Discussion

Across several different ways of manipulating a collective problem-solving situation, from partial to full copying, from combining existing solutions to randomly modifying a single solution, and from locally to globally connected agents, we found a single underlying factor that unifies and systematizes the results—diversity of solutions. The explanatory power of diversity is shown by the largely overlapping lines in Figure 8. Furthermore, there is a systematic, monotonically increasing relation between problem difficulty and the optimal level of diversity.

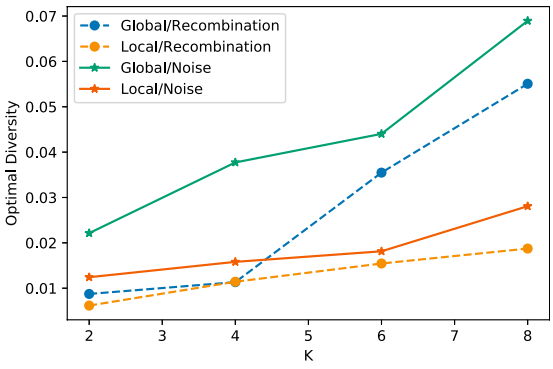

Figure 9. Optimal diversity across problem difficulty. Each point represents the diversity associated with the amount of copying that produced the maximum final performance, averaged over 1000 repetitions, for global and local, recombination and noise conditions. The data was smoothed with a Gaussian filter prior to finding the maximum final performance and associated diversity. The optimal diversity increases monotonically with problem difficulty.

Throughout our analysis, we characterized how the different conditions affect diversity so that strategic and environmental components of social learning strategies may be intelligently assembled to promote optimal group performance. In general, we found that the diversity maintained by a strategy decreases as more bits are copied. Therefore, varying the number of bits copied offers a promising approach to achieving high-performing diversity. For example, copying with noise can generate excessive diversity that hurts performance; copying more bits and consequently adding less noise can mitigate this. Additionally, we found that global networks, when combined with partial copying, maintain diversity but still converge more rapidly than local networks. This feature would be desirable for problems with shorter time horizons.

Although we only analyzed in depth the diversity of populations in the two extreme network conditions, using more realistic networks with varying levels of community structure, we demonstrated that there is a robust interaction between clustering and copying structure. We saw a general advantage for more clustered communities when agents fully copy, but an advantage for less clustered communities when agents partially copy. This is consistent with there being an optimal level of diversity for solving problems. This optimal level can be achieved by combining partial copying, which increases diversity, with less clustering, which decreases diversity; or by combining full copying, which decreases diversity, with more clustering, which increases diversity.

Our work draws from and contributes to several different areas of study within social learning. First, we strengthen the recommendation for varying the degree of copying when modeling social learning, as we have shown that partial copying improves group performance considerably and is a common human behavior. Empirical work has shown that humans choose to copy partially (Derex et al., 2015; Posen and Martignoni, 2018), even though it is risky compared to the improvement guaranteed by fully copying a better solution. Why accept the risks of partial copying? Rogers' paradox (Rogers, 1988) describes performance deficits when there are *too many* individuals perfectly imitating others. The social learners exploit good solutions but contribute no new information, stalling innovation. Confirming the paradox, Rendell et al. (2010) showed that when social learners group together and copy each other, performance degrades in these local groups as they all quickly converge to the same sub-optimal solution. Adaptation of others' solutions, that is, partial copying, is one strategy that resolves this paradox. Wisdom and Goldstone (2010) found that individuals *do* cluster around similar solutions, but innovations are introduced by individuals adapting those solutions. By collaborating locally, the solutions available from peers are more likely to be compatible with the focal individual's solution, allowing them to adapt others' solutions with less risk. Corroborating these findings, Miu et al. (2018) showed that collective improvement is achieved mostly through small, iterative tweaks of the best solution and less commonly, radical innovations of that solution. Yet, aside from a couple of notable exceptions (Csaszar and Siggelkow, 2010; Posen et al., 2013; Posen and Martignoni, 2018), most simulations have modeled copying as full or have had noisy copying (Fang et al., 2010; Goldstone et al., 2013; Lazer and Friedman, 2007; Rendell et al., 2010), which we have shown behaves differently than partial copying.

