How much to copy from others? The role of partial copying in social learning

Chelsea M. Campbell (campbchm@iu.edu)
Cognitive Science Program, Indiana University
Bloomington, IN 47406 USA

Eduardo J. Izquierdo (edizquie@iu.edu)
Cognitive Science Program, Indiana University
Bloomington, IN 47406 USA

Robert L. Goldstone (rgoldsto@iu.edu)
Cognitive Science Program, Indiana University
Bloomington, IN 47406 USA

Abstract

One of the major ways that people engage in adaptive problem solving is by copying the solutions of others. Most of the work on this field has focused on three questions: when to copy, who to copy from, and what to copy. However, how much to copy has been relatively less explored. In the current research, we are interested in the consequences for a group when its members engage in social learning strategies with different tendencies to copy entire or partial solutions and different complexities of search problems. We also consider different network topologies that affect the solutions visible to each member. Using a computational model of collective problem solving, we demonstrate that strategies where social learning involves partial copying outperform strategies where individuals copy entire solutions. We analyze the exploration/exploitation dynamics of these social learning strategies under the different conditions.

Keywords: social learning; individual learning; copying; explore/exploit; imitation; network topology

Introduction

There are two major ways that people engage in adaptive problem solving: trial-and-error individual learning and copying the solutions of others. There is much empirical work studying social learning (Derex & Boyd, 2016; Mason & Watts, 2012; Mason, Jones, & Goldstone, 2008; Wisdom, Song, & Goldstone, 2013; Wisdom & Goldstone, 2011) and a number of computer simulations (Lazer & Friedman, 2007; Fang, Lee, & Schilling, 2010; Barkoczi & Galesic, 2016; Rendell et al., 2010), focusing mainly on *what* individuals should copy, *who* they should copy from, and *when* they should copy. The most common "what to copy" strategies are copy either the best or the most popular solution (Barkoczi & Galesic, 2016). "Who to copy" refers most commonly to what type of communication network individuals should be situated in. Asking *when* one should copy clearly positions social learning as an explore-exploit problem. Exploiting others' solutions prematurely can cause the group to converge on a sub-optimal solution, but once an optimal solution is found, copying is essential. It is evident that when searching for good social learning strategies, there is a lot to take into account, and here we suggest one more: *how much* to copy?

Several computer simulations have addressed how different social learning strategies perform in problem spaces of varying difficulty. Most have focused on the effect of the communication network on group performance, finding that locally-connected groups excel in complex problem spaces while globally-connected groups excel in simple ones (Derex & Boyd, 2016; Lazer & Friedman, 2007; Fang et al., 2010; Wisdom et al., 2013) (going forward we will refer to these network types as simply local or global). When the best solution is easy to find, global groups excel as information about the optimal solution spreads quickly. When the best solution is hard to find, local groups excel because information spreads slowly, maintaining solution diversity longer, and the group can explore more before converging on the best solution. Although there is an impressive range of strategies that have been studied, most every strategy that performs well in complex problem spaces does so by maintaining diversity. In searching for other ways of maintaining diversity we were inspired by empirical work that has found that people do not copy entire solutions (Caldwell, Cornish, & Kandler, 2016; Derex, Feron, Godelle, & Raymond, 2015), though most computer simulations have modeled copying as perfect imitation. In the paper, we are interested in the consequences for the group when its members engage in social learning strategies with different tendencies to copy entire or partial solutions.

Model

Following previous work (Barkoczi & Galesic, 2016; Fang et al., 2010; Lazer & Friedman, 2007), we modeled social learning using a group of individuals exploring a problem space through social and individual learning. In this section, we describe the problem space, the learning strategy, and the network of collaborations.

