

Models and Anti-Models:
The Structure of Payoff-Dependent Social Learning*

Charles Efferson^{1,2,4}, Rafael Lalive^{2,3}, Peter J. Richerson^{4,5}, Richard
McElreath^{4,6}, and Mark Lubell^{4,5}

¹Santa Fe Institute

²Institute for Empirical Research in Economics, University of Zürich

³Department of Econometrics and Political Economy, University of
Lausanne

⁴Graduate Group in Ecology, University of California, Davis

⁵Environmental Science and Policy, University of California, Davis

⁶Anthropology, University of California, Davis

November 7, 2006

Running title: Payoff-dependent social learning

Word count: \approx 3500

***Corresponding author:** Charles Efferson, Santa Fe Institute, 1399 Hyde Park Rd., Santa Fe, NM 87501, *cmefferson@santafe.edu*.

Abstract: Theory shows that the structure of social learning should be critically important to the evolution of norms and behaviors in human societies. Imitating success is a type of payoff-dependent social learning that has received considerable attention, but other types of payoff-dependent social learning are also feasible. Avoiding unsuccessful behaviors, for example, can have theoretical advantages under certain circumstances. We conducted an experiment that allowed players to choose between learning about either successful or unsuccessful behaviors. In addition, the experiment allowed us to describe precisely how players used this social information by separating individual learners and social learners. Players chose between two technologies repeatedly. Payoffs were random, but one technology had a higher expected payoff. Individual learners knew their realized payoffs after each choice, while in each period social learners could only choose to know one of the following: a) the technology producing the highest payoff among individual learners or b) the technology producing the lowest payoff. Social learners never knew their realized payoffs. When social learners chose to know the technology producing the highest payoff, a model of imitating this successful behavior matches the data very closely. When social learners chose to know the technology producing the lowest payoff, which they did roughly a quarter of the time, they tended to choose the opposite technology in early periods, while increasingly choosing the same technology in late periods. This use of social information improved the performance of social learners relative to simply randomizing over the two technologies, but it was not theoretically optimal.

Key words: cultural evolution, payoff-dependent social learning, experimental games

1 Introduction

In the study of social learning, a fundamental hypothesis is that humans do not learn from other humans in a completely random way (Boyd and Richerson, 1985, 2005; Henrich and McElreath, 2003; Richerson and Boyd, 2005). Instead people discriminate when deciding with whom to associate and how to use the information embedded in their social group. Simple forms of nonrandom social learning, via feedbacks between the individual and group levels, can produce complex behavioral dynamics. In particular, theory and experiments suggest that nonrandom social learning can lead to multiple equilibria (Anderson and Holt, 1997; Banerjee, 1992; Baum *et al.*, 2004; Bikhchandani *et al.*, 1992; Boyd and Richerson, 1992; Bowles, 2004; Salganik *et al.*, 2006), extreme behaviors (Boyd and Richerson, 1985), cultural group selection (Henrich, 2004), ethnically marked social groups (Boyd and Richerson, 1987; McElreath *et al.*, 2003), and occasionally even maladaptive behaviors (Boyd and Richerson, 1985; Richerson and Boyd, 2005).

To say that social learning is not random, however, is incomplete because even small differences between nonrandom forces at the individual level produce radically different predictions at the aggregate level (Efferson and Richerson, 2006). Controlled experiments offer a way to characterize the relevant forces precisely. One type of nonrandom social learning that has received considerable attention postulates that individuals tend to imitate or reproduce behaviors that were successful in the past (Offerman and Sonnemans, 1998; Henrich and Gil-White, 2001; Camerer, 2003; Offerman and Schotter, 2005). Indeed, all of evolutionary game theory rests on a particular version of this postulate (Weibull, 1995; Gintis, 2000; Bowles, 2004). Imitating success, however, can take various forms. One could, for example, adopt the behavior estimated to have the highest mean payoff in the population. One could alternatively adopt the behavior that most recently produced the highest payoff in the population. Moreover, completely different types of payoff-dependent social learning are possible. One could use behaviors that have previously produced low payoffs as negative examples (i.e. behaviors to avoid). Here we present an integrated approach that combines theory and experiments and allows us to identify some of the key features of payoff-dependent social learning. Specifically, we address dramatically different types of payoff-dependence:

imitating successful behaviors and avoiding unsuccessful behaviors.

In the electronic supplement we develop a model of imitating successful behaviors and avoiding unsuccessful behaviors that formalizes the relationship between these two types of social learning and specifies when one is better than the other. As in the experiment we describe shortly, the setting involves two behaviors or “technologies.” Technology is an abstract term here and refers to anything (e.g. a hunting technique) that transforms inputs (e.g. time) into payoffs (e.g. food). Payoffs are stochastic, but one technology is optimal because its payoff distribution has a higher mean. Individuals know this, but they do not know with certainty which technology is best. In this situation, using information about the relationship between payoffs and choices can filter noisy feedback at the individual level into a valuable social signal.