The second area we contribute to is the study of solution diversity during social learning. While others have mostly looked at diversity over time (Bernstein et al., 2018; Curran, O'Riordan and Sorensen, 2007; Fang et al., 2010; Goldstone et al., 2013; Lazer and Friedman, 2007; Posen et al., 2013), we additionally studied the *overall* diversity maintenance of each of our conditions and uncovered strong relationships between diversity, task difficulty, and performance.

Lastly, our finding that global networks can preserve diversity is contrary to many theories that would say that more globally connected networks increase conformity (Derex and Boyd, 2016; Fang et al., 2010; Lazer and Friedman, 2007; Mason et al., 2008; Nahum et al., 2015). However, when copying is partial, global networks uniquely facilitate the combination of disparate solutions, resulting in novel, highly diverse solutions.

Our results, in combination with important empirical findings, promise several interesting avenues for future work. One possible extension to the current work would be to allow individuals to adjust their social learning strategy based on cues from the environment. This is a well-established behavioral flexibility in humans (Derex et al., 2015; Heyes, 2016; Kendal et al., 2018; Laland, 2004; Rendell et al., 2011; Wisdom et al., 2013) and was a common approach among the top strategies in a simulated tournament of social learning strategies (Rendell et al., 2010). For example, while we fix the number of bits that can be changed via social or individual learning, it would be interesting to allow the

amount copied or changed be influenced by additional environmental and social cues. It is well-documented that social learners are more influenced by higher performing solutions (Derex and Boyd, 2016; Derex et al., 2015; Heyes, 2016; Wisdom et al., 2013) and that both types of learners are more likely to make extensive changes to their solutions when the fitness landscape of the task is simple (Derex et al., 2015). Particularly interesting given our results would be to allow agents to dynamically adjust the amount they are copying based on the population's diversity. This would also be highly applicable to research on diversity-guided evolutionary algorithms (Ursem, 2002). Ultimately, the results from our simulations allow us to make a concrete theoretical prediction: if learners are allowed to select how much to copy, and if the performance of the group has an effect on selection, then given the overall benefits of partial copying to group performance, we would expect partial copying to emerge naturally over time among learners.

Another direction of future study could be investigating the claim that the quality of knowledge is more important than the quantity (Posen et al., 2013). While we show that a moderate amount of diversity leads to peak performance, it could be that this quantity is correlated with the preservation of useful knowledge. There are many established measures of diversity (Burke et al., 2004; Curran et al., 2007), some which may potentially be adapted to better capture the quality of diversity.

The majority of the work in social learning has been focused on three key questions: *who* individuals should copy from, *when* they should copy, and *what* behaviors they should copy (Kendal et al., 2018; Laland, 2004; Rendell et al., 2011). Here, we suggest one more, *how much* to copy. The inclusion of this question has broad implications for metacognitive social learning strategies (Heyes, 2016). If members of a group have different solutions, it may be worthwhile to share and combine ideas and start to narrow down options. If, instead, the group is becoming an echo chamber, copying less to increase the diversity of ideas may be beneficial. Our results highlight that diversity can be highly indicative of the success of a strategy, and further that partial copying can be finely adjusted to offer robust and predictable improvements to diversity and, in turn, group performance outcomes.

Acknowledgments

The authors would like to thank Mirta Galesic, Daniel Barkoczi, Henrik Olsson, Coty Gonzalez, and Edgar Andrade-Lotero for valuable conversations about this work.

Author contributions

E.I. and R.G. designed the study. E.I. and C.C. implemented the experiments and created the figures. All authors analyzed the data. C.C. wrote the manuscript and all authors commented on the paper.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This material is based on work supported by the National Science Foundation under Grant No. 1845322 and Indiana University's Hutton Honors College and Cognitive Science Program.