Problem space. The individuals in the group are solving problems generated through the tunably rugged NK-landscape first developed by Kauffman and Weinberger (1989). In the model, an NK landscape represents the problem space and is determined by the number of dimensions (N) in the space and the number of epistatic interactions between the dimensions (K). Each dimension corresponds to a locus in the solution and the value at that locus determines a contribution to the overall score for that solution. The contribution of a specific dimension, however, also depends on the values at K other dimensions. In this way, the parameter K determines the "smoothness" of the problem space. The simplest problem space, $K = 0$, contains a single global optimum, but as K increases, the problem space becomes more rugged. When K is at its highest possible ($K = N - 1$), the problem space is effectively random. For the simulations presented below, we fixed the dimensionality ($N = 15$) and varied systemically the ruggedness of the problem space (K between 0 and 14). For each problem space, scores were normalized 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 (Lazer & Friedman, 2007; Barkoczi & Galesic, 2016), we elevated 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 8 widens the distribution of the upper range of payoffs.

Social and individual learning. We modeled a group of 100 individuals exploring the problem space through social and individual learning. For each problem, the group started with initially random solutions and associated scores. At each time step of the simulation, individuals observed the solution and score of one randomly chosen neighbor from their network of collaborators. If the neighbor had a better scoring solution than the individual, the individual copied the neighbor's solution entirely. If the alternative solution was not better scoring, the individual attempted to learn on its own by "flipping" one random bit in its own solution. It kept the change only if it improved the score of its solution, abandoning it otherwise. The primary difference between our model and previous modeling work was our social learning strategy. In most previous studies, when an individual was copying from a better-performing neighbor, they adopted 100% of that neighbor's solution. In contrast, we also considered cases where the individual adopted the better individual's solution only partially. In the partial copying conditions, the individual copied each bit from the better individual's solution with a 50% probability. In one experiment, we also systematically explored the amount of copying across the full range, from 0 to 15 bits. In a follow-up experiment, we examined a condition where we again varied the number of bits copied from 0 to 15, but the rest of the bits that were not copied were set to a random bit value.

Collaboration network. Individuals in a group were connected to each other through a network of collaborators. In other words, each individual had specifically assigned neighbors with which they could collaborate. In addition to manipulating the problem complexity and the amount of copying, we followed others in manipulating the structure of this collaboration network (Derex & Boyd, 2016; Fang et al., 2010; Lazer & Friedman, 2007; Mason et al., 2008; Mason & Watts, 2012). Manipulating the network of collaborations effectively alters the efficiency of information spread in that group. In this paper, we report on two network architectures: global groups, in which every member is connected to every other member in the group (i.e., high 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 immediate neighbors (i.e., low information spread). Networks of collaboration in the real world are likely to fall somewhere in between these two extremes. Although we systematically varied the size of the neighborhood for each, from purely local to fully global, we report only on the architectures at the two extremes of the spectrum for simplicity.

Measuring performance. In order to measure the performance of the group on any one condition (i.e., problem difficulty, degree of copying, and collaboration network), we follow previous work in examining the average score across the individuals in a group at the end of the learning trials. In preliminary work, we also analyzed the performance of the best individual in the group. However, the goal of the present paper will be to examine the advantages of different degrees of copying for the group as a whole. 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 1,000 problem spaces.

Results

To examine the efficacy of partial copying, we first analyzed the four combinations of copying strategy and network topology across task difficulty. Finding partial copying to be the most successful, we considered the dynamics of the learning process over time to explain this result. We then varied the amount of bits copied to see if a specific amount of copying was optimal. We compared the results of this and a similar experiment to verify that the advantage of partial copying was from mixing two solutions, and not just from adding random noise to a good solution.

Should individuals copy entire solutions?

Copying entire solutions saves an individual time and mitigates potential risks from mixing partial solutions, but this may not be the best strategy for the group. In the first set of experiments, we analyzed group performance when individuals copied entire solutions compared to only part of solutions (i.e., a random 50% of the solution). Because recent theoretical work has highlighted the effect of the network communication structure on group performance, we considered conditions where individuals were connected globally and locally.

Final group performance. The first step in our analysis was to examine group performance for four conditions across the two dimensions of interest: amount of copying (full or partial) and connectivity of the group (global or local).