This model has two conclusions relevant to the experiment described below. Namely, the value of avoiding unsuccessful behaviors versus imitating successful behaviors depends on both the shapes of the payoff distributions and interactions with other decision-making forces. To illustrate how the shapes of the distributions can matter, consider a situation in which the optimal behavior has a symmetric distribution with a small variance, while the sub-optimal distribution is right-skewed with a large variance. This scenario could apply in various winner-take-all settings. For example, if facing a choice between pursuing a career as a rock star or going to medical school, the market for aspiring rock stars will tend to be winner-take-all in the sense of Frank and Cook (1995). A few young rock-star hopefuls will be wildly successful far beyond anything an ordinary doctor can achieve, but most will bomb miserably. In short the *ex ante* payoff distribution for those struggling to be the next pop idol has a low mean and a fat right tail. Going to medical school, though relatively mundane and without potential for the adoration of countless teens, brings a higher expected payoff. In this case, imitating success (e.g. Britney Spears) is on average a disaster, while avoiding failure (e.g. the unemployed rock drummer you met at a retirement party last week-end) is all in all a good use of payoff-dependent social information.

To see how other decision-making forces matter, consider two payoff distributions that are identical except for their expectations. Now imagine some force is biasing choices toward the sub-optimal behavior (e.g. the optimal behavior is

a recent innovation in a population of conformists). In this case avoiding the behavior of the most unsuccessful individual in the social group is better than imitating the most successful simply because both individuals probably exhibit the sub-optimal behavior. Alternatively, if some force (e.g. individual learning) is biasing choices toward the optimal behavior, then imitating success is better for the opposite reason. When no additional forces bias choices one way or another, the two forms of payoff-dependent social learning are equivalent on average.

The intuition behind this result is straightforward. Imagine everyone chooses the sub-optimal technology in a period. In this case choosing the technology that did not produce the worst payoff is certain to identify the optimal technology in the next period. Analogously, when everyone chooses the optimum, choosing the technology that produced the best payoff is certain to identify the optimum. Exactly the same reasoning applies when individuals in the social group exhibit a mix of choices, but the effects are weaker. In particular, under a mix of choices either the optimal technology or the sub-optimal technology can produce the highest payoff and/or the lowest payoff in a social group. All combinations are possible, but the probability distributions over the various outcomes are typically not uniform. This is what gives payoff-dependent social learning its value. Importantly, however, if a bias toward the sub-optimum is initially present, avoiding unsuccessful behaviors increases the frequency of optimal choices through time. This reduces the bias toward the sub-optimum and correspondingly reduces the value of avoiding unsuccessful behaviors as a social learning algorithm. Imitating success does not have this problem.

2 Experimental methods

With 72 students at the University of Zürich and the Swiss Federal Institute of Technology, we conducted the following experiment. In each period each player faced an individual choice between one of two technologies, which in this case were simply the colors “red” and “blue.” Payoffs followed normal distributions, but one color was optimal in that its payoff distribution had a higher expectation. Players did not know which color was better. They made choices for multiple blocks of 25 periods each. Each block of 25 periods had a randomly selected

optimal color, but all players who played together had the same optimum. All of this was explained in the instructions before beginning a experimental session. In addition, participants viewed an extensive demonstration before the beginning of the experiment. The demonstration produced several animated histograms that gave subjects an intuition for how the random payoffs were generated, even if they did not have a formal understanding of probability theory. The supplement provides additional details.

Players within a session were divided into two groups that played simultaneously. In one group of 5 players, each player individually chose one of the two colors in each period and immediately received private information about her realized payoff. These players did not have any information about other players, and thus we refer to them as *individual learners*. In the other group, composed of 7 *social learners*, in each period each player chose to access one of two types of social information: 1) the color that produced the lowest payoff among the group of individual learners or 2) the color that produced the highest payoff. This information was available after all individual learners had made their choices in a period but before a given social learner had made her choice. After communicating the social information, each social learner made a choice between the two colors privately and received a payoff. Realized payoffs, however, were never communicated to players in this group, and thus individual learning was not possible. The entire experiment was conducted on a local computer network using z-Tree (Fischbacher, 1999).

If a given social learner chose to learn which color produced the highest payoff among individual learners, and then she chose the same color, this can be viewed as one form of imitating success. If she chose to learn about the color producing the worst payoff and then chose the opposite color herself, this can be viewed as using failure as a negative behavioral example. Although this latter type of social learning has received little attention, it could be related to the finding that experimental subjects show an extreme aversion to losses (Kahneman and Tversky, 1979) in that avoiding unsuccessful behaviors could limit one's vulnerability to losses from the status quo. Avoiding unsuccessful behaviors could also be related to situations in which losses are catastrophic, a concept formalized as risk dominance in game theory (Bowles, 2004). In addition to imitating success and

avoiding unsuccessful behaviors, social learners in the present experiment could also imitate the worst or do the opposite of the best. We discuss these possibilities below and in the supplement. To simplify the discussion, we will refer to imitating or doing the opposite of the best (IB or DOB) and imitating or doing the opposite of the worst (IW or DOW). Figure 1 summarizes the probabilities with which these four types of social learning identify the optimal color.