Data availability statement

This article earned Open Data and Open Materials badges through Open Practices Disclosure from the Center for Open Science. Simulation code is available at: https://github.com/chelseac704/PartialCopying. Data files are available at https://drive.google.com/drive/folders/144fQdlHhPqiDbSYtQBPFLVk9PaDdVo8D?usp=sharing.

ORCID iD

Chelsea M Campbell https://orcid.org/0000-0002-8691-0747

Supplemental material

Supplemental material for this article is available online.

References

Barkoczi D, Analytis P and Wu C (2016) Collective search on rugged landscapes: a cross-environmental analysis. *Proceedings of Annual Conference of the Cognitive Science Society* 38: 918–923.

Barkoczi D and Galesic M (2016) Social learning strategies modify the effect of network structure on group performance. *Nature Communications* 7: 13109.

Bernstein E, Shore J and Lazer D (2018) How intermittent breaks in interaction improve collective intelligence. *Proceedings of the National Academy of Sciences* 115(35): 8734–8739.

Bollobás B, Borgs C, Chayes J, et al. (2003) Directed scale-free graphs. In: *Proceedings of the Annual ACM-SIAM Symposium on Discrete Algorithms*. New York, NY: Association for Computing Machinery (ACM), 132–139.

Burke EK, Gustafson S and Kendall G (2004) Diversity in genetic programming: an analysis of measures and correlation with fitness. *IEEE Transactions on Evolutionary Computation* 8(1): 47–62.

Caldwell CA, Cornish H and Kandler A (2016) Identifying innovation in laboratory studies of cultural evolution: rates of retention and measures of adaptation. *Philosophical Transactions of the Royal Society B: Biological Sciences* 371(1690): 20150193.

Csaszar FA and Siggelkow N (2010) How much to copy? Determinants of effective imitation breadth. *Organization Science* 21(3): 661–676.

Curran D, O'Riordan C and Sorensen H (2007) Evolutionary and lifetime learning in varying nk fitness landscape changing environments: an analysis of both fitness and diversity. *Proc. Natl. Conf. on A.I* 1: 705–711.

Derex M and Boyd R (2016) Partial connectivity increases cultural accumulation within groups. *Proceedings of the National Academy of Sciences* 113(11): 2982–2987.

Derex M, Feron R, Godelle B, et al. (2015) Social learning and the replication process: an experimental investigation. *Proceedings of the Royal Society B: Biological Sciences* 282(1808): 20150719. DOI: 10.1098/rspb.2015.0719.

Fang C, Lee J and Schilling MA (2010) Balancing exploration and exploitation through structural design: the isolation of subgroups and organizational learning. *Organization Science* 21(3): 625–642.

Girvan M and Newman MEJ (2002) Community structure in social and biological networks. *Proceedings of the National Academy of Sciences* 99(12): 7821–7826.

Goldstone RL, Wisdom TN, Roberts ME, et al. (2013) Learning along with others. In: *Psychology of Learning and Motivation*. Cambridge, MA: Academic Press, 58, 1–45.

Gomez CJ and Lazer DMJ (2019) Clustering knowledge and dispersing abilities enhances collective problem solving in a network. *Nature Communications* 10(1): 5146–5211.

Heyes C (2016) Who knows? Metacognitive social learning strategies. *Trends in Cognitive Sciences* 20(3): 204–213.

Kauffman S and Levin S (1987) Towards a general theory of adaptive walks on rugged landscapes. *Journal of Theoretical Biology* 128(1): 11–45.

Kauffman SA and Weinberger ED (1989) The NK model of rugged fitness landscapes and its application to maturation of the immune response. *Journal of Theoretical Biology* 141(2): 211–245.

Kendal RL, Boogert NJ, Rendell L, et al. (2018) Social learning strategies: bridge-building between fields. *Trends in Cognitive Sciences* 22(7): 651–665.

Laland KN (2004) Social learning strategies. *Animal Learning & Behavior* 32(1): 4–14.