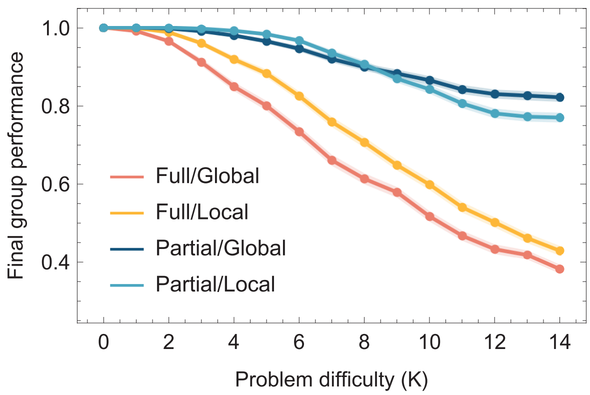

Figure 1: Final group performance across problem difficulty. Each point represents performance of the group of 100 individuals after 2,500 learning steps, averaged over 1,000 repetitions. Shaded area represents standard error around the mean.

Specifically, we examined the average score of all the final solutions in the group after 2,500 learning trials across all possible problem difficulties (Fig. 1). For the simplest possible problem space (K=0), all conditions consistently found the global optimum. As problem difficulty increased, group performance decreased across conditions, decreasing more drastically for full copying strategies than partial copying strategies. For most problem spaces, partial copying outperformed full copying conditions. When individuals copied fully, the local groups consistently outperformed global groups across problem difficulty. Finally, the difference in performance afforded by the local network structure was smaller for partial copying than for full copying groups. Although within partial copying, the effect of network structure was smaller, local groups had a small advantage for intermediate problem spaces (K between 4 and 7) and global groups had an advantage on the hardest and most random problem spaces (K>10).

Group performance dynamics. The second step in our analysis was to consider the dynamics of group performance over time (Fig. 2). In order to examine the asymptotic behavior of these strategies, we ran each simulation until we observed no more learning. However, we focused our analysis on the initial transient since that is when most of the learning occurs. For simplicity, we analyzed only the dynamics in one intermediate problem difficulty (K=6), where the difference in performance between full and partial copying is large, but the problem space is not too random. When individuals copied entire solutions, global groups learned fast and got stuck quickly; local groups moved slower but arrived at a better solution. This is consistent with what has been reported in previous work (Lazer & Friedman, 2007; Barkoczi & Galesic, 2016). When copying was partial, performance for both local and global groups improved more slowly initially, but with time they outperformed the full copying strategies by a significant margin. So although full copying increased group performance over the short term, partial copying led to better asymptotic performance overall for both of the network topologies we studied. Finally, in the partial copying conditions, the global groups reached the optimum earlier than the local ones; but the local groups eventually matched their performance.

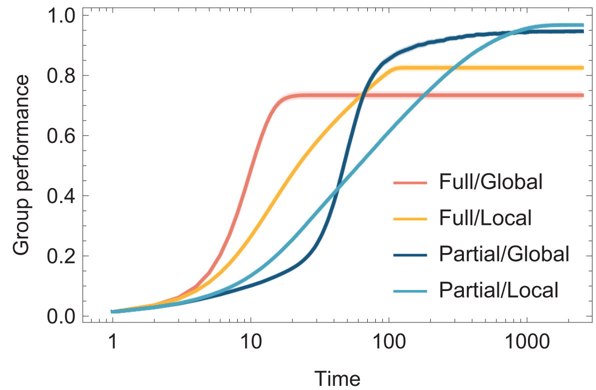

Figure 2: Group performance over time. Time shown on a log scale because most of the learning occurs in the initial period. For simplicity, we only visualize an intermediate problem difficulty (K=6). Each trace represents group performance over time, averaged over 1,000 repetitions, for full and partial, global and local conditions. Shaded area represents standard error around the mean.

Learning choice dynamics. To answer the question of why partial copying outperformed full-copying, we examined how the learning choices in the group changed over time (Fig. 3). In the full-copying conditions, individual learning and copying disappeared quickly; individuals were rapidly losing opportunities to improve solutions. One problem was that copying too early, especially when copying is full, leads to a quick loss of diversity of solutions in the population. As more of an individual's neighbors shared their same solution, opportunities for copying diminished. As these groups converged prematurely on local optima, local improvements became scarce so individual learning too slowed to a stop. In the partial copying conditions, both forms of learning were maintained for longer and in larger proportions of the group. For one, partial copying does not lead to the pronounced diversity loss of full-copying, so the local partial copying group still did well despite copying early on. The global partial copiers did even better by delaying copying and instead exploring more of the problem space before exploiting the best solutions they had found so far.