3 Results and Discussion

As we show in the supplement, the respective values of imitating the best and doing the opposite of the worst depend on interactions with other forces. In the experiment, the force most important is individual learning. If individual learning is effective, and thus individual learners increasingly choose the optimal color as time passes, imitating the best will correspondingly become more effective than doing the opposite of the worst. As Figure 2 shows, individual learning was effective. Individual learners chose the two colors with approximately equal probability in early periods, but the proportion of individual learners choosing the optimal color increased through time.

In the present experiment, social learners could do the opposite of the worst, imitate the worst, do the opposite of the best, or imitate the best. Moreover, a given social learner's choice in a given period was either optimal or sub-optimal. This scheme provides 8 categories for the the choices of social learners. Figure 3 summarizes how choices were distributed among these 8 categories. The most common outcome by far was a social learner imitating the best and making an optimal choice as a result. The second most common outcome was doing the opposite of the worst and making an optimal choice. Interestingly, when accessing the color of the worst payoff, social learners also imitated the worst at a relatively high rate in a way that systematically yielded optimal choices. We return to this result shortly.

The model in the supplement provides conditional theoretical predictions about the optimality of choices by players who imitate the best or do the opposite of the worst. These predictions are summarized in Figure 1. Specifically, if social learners make a large number of choices by imitating the best, for any given proportion

of individual learners choosing optimally, the resulting choices of social learners will be optimal in a way described precisely by the imitate-the-best curve in Figure 1. Analogously, if social learners make a large number of choices by doing the opposite of the worst, the resulting choices will be optimal in a way described precisely by the do-the-opposite-of-the-worst curve. This is so because the curves in Figure 1 are simply summary statements about how the stochastic process in the experiment works. So long as the number of observations is adequately large, deviations from these predictions can only occur because players sometimes do the opposite of the best or imitate the worst.

Figures 4 and 5 show the theoretical predictions and associated data. Figure 4 shows that a model of imitating the best matches actual play in the experiment closely given player-period combinations when the social learner chose to learn about the color producing the best payoff among individual learners. This tells us that social learners in these situations exhibited a strong tendency to imitate the best, which we also see in Figure 3. Occasionally players did the opposite of the best, and these choices are responsible for the small deviations between theory and data.

In sharp contrast, Figure 5 shows that when the social learner wanted to know the worst color, theory and data do not match at all. The reason is because social learners showed a marked tendency to imitate the worst. To evaluate this finding, note that if individual learners are choosing the two colors with roughly equal probability, as in early periods, doing the opposite of the worst is better than imitating the worst. As individual learners accumulate information about which color is best, however, and thus focus on choosing that color, imitating the worst actually becomes better than doing the opposite of the worst. We discuss this further in the supplement, but the basic idea is intuitive. If almost all individual learners are choosing the optimum, any imitative strategy is better than using failure as a negative behavioral example, even if one imitates failure. If a social learner makes these kinds of assumptions about individual learners, she should switch from doing to opposite of the worst to imitating the worst given that she has chosen to access the color producing the worst payoff. As the social learner accumulates prior information from earlier periods, and thus the marginal value of social information declines with each new period, this could also accentuate the

inclination to switch to imitating the worst.

When to switch, however, is a complex problem. If social learners vary in terms of their assumptions about individual learners, their switching points will vary, and this will lead to a thorough mix of imitating the worst and doing the opposite of the worst. In particular, imagine that social learners accessing the worst color tend on average to switch at some point from doing the opposite of the worst to imitating the worst, but a notable mix of the two strategies is generally present. In early periods, when optimal choices are more likely to be at low frequencies among individual learners, imitating the worst will suppress optimal choices among social learners relative to simply doing the opposite of the worst. In later periods, however, as optimality increases among individual learners, imitating the worst will actually encourage optimal choices relative to doing the opposite of the worst. The resulting mix of social learning strategies in all periods will flatten the imitation curve, as shown in Figure 5. If the trends governing a switch from doing the opposite of the worst to imitating the worst are strong enough, however, choices will be biased toward the optimal color relative to simply randomizing over the two colors with equal probability. One can also see this in Figure 5; social learners accessing the worst color did better than simply randomizing.

Figure 6 shows the absolute proportions of social learners doing the opposite of the worst and imitating the worst by period. Trends in choices follow the pattern described above. In early periods, doing the opposite of the worst is more common than imitating the worst. The tendency to do the opposite of the worst, however, falls significantly, and the tendency to imitate the worst rises significantly. By the final period, social learners are doing each in roughly equal proportions.

This discussion, however, does not address why social learners wanted to know about unsuccessful behaviors in the first place. Altogether, social learners used social information to improve their rate of choosing optimally relative to simply randomizing over the two colors with equal probability. Using a mean proportion of optimal choices for each of the 42 social learners, a simple t test with a null proportion of 0.5 confirms this conclusion ($t_{41} = 14.4085$, $P = 1.2138 \times 10^{-17}$, 95% confidence interval: 0.769 ± 0.038). Nonetheless, in this particular experiment social learners had little reason to access the color producing the worst payoff

among individual learners. To know definitively what type of social information would be better in a given period, social learners would have had to know how many of the 5 individual learners chose optimally in that period. Social learners did not have that information. Consider, however, the following weak assumptions about individual learning, which correspond well with what individual learners actually did in the experiment. Assume individual learners start off choosing each color with equal probability. Through time, they accumulate information based on private feedback about their realized payoffs and thus learn which of the two colors is better. As a result, the probability that individual learners choose the optimal color increases through time at some unspecified rate from an initial probability of 0.5.