Lazer D and Friedman A (2007) The network structure of exploration and exploitation. *Administrative Science Quarterly* 52(4): 667–694.

Luhmann CC and Rajaram S (2015) Memory transmission in small groups and large networks. *Psychological Science* 26(12): 1909–1917.

Mason WA, Jones A and Goldstone RL (2008) Propagation of innovations in networked groups. *Journal of Experimental Psychology: General* 137(3): 422–433.

Mason W and Watts DJ (2012) Collaborative learning in networks. *Proceedings of the National Academy of Sciences* 109(3): 764–769.

Miu E, Gulley N, Laland KN, et al. (2018) Innovation and cumulative culture through tweaks and leaps in online programming contests. *Nature Communications* 9(1): 2321–2328.

Nahum JR, Godfrey-Smith P, Harding BN, et al. (2015) A tortoise-hare pattern seen in adapting structured and unstructured populations suggests a rugged fitness landscape in bacteria. *Proceedings of the National Academy of Sciences* 112(24): 7530–7535.

Posen HE, Lee J and Yi S (2013) The power of imperfect imitation. *Strategic Management Journal* 34(2): 149–164.

Posen HE and Martignoni D (2018) Revisiting the imitation assumption: why imitation may increase, rather than decrease, performance heterogeneity. *Strategic Management Journal* 39(5): 1350–1369.

Rendell L, Boyd R, Cownden D, et al. (2010) Why copy others? Insights from the social learning strategies tournament. *Science* 328(5975): 208–213.

Rendell L, Fogarty L, Hoppitt WJE, et al. (2011) Cognitive culture: theoretical and empirical insights into social learning strategies. *Trends in Cognitive Sciences* 15(2): 68–76.

Rogers AR (1988) Does biology constrain culture. *American Anthropologist* 90(4): 819–831.

Ursem RK (2002) Diversity-guided evolutionary algorithms. In: *International Conference on Parallel Problem Solving from Nature*. Berlin, Germany: Springer, 462–471.

Watts DJ and Strogatz SH (1998) Collective dynamics of 'small-world' networks. *Nature* 393: 440–442.

Wisdom TN and Goldstone RL (2011) Innovation, imitation, and problem-solving in a networked group. *Nonlinear Dynamics, Psychology, and Life Sciences* 15(2): 229–252.

Wisdom TN, Song X and Goldstone RL (2013) Social learning strategies in networked groups. *Cognitive Science* 37(8): 1383–1425.

Wisdom TN and Goldstone RL (2010) Social learning and cumulative innovations in a networked group. In: *International Conference on Social Computing, Behavioral Modeling, and Prediction*. Berlin, Germany: Springer, 32–41.

Appendix

What follows are tables outlining the parameters of the model (Table A1), a comparison between our replication of Barkoczi and Galesic's (2016) model and their original model (Table A2), and important similarities and differences between our results and Barkoczi and Galesic's results (Table A3).

Table A1. Model parameters.