Within the partial copying condition, it is interesting that there was more copying and individual learning for global groups than local ones. For local groups, the solutions in a region in the social network all began to look similar as individuals could only copy from their immediate neighbors or search locally. As a result, it became less likely that their neighbors would have better solutions for them to copy. The individuals in the global condition did not have this disadvan-

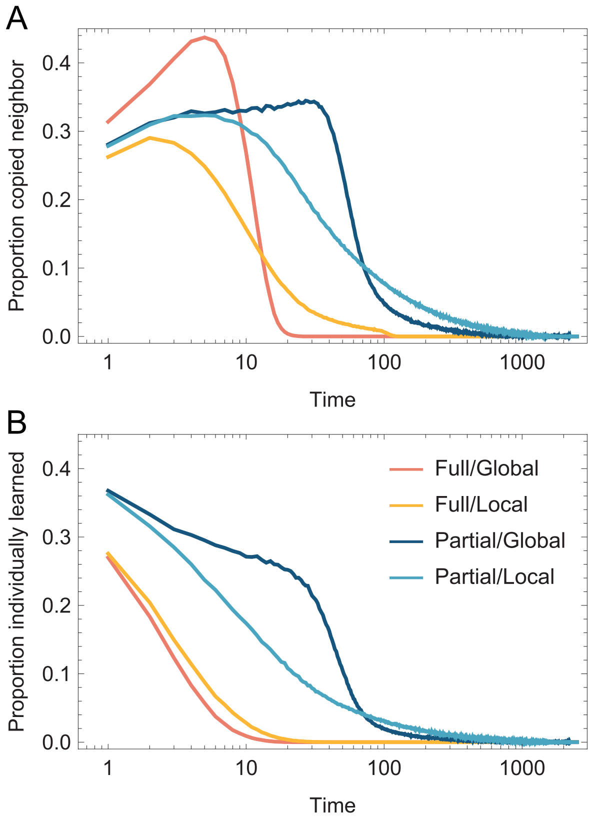

Figure 3: Learning choice dynamics (K=6). [A] Proportion of individuals in the population that learned by copying a neighbor. [B] Proportion of individuals in the population that learned individually. Note that these two measures do not sum to 1 because when neither copying nor individual learning led to an improvement, the individual did not learn during that trial. As the population converged, both types of learning diminished. Each trace represents group performance over time, averaged over 1,000 repetitions, for full and partial, global and local conditions.

tage, as they could potentially copy from anyone in the population, increasing their odds of choosing to copy from a better neighbor. This also explains why the difference in copying frequency between the two partial conditions was absent until about 10 learning trials had passed. All the solutions were random at the beginning so the chances of choosing a better-performing neighbor were the same for both groups until the local neighborhoods started to conform. Global groups not only copied more, but also learned more than local groups when there was partial copying. Movement is limited in local groups, causing many individuals to end up stuck on local optima where learning stalled. Individuals in global groups, who could potentially move anywhere in the problem space by partially copying a distant solution, could discover new regions with many more opportunities to individually learn.