These assumptions mean that, in period 1, the choices of individual learners should be symmetrically distributed around the intersection point of the two curves in Figure 1. In this case imitating the best and doing the opposite of the worst are equivalent on average. As soon as individual learning occurs, however, it increases the mass to the right of the intersection point and decreases the mass to the left of the intersection point. Imitating the best is always better than doing the opposite of the worst when this happens. Moreover, imitating the worst also cannot compete with imitating the best. As Figure 1 shows, for any proportion of individual learners choosing optimally in the open interval $(0.5, 1)$, the curve for imitating the best is always farther from a probability of 0.5 than the curve for doing the opposite of the worst. This means that any bias toward the optimal color among individual learners, no matter how small, renders social information about the best color *more definitive* than social information about the worst color. A Bayesian in this situation, for example, would never choose to know about the color producing the worst payoff in any period after $t = 1$; the information is simply less useful.

Actual social learners, however, showed an interest in both successful and unsuccessful behaviors. Most commonly, they chose to learn about successful behaviors, which they then showed a strong tendency to imitate. Nonetheless, a notable proportion of the time social learners relinquished their opportunity to learn about success and instead chose to learn about the behavior that had most recently produced the worst payoff. Social learners, however, did not simply choose

the opposite behavior in this case. Rather they tended to choose the opposite of the worst behavior in early periods, while imitating the worst behavior in late periods. This finding is in contrast to the typical assumption that the way people use social information changes on a time scale much slower than that of behavior itself. All in all, social learners used payoff-dependent social information effectively; they did better than random. The use of social information, however, was not optimal because of the on-going interest in behaviors producing low payoffs even when information about high payoffs was more definitive. This finding suggests the possibility that the use of payoff-dependent social information was algorithmic in some way that was good but not optimal in this experimental setting.

Acknowledgements

The authors would like to thank all the members in the chair of Ernst Fehr at the Institute for Empirical Research in Economics for assistance and comments. A particular thanks goes to Urs Fischbacher for his z-Tree program to produce animated histograms. CE thanks the U.S. Environmental Protection Agency (STAR Fellowship # U-91615601-3) for financial support. This research was funded by a U.S. National Science Foundation Grant (Award # 0340148 to PJR, RM, and ML).

References

- Anderson, L. R. and Holt, C. A. (1997). Information cascades in the laboratory. *The American Economic Review*, **87**(5), 847–862.
- Banerjee, A. V. (1992). A simple model of herd behavior. *The Quarterly Journal of Economics*, **107**(3), 797–817.
- Baum, W. M., Richerson, P. J., Efferson, C. M., and Paciotti, B. M. (2004). Cultural evolution in laboratory microsocieties including traditions of rule giving and rule following. *Evolution and Human Behavior*, **25**(5), 305–326.
- Bikhchandani, S., Hirshleifer, D., and Welch, I. (1992). A theory of fads, fashion, custom, and cultural change as information cascades. *The Journal of Political Economy*, **100**(5), 992–1026.

- Bowles, S. (2004). *Microeconomics: Behavior, Institutions, and Evolution*. New York: Russell Sage.
- Boyd, R. and Richerson, P. J. (1985). *Culture and the Evolutionary Process*. Chicago: University of Chicago Press.
- Boyd, R. and Richerson, P. J. (1987). The evolution of ethnic markers. *Cultural Anthropology*, **2**(1), 65–79.
- Boyd, R. and Richerson, P. J. (1992). How microevolutionary processes give rise to history. In M. H. Nitecki and D. V. Nitecki, editors, *History and Evolution*, pages 178–209. Albany: SUNY Press.
- Boyd, R. and Richerson, P. J. (2005). *The Origin and Evolution of Cultures*. Oxford: Oxford University Press.
- Camerer, C. F. (2003). *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton: Princeton University Press.
- Efferson, C. and Richerson, P. J. (2006). A prolegomenon to nonlinear empiricism in the human behavioral sciences. *Biology and Philosophy*, **in press**.
- Fischbacher, U. (1999). z-Tree: toolbox for readymade economic experiments. IEW Working paper 21, University of Zuerich.
- Frank, R. H. and Cook, P. J. (1995). *The Winner-Take-All-Society: Why the Few at the Top Get So Much More Than the Rest of Us*. New York: Penguin.
- Gintis, H. (2000). *Game Theory Evolving*. Princeton: Princeton University Press.
- Henrich, J. (2004). Cultural group selection, coevolutionary processes and large-scale cooperation. *Journal of Economic Behavior and Organization*, **53**(1), 3–35.
- Henrich, J. and Gil-White, F. (2001). The evolution of prestige: freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. *Evolution and Human Behavior*, **22**(3), 165–196.
- Henrich, J. and McElreath, R. (2003). The evolution of cultural evolution. *Evolutionary Anthropology*, **12**(3), 123–135.