Parameter	Interpretation	Value(s)	Justification
Population size	How many agents make up a population	100 (50, 150 in Supplemental Figure 3)	Following Barkoczi and Galesic (2016)
N	The length of bit string solutions	15 (8, 12 in Supplemental Figure 3)	Following Barkoczi and Galesic (2016)
K	The number of epistatic interactions between bits	0 through 14, but focusing on K=6 for most analyses	Real-world problems vary widely in difficulty, and we wanted to thoroughly evaluate the performance of partial copying across task difficulty. We focus on K=6 for analyzing behavior over time because it determines a landscape that is complex without being too random
t	Simulation length/number of time steps	2500	Most prior studies only ran simulations for short time lengths, we are interested additionally in the long run performance
Repetitions	How many times the same condition is rerun, we then average over the repetitions	1000	Following Barkoczi and Galesic (2016)
Copying amount	Probability of copying each individual bit from another solution	Partial copying: 0.5, also varied through 1/N (copying 1 bit) to 1 (full copy)	Empirical work has demonstrated that humans are not perfect imitators; when given the option to partially copy, we do so
Neighborhood size (network topology)	How many adjacent agents each individual is connected to	Neighborhood size of 2 in the local case; 99 in the global case (we examine the full range from local to global in Supplemental Figure 2)	We focus our analysis on two extremes of network connectivity, which differ greatly in efficiency of information spread
Community clustering (in-group prob.)	The probability that any two vertices in the same community will be connected	P_{in} = [0.2, 0.8, 0.95, 0.9875]	Our method of generating these networks is similar to the method of Girvan and Newman (2002). These four networks range from a random network to a highly clustered network, covering an appropriate variety of clustering in between
Small-world graph: average degree	How many adjacent neighbors each agent is connected to	4	Following Luhmann and Rajaram (2015)
Small-world graph: rewiring probability	The probability of rewiring each edge	0.1	Following Luhmann and Rajaram (2015)
Scale-free graph: *alpha*	Probability of adding a new node connected to an existing node (chosen randomly according to in-degree distribution)	0.41	Following Bollobás, Borgs, Chayes and Riordan (2003), who identified parameters that define a scale-free network that mimics the observed power laws for the web graph

(continued)

Table A1. (continued)

Parameter	Interpretation	Value(s)	Justification
Scale-free graph: *beta*	Probability of adding an edge between two nodes	0.54	Following Bollobás et al. (2003)
Scale-free graph: *gamma*	Probability of adding a new node connected to an existing node (chosen randomly according to out-degree distribution)	0.05	Following Bollobás et al. (2003)
Scale-free graph: $delta_{in}$	Bias for choosing nodes from in-degree distribution	0.2	Following Bollobás et al. (2003)
Scale-free graph: $delta_{out}$	Bias for choosing nodes from out-degree distribution	0	Following Bollobás et al. (2003)

Table A2. Model design similarities and differences between our model and Barkoczi and Galesic's (2016) model.

Population size	We follow Barkoczi and Galesic in fixing the population size to 100 agents for all main experiments, but we additionally study populations of size 50 and 150 in Supplemental Figure 3.
N (solution length)	We follow Barkoczi and Galesic in fixing the solution length to 15, but additionally study solution length 8 and 12 in Supplemental Figure 3.
K (problem complexity)	Barkoczi and Galesic only examine landscapes where K=0 and K=7; we vary K through every possible value from K=0 to K=14.
t (time steps)	We differ in that Barkoczi and Galesic ran simulations for 200 timesteps, while we ran them for 2500 timesteps, which was the point at which all conditions stopped improving.
Repetitions	We follow Barkoczi and Galesic in running simulations for 1000 repetitions.
Social learning strategy	Our social learning strategy differs from that of Barkoczi and Galesic in that agents only observe the solution of 1 other agent each time step, instead of looking at the solutions of 3 or 9 other agents—two conditions that the authors examined. They also compared two strategies where agents either attempted to copy from the best performing agent they observed, or attempted to copy the most frequent solution they observed. Our models are similar in that agents would only copy from their chosen neighbor if that neighbor had a better-scoring solution. If the alternative solution was not better than their own, they attempted individual learning in which they flipped a single random bit, keeping the change only if it granted an improvement.
Copying amount	Barkoczi and Galesic only considered perfect, full copying. We additionally consider partial copying conditions throughout the full range of copying a single bit to copying all bits.
Network connectivity	In addition to the global and local networks we examine here, Barkoczi and Galesic examined eight other network structures that cover the spectrum from local to global. We instead vary the number of adjacent agents individuals are connected to through the full range, look at one random network and 3 with varying amounts of clustering in communities, a Watts-Strogatz small-world network, and a scale-free network.

Table A3. Comparison of relevant results from this and Barkoczi and Galesic's (2016) work.

Both models show that all strategies perform well in simple problem spaces
Barkoczi and Galesic find that social learning strategies modify the effect of network structure. We find that copying amount can also modify the effect of network structure, such that full copying in global networks leads to diversity loss, but partial copying in global networks leads to high diversity maintenance.