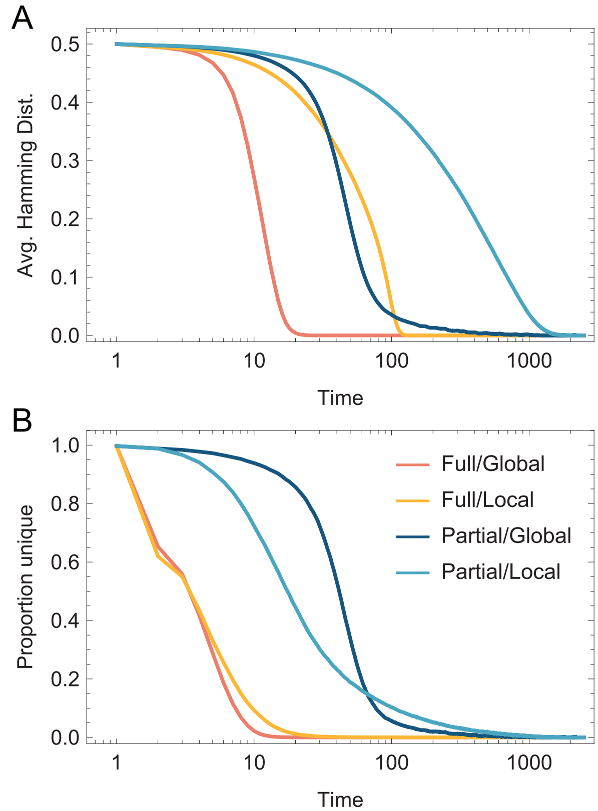

Figure 4: Diversity of the individuals in the group (K=6). [A] Average Hamming distance between every pair of individuals in the group. [B] Proportion of unique solutions in the group. Each trace represents group performance over time, averaged over 1,000 repetitions, for full and partial, global and local conditions.

Diversity: different forms of exploitation and exploration. To explain the difference in performance between these social strategies, we considered the maintenance and generation of diversity in the population and their explore-exploit dynamics. Individual learning corresponds to local exploration as individuals take small steps from their current position. Copying entire solutions corresponds to exploitation. Partial copying, however, amounts to a unique blend of exploration and exploitation. Mixing two solutions can be interpreted as exploitation of the solution elements being copied, but also exploration of a greater set of potential solutions than are reachable in a single step of individual learning. Because mixing solutions can position an individual in a previously unoccupied region of the problem space, it carries the potential of broadening the pool of unique solutions in the population.

In order to characterize the diversity of solutions in these groups in the different conditions, we examined two complementary measures: the overall spread of the solutions in space using the average Hamming distance between every pair of individuals in the group (Fig. 4A) and the proportion of solutions that were unique in the group (Fig. 4B). In both the full and partial copying conditions, the distance between individ-

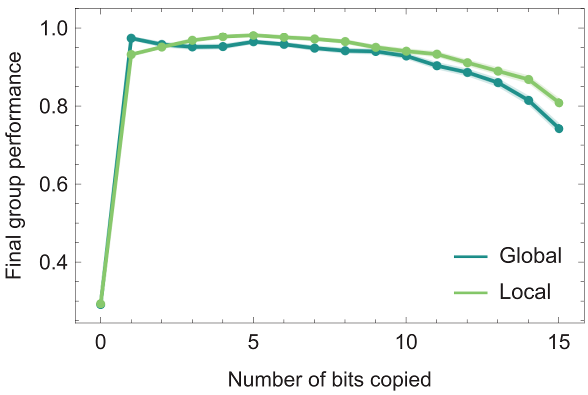

Figure 5: Final group performance as a function of the number of bits copied (K=6). We consider global and local conditions. Each point represents final group performance after 2,500 learning steps, averaged over 1,000 repetitions. Shaded area represents standard error around the mean.

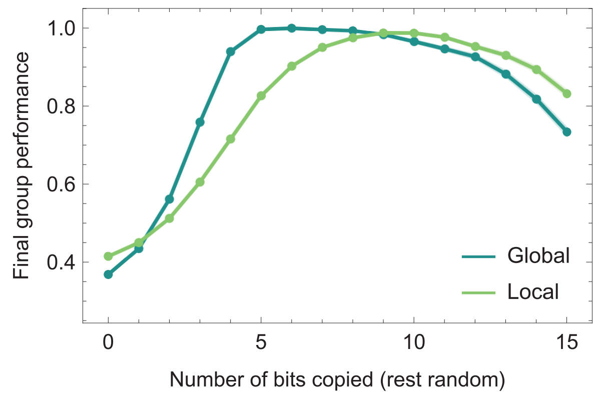

Figure 6: Final group performance as a function of the number of bits copied when the rest of the bits are randomized (K=6). We consider global and local conditions. Each point represents final group performance after 2,500 learning steps, averaged over 1,000 repetitions. Shaded area represents standard error around the mean.