- Kahneman, D. and Tversky, A. (1979). Prospect theory: an analysis of decision under risk. *Econometrica*, **47**(2), 263–292.
- McElreath, R., Boyd, R., and Richerson, P. J. (2003). Shared norms and the evolution of ethnic markers. *Current Anthropology*, **44**(1), 122–129.
- Newey, W. K. and West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, **55**(3), 703–708.
- Offerman, T. and Schotter, A. (2005). Sampling for information or sampling for imitation: an experiment on social learning with information on ranks. Mimeo, New York University.
- Offerman, T. and Sonnemans, J. (1998). Learning by experience and learning by imitating successful others. *Journal of Economic Behavior and Organization*, **34**(4), 559–575.
- Richerson, P. J. and Boyd, R. (2005). *Not By Genes Alone: How Culture Transformed the Evolutionary Process*. Chicago: University of Chicago Press.
- Salganik, M. J., Dodds, P. S., and Watts, D. J. (2006). Experimental study of inequality and unpredictability in an artificial cultural market. *Science*, **311**(10 February), 854–856.
- Weibull, J. W. (1995). *Evolutionary Game Theory*. Cambridge: The MIT Press.

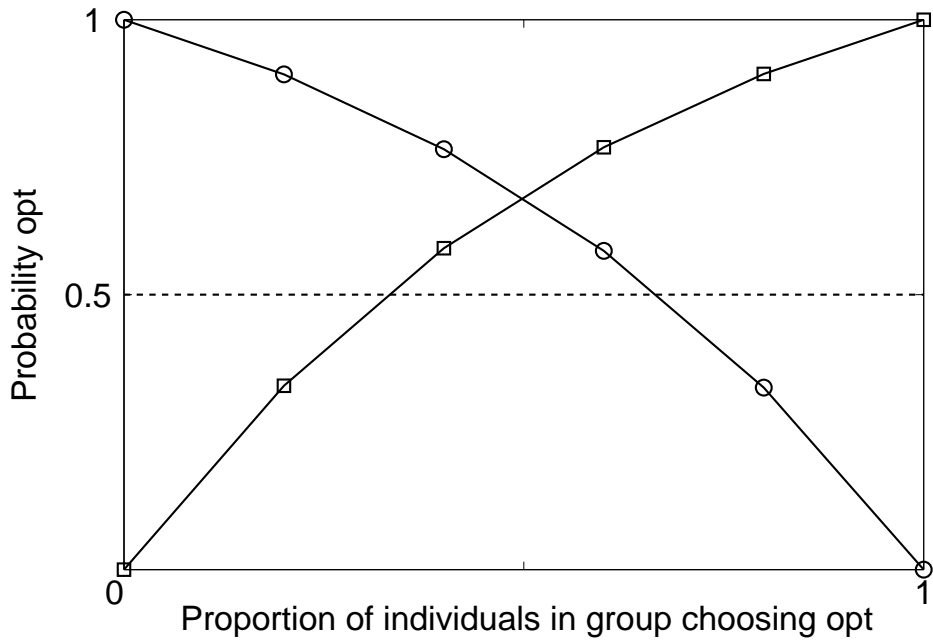


Figure 1: The line with squares shows the probability that imitating the best identifies the optimal color under the specific payoff distributions used in the experiments. The curve is a function of the proportion of individual learners choosing the optimal color. Note that the residual probability (i.e. from the curve upward) is the probability that imitating the best identifies the sub-optimal color and thus the probability that doing the opposite of the best identifies the optimal color. The line with circles shows the analogous probability that doing the opposite of the worst identifies the optimal color. The residual probability in this case is the probability that doing the opposite of the worst identifies the sub-optimal color and thus the probability that imitating the worst identifies the optimal color. As a reference against which to judge the statistical biases the different types of social learning create, a constant probability of 0.5 is also shown.

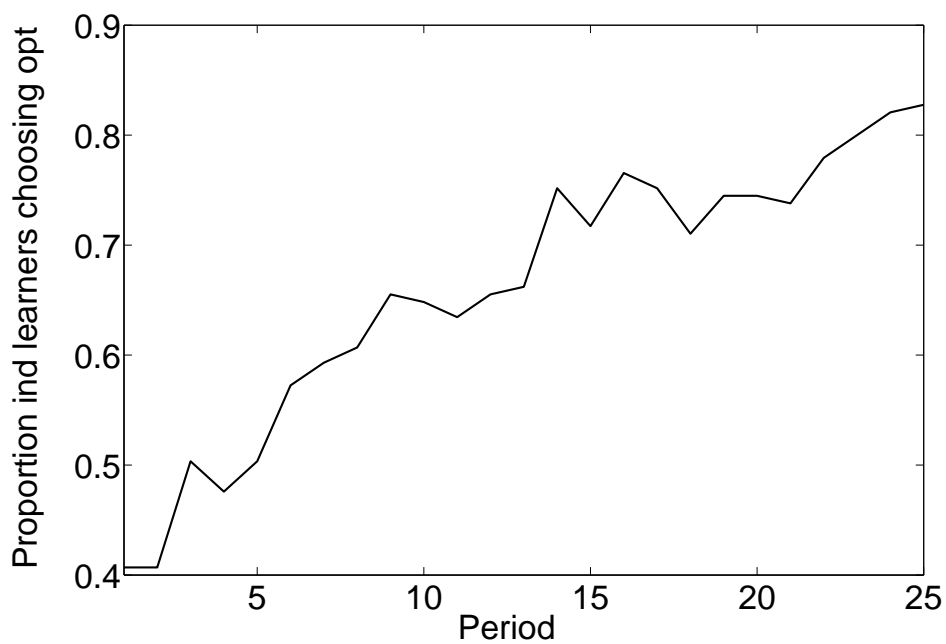


Figure 2: The proportion of individual learners choosing the optimal technology. The upward trend is highly significant ($p < 0.01$) when we regress the proportion choosing optimally on period using the method of Newey and West (1987) to correct for heteroscedasticity and autocorrelation up to lag 3. In this case, the estimated coefficient for period is 0.016 and the R^2 value is 0.903.

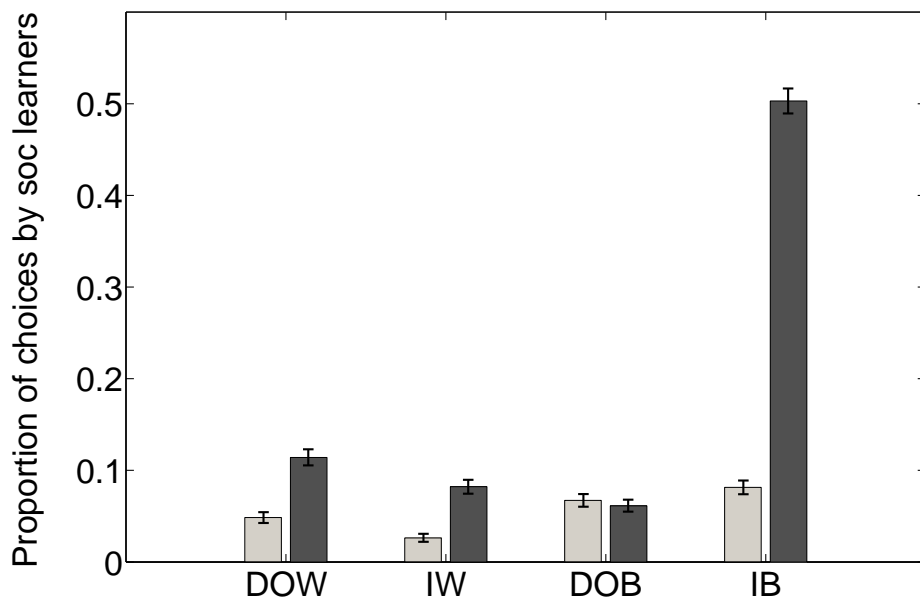


Figure 3: Categorizing choices among social learners. Choices are either optimal (dark gray bars) or sub-optimal (light gray bars), and they can be categorized with respect to the associated group of individual learners as doing the opposite of the worst (DOW), imitating the worst (IW), doing the opposite of the best (DOB), or imitating the best (IB). Error bars represent 95% bootstrapped confidence intervals.

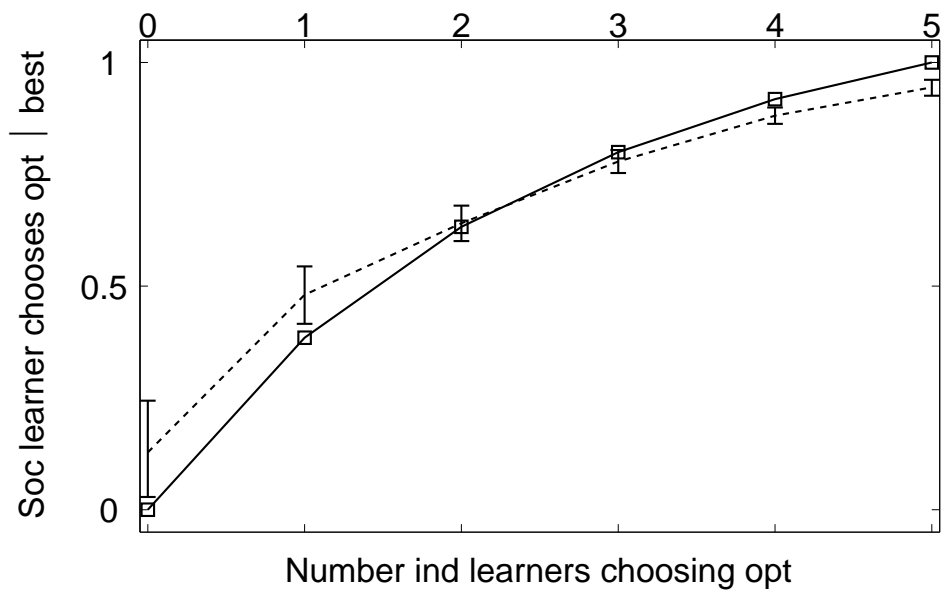


Figure 4: The solid line with squares shows the theoretical probability of an optimal choice for a social learner who accesses the color producing the best payoff among individual learners and then imitates this same color herself. The dashed line shows the actual data (i.e. proportions) with 95% bootstrapped confidence intervals for player-period combinations when the social learner chose to know which color had produced the best payoff among individual learners.