uals in the global groups became smaller faster than in the local groups. This effect is consistent with the explanation above: global groups lose diversity more rapidly than local ones. Although the Hamming distance is useful to capture the size of dispersion of the solutions in the group, it can lead to misleading comparisons if the groups have different levels of clustering around groups of solutions. The proportion of unique solutions in the group allowed us to further characterize the diversity in these different conditions (Fig. 4B). Although the average Hamming distance was different for the global and the local conditions in the full-copying conditions, their proportion of unique solutions was similar. In other words, both groups were clustered around a similar number of unique solutions, but the clusters in the local groups were spread out over a wider range of the space. In the partial copying conditions, the effect was different. Although the local group was dispersed over a wider range of the space (as shown by the Hamming distance), it was the global group that had the highest proportion of unique solutions. The combination of global-connectedness and partial copying conferred the highest amount of diversity to the population. Copying mixes solutions that are less related when the network is global rather than local. In the local condition, copying mixes similar solutions. It was also useful to compare the full/local and the partial/global conditions. The average Hamming distance between them was similar, but the proportion of unique solutions was different: much higher for the partial/global. So although they were dispersed in similarly-sized clouds in the problem space, the full/local was highly clustered around a small group of unique solutions, while the partial/global was more widely spread within the same space.

How much to copy from others?

In the previous section, we demonstrated that strategies where individuals copy 50% of a neighbor's solution outperform strategies where the individual is forced to copy the neighbor's entire solution. However, how much copying is best for the group? In order to better understand this, we examined the performance of partial copying while we systematically varied the amount of bits that were copied from the better performing neighbor (Fig. 5). In the global groups, copying a single bit from the better performing neighbor improved the social learning strategy dramatically relative to no copying. Each additional bit copied led 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 was a similar effect, except that copying a few more bits was useful. This is because nearby solutions are more alike, so there is not as much disruption as in the global condition. The qualitative shape of this curve was the same for different problem difficulties (Ks); what changed was the overall group performance (higher performance for simpler problems and vice versa). In summary, copying some of your neighbor's solution is essential, but copying too much of their solution can be detrimental. Also, local groups can copy more of their neighbor's solution before incurring a detriment.

Mixing solutions or copying with noise?

Although we hypothesized that the advantage observed in the partial copying groups was due to mixing two solutions, it is possible that it simply came from adding random noise to an already good solution. To investigate this, we evaluated a new condition where, when copying, agents created a new solution by combining a better-performing neighbor's solution with random bits (Fig. 6), as opposed to normal partial copying where they combined their neighbor's solution with their own. When only a few bits were being copied (1-4 in the global case; 1-7 in the local case), mixing the current solution with a better one was the best strategy. However, when many bits were being copied from the better-performing neighbor, discarding the current solution for random bits was the best strategy. As expected, when nearly all bits were being copied, there was little or no difference between mixing solutions or copying with noise. Also, mixing solutions was better for local than global networks because of the high similarity of solutions in the local region of the social network. Local copying is more likely to produce viable solutions that preserve important interactions between bits, while mixing solutions in the global case carries the risk of breaking important interactions because very different solutions could be blended.

Discussion

Social learning has been widely shown to be advantageous over individual learning but can lead populations to converge prematurely on sub-optimal solutions. Social learning is an explore-exploit problem; agents must balance exploring new solutions and exploiting good solutions. Recent studies have searched for a way to balance this trade-off and reap the benefits of social learning while avoiding the drawbacks. These proposed methods are successful because they help the group maintain the global diversity of solutions longer. Populations will eventually converge on a solution, but strategies that maintain diversity keep populations exploring longer and allow more of a chance for the population to converge on the *optimal* solution. For this reason, studying the long-term performance of social learning strategies is critical. In previous work, simulations had been run for somewhere between 100 and 200 learning steps (Lazer & Friedman, 2007; Barkoczi & Galesic, 2016), but this was not long enough to show the asymptotic behavior of the social learning strategies. The dynamic aspects of the strategies should be taken into consideration as seriously as their asymptotic performance, as this helps demonstrate the diversity-maintenance ability of a given strategy.

Network topology. One of the most explored suggestions for maintaining diversity is to mediate information flow by embedding learners in inefficient networks (Derex & Boyd, 2016; Lazer & Friedman, 2007) or clustered networks, a special type of inefficient network (Fang et al., 2010). Derex and Boyd (2016) found that individuals in efficient networks tend to copy successful individuals more often, and Wisdom, Song, and Goldstone (2013) found that people are also more influenced by higher-performing individuals, copying larger parts of their solutions. Copying is beneficial, but as we have shown, too much copying cannot maintain the diversity of solutions in the population as well and can cause the group to get stuck on a sub-optimal peak. Embedding the group in an inefficient network (such as the local network we used) can prevent this by slowing down information propagation, but diversity maintenance via inefficient networks may not always be the best solution.

The complexity and size of the search space affects the performance of efficient and inefficient networks. Mason, Jones, and Goldstone (2008) found that efficient networks perform the best in a unimodal problem space – one with a single global optimum and no local optima. In a multimodal problem space, inefficient networks performed best. This result is confirmed by Mason and Watts (2012), who found that efficient networks resulted in better group performance for their comparatively limited problem space. For simple problem spaces, which have fewer local optima to get stuck in, maintaining diversity is less important compared to quickly propagating good solutions, which is a strength of efficient networks. It follows, then, that when considering how to connect a group, the difficulty of the problem should be considered (Goldstone, Wisdom, Roberts, & Frey, 2013).

Studies have demonstrated that one should also take into account the individuals' strategy when considering how to connect a group (Barkoczi & Galesic, 2016; Barkoczi, Analytis, & Wu, 2016). They found that inefficient networks are best when individuals copy the best solution of their neighbors, but efficient networks are best when individuals copy the most frequent solution. Both of these social learning strategies are common human biases (Derex et al., 2015; Derex & Boyd, 2016; Heyes, 2016; Kendal et al., 2018; Wisdom et al., 2013), but we chose to consider only the best member strategy in our model because in a large search space, an individual's neighbors are likely to all have unique solutions, making the conformity strategy unfeasible.

Alternate diversity-maintenance strategies. Inefficient networks are not the only way to maintain solution diversity in complex problem spaces. Other strategies preserve diversity by mitigating a problematic bias: humans tend to do more copying initially, when they are unfamiliar with a problem and uncertain what a good solution may be (Wisdom et al., 2013). Copying early can cause a population to get stuck on a sub-optimal peak; it is much better to explore early and exploit later (Yahosseini, Reijula, Molleman, & Moussaid, 2018). One suggestion is to limit social influence by having intermittent breaks, during which participants must explore solutions on their own (Bernstein, Shore, & Lazer, 2018). We have demonstrated the benefits of copying only parts of solutions. While partial imitation has been ubiquitous in empirical work, its advantages over perfect imitation have not been fully explored.

Imitation. Although copying might seem to *increase* conformity and decrease solution diversity, it actually maintains diversity longer than individual learning and facilitates innovation (Wisdom & Goldstone, 2011; Derex & Boyd, 2016). One reason is that humans are not perfect imitators. We copy erroneously (Caldwell et al., 2016) and we do not copy completely, adapting information from others (Derex et al., 2015). These mutations and adaptations can lead to the discovery of new solutions, which explains why social learners end up exploring more of a search space than individual learners (Derex et al., 2015). Some simulations of social learning have added noise to copied solutions to model imitation error (Goldstone et al., 2013; Rendell et al., 2010), but this is distinct from partial copying. Partial copying by our definition results in the combination of two existing solutions in the population. Partial copying and imitation error are similar, but their subtle difference is important as we have shown that when copying a better solution, keeping a little of one's own solution is better under some conditions than adding the same amount of random noise. However, our partial copying is not completely absent from simulation work. Fang, Lee, and Schilling (2010) and Lazer and Friedman (2007) included a version of partial copying in their models, although they called it "the proclivity or ability of individuals to learn from one another" and "error in copying", respectively. In both models, simulated individuals copied a component from another individual's solution with a probability, resulting in a new solution that was a combination of the two solutions. Lazer and Friedman found that introducing this "error" increased long run performance compared to perfect copying because it expanded the set of potential solutions available to the copier. In this research, we have replicated and expanded on this finding.