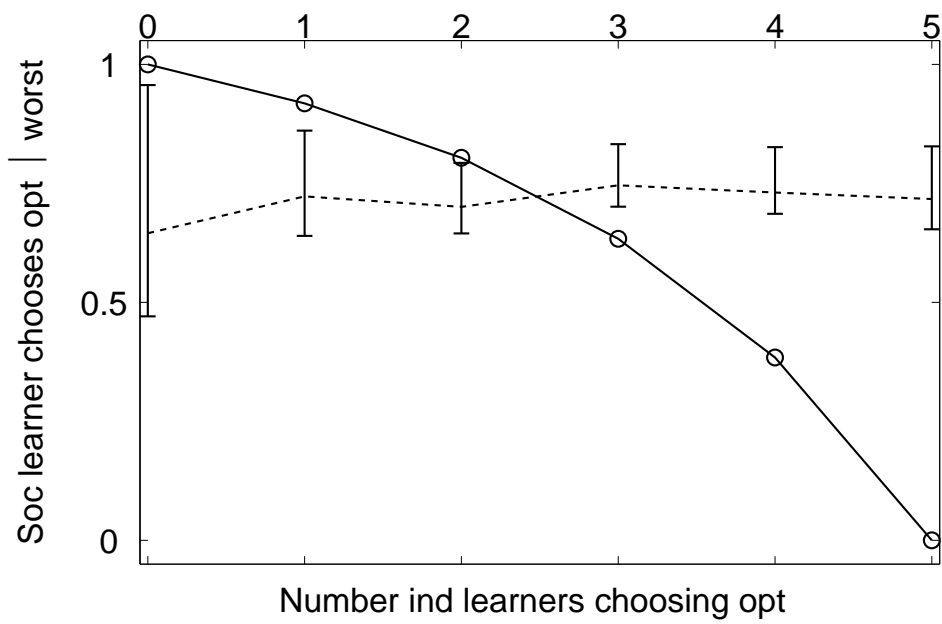


Figure 5: The solid line with circles shows the theoretical probability of an optimal choice for a social learner who accesses the color producing the lowest payoff among individual learners and then chooses the opposite color herself. The dashed line shows the actual data (i.e. proportions) with 95% bootstrapped confidence intervals for player-period combinations when the social learner chose to know which color had produced the worst payoff among individual learners.

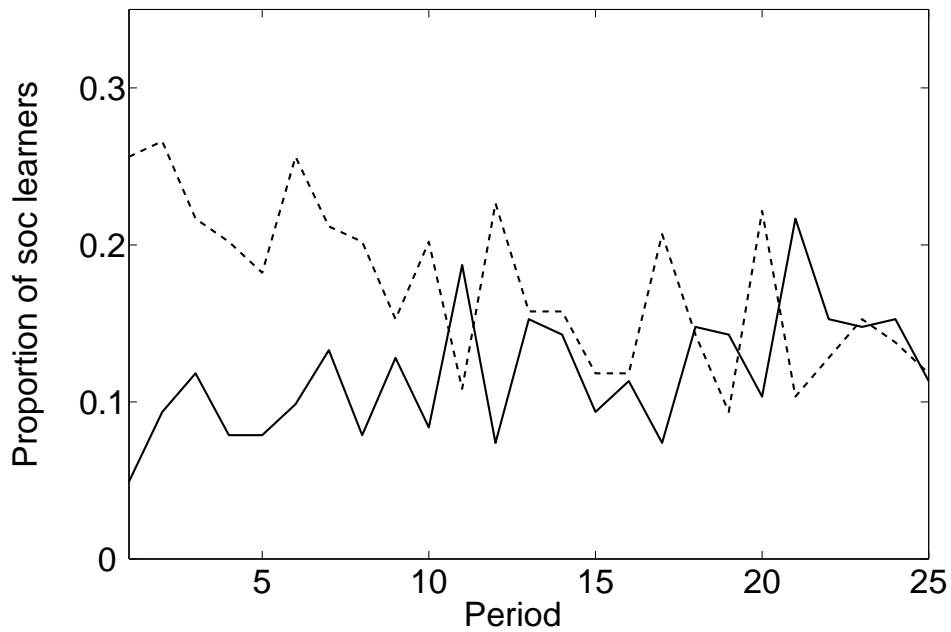


Figure 6: The proportions of social learners either imitating the worst or doing the opposite of the worst by period. In both cases, we regressed the proportion of players using one of the two heuristics against period using the method of Newey and West (1987) to correct for heteroscedasticity and autocorrelation up to lag 3. In the case of doing the opposite of the worst, the trend is significant ($p < 0.05$), and the estimated coefficient for period is -0.004. In the case of imitating the worst, the trend is highly significant ($p < 0.01$), and the estimated coefficient for period is 0.003.

A Experimental methods

The experiments were conducted in the laboratory of the Institute for Empirical Research in Economics at the University of Zürich. Experiments were implemented entirely on a local computer network using the z-Tree software developed by Fischbacher (1999). We recruited a total of 72 undergraduate students from the University of Zürich and the Swiss Federal Institute of Technology. We ran a total of two sessions with 36 students in each session.

Students in each session were divided into 3 “worlds.” Each world consisted of two groups. A group of 5 individuals received individual payoff information as described in detail below. These players were the *individual learners*. A group of 7 players who played simultaneously received no individual payoff information. Instead, in each period each of these *social learners* could choose to learn which of the two colors had produced the lowest payoff among the individual learners or which color had produced the highest payoff.

Subjects were first informed they were participating in a laboratory experiment at the University of Zürich. Communication between subjects was not permitted. Students earned points in the experiment, and 150 points was worth one Swiss franc (about 0.78 USD or 0.64 EUR).

Subjects were instructed that their task was to choose either a “blue” technology or a “red” technology in each period. Technologies generated points at random according to specific random processes. One technology, the optimal technology, had a higher average payoff but was otherwise like the sub-optimal technology. The color of the optimal technology was chosen at random with probability 50%.

The sub-optimal technology generated draws from a normal distribution with an expectation of 30 points and a standard deviation of 12 points. The optimal technology had a normal payoff distribution with an expectation of 38 points and a standard deviation of 12 points. Both distributions were truncated at 0 and 68. Truncation means

that the truncated and untruncated distributions had slightly different expectations and standard deviations. Payoffs were rounded to integer values, and thus the set of possible payoffs for both colors was $\{0, 1, \dots, 68\}$. We did not describe the payoff distributions to subjects in technical terms, but before they began the actual experiment we did provide them with an intuitively accessible demonstration of the random processes governing payoffs.

Specifically, the instructions before the experiment paid particular attention to the possibility that subjects may not have understood the formal concept of payoffs that follow probability distributions. Before the experiment started, subjects saw a demonstration of the two random technologies. In the demonstration, the optimal color was first determined with probability 50%. The optimal color was the same for each subject within a world but potentially different across worlds. Once an optimal color had been determined, the computer would take 250 draws from each of the two probability distributions for each subject individually. Two horizontal number lines from 0 to 68 appeared one beneath the other on the screen with one number line for each color. For each draw producing a specific value (e.g. 27 points for a red choice), the computer would place a little box (colored red or blue) along the appropriate number line (e.g. a red box at 27 for the number line being used to plot red draws). For multiple draws producing the same payoff, boxes were stacked on top of each other. As a consequence, students essentially watched a histogram being built draw by draw on the screen in front of them. This allowed them an intuitive sense of the stochastic process even if they had no training in probability theory or data analysis. Moreover, while the histograms were being built, they knew which color was optimal (but *only* for the demonstration). They could thus see, as an example, that blue was producing payoffs centered around 38, while red was producing payoffs centered around 30. The histogram was explained to them in writing, and subjects could read the explanation repeatedly while the histograms were being built. After the first demonstration, an optimal color was selected, and the

demonstration was repeated. The whole demonstration phase took 5-10 minutes.

After the demonstration had been completed, subjects were informed that one repetition of the experiment would last for 25 periods. The timing was as follows. The computer first assigned the optimal color at random with probability 50%. The optimal color stayed the same for all 25 periods and was identical for each subject within a world. In each period, subjects chose between red and blue by indicating the desired color on the computer screen and clicking “OK.” Immediately after making a choice, individual learners were informed privately about the realized number of points received. They also knew that the points from each period would be added up to yield a total payoff at the end of the entire experimental session.

Social learners were informed of the fact that in “the other part of the laboratory” a group of five individuals was facing the same two technologies with the same optimal color. Importantly, social learners also knew that subjects in the other group knew how many points they were making after each choice. In essence social learners knew that the players in the other group were receiving individual feedback.

Within each period, social learners were first given the choice to learn which color had just produced the lowest payoff in the other group or which color had produced the highest payoff. After making their choices, the relevant color was communicated to each social learner privately. Social learners then made their own choices. Social learners knew they would not receive information about their own payoffs until the entire session was over for the day.

The experiment was repeated three times for two worlds, five times for one world, and six times for three worlds, with the differences depending on the speed of the players and the time available. For those players repeating the experiment fewer times, show-up fees were adjusted spontaneously to equilibrate total payoffs across worlds as much as possible. The possibility of repeating the experiment less than six times and adjusting the show-up fee was not discussed before the experimental sessions.

After all repetitions, subjects answered a questionnaire that recorded basic socio-demographic characteristics like gender, age, and academic major