Conclusions

Work on social learning has primarily focused on three questions: when to copy, who to copy from, and what to copy (Kendal et al., 2018). In this paper, we focus on *how much* to copy from others, considering the benefits for the group when individuals copy better solutions only partially. We highlight our key findings. First, partial copying benefits the group, even though it comes at a risk to the individual; copying an entire solution guarantees improvement for the copier, but the mix that results from copying partially does not. Second, the network of communication does not have as large an effect on group performance for partial compared to full copying strategies. Third, copying some bits is essential for group performance; copying too many bits is detrimental to the group, yet local groups can afford to copy more before incurring in the detriments. Finally, partial copying strategies allow for different forms of exploration and exploitation than copying entire solutions or copying with noise. Our findings suggest additional experimental work to systematically study the effect of partial copying on group performance.

Data availability

Simulation code and relevant data files are available at: https://github.iu.edu/EASy/CampbellCogSci2020.

Acknowledgments

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.

References

Barkoczi, D., Analytis, P., & Wu, C. (2016). Collective search on rugged landscapes: A cross-environmental analysis. *Proc. Cogn. Sci. Soc. Conf.*, 918-923.

Barkoczi, D., & Galesic, M. (2016). Social learning strategies modify the effect of network structure on group performance. *Nat. Commun.*, *7*, 13109.

Bernstein, E., Shore, J., & Lazer, D. (2018). How intermittent breaks in interaction improve collective intelligence. *Proc. Natl. Acad. Sci. USA*, *115*(35), 8734–8739.

Caldwell, C., Cornish, H., & Kandler, A. (2016). Identifying innovation in laboratory studies of cultural evolution: Rates of retention and measures of adaptation. *Philos. Trans. R. Soc. B.*, *371*.

Derex, M., & Boyd, R. (2016). Partial connectivity increases cultural accumulation within groups. *Proc. Natl. Acad. Sci. USA*, *113*(11), 2982–2987.

Derex, M., Feron, R., Godelle, B., & Raymond, M. (2015). Social learning and the replication process: An experimental investigation. *Proc. R. Soc. B.*, *282*.

Fang, C., Lee, J., & Schilling, M. (2010). Balancing exploration and exploitation through structural design : The isolation of subgroups and organizational learning. *Organ. Sci.*, *21*(3), 625–642.

Goldstone, R., Wisdom, T., Roberts, M., & Frey, S. (2013). Learning along with others. In *Psychology of Learning and Motivation* (Vol. 58, p. 1–45). Academic Press.

Heyes, C. (2016). Who knows? Metacognitive social learning strategies. *Trends Cogn. Sci.*, *20*(3), 204–213.

Kauffman, S., & Weinberger, E. (1989). The NK model of rugged fitness landscapes and its application to maturation of the immune response. *J. Theor. Biol.*, *141*(2), 211–245.

Kendal, R., Boogert, N., Rendell, L., Laland, K., Webster, M., & Jones, P. (2018). Social learning strategies: Bridge-building between fields. *Trends Cogn. Sci.*, *22*(7), 651–665.

Lazer, D., & Friedman, A. (2007). The network structure of exploration and exploitation. *Adm. Sci. Q.*, *52*(4), 667–694.

Mason, W., Jones, A., & Goldstone, R. (2008). Propagation of innovations in networked groups. *J. Exp. Psychol. Gen.*, *137*(3), 422–433.

Mason, W., & Watts, D. (2012). Collaborative learning in networks. *Proc. Natl. Acad. Sci. USA*, *109*(3), 764–769.

Rendell, L., Boyd, R., Cownden, D., Enquist, M., Eriksson, K., Feldman, M., & Laland, K. (2010). Why copy others? Insights from the social learning strategies tournament. *Science*, *328*(5975), 208–213.

Wisdom, T., & Goldstone, R. (2011). Innovation, imitation, and problem-solving in a networked group. *Nonlinear Dynam. Psychol. Life Sci.*, *15*(2), 229–252.

Wisdom, T., Song, X., & Goldstone, R. (2013). Social learning strategies in networked groups. *Cogn. Sci.*, *37*(8), 1383–1425.

Yahosseini, K., Reijula, S., Molleman, L., & Moussaid, M. (2018). Social information can undermine individual performance in exploration-exploitation tasks. *PsyArXiv Preprint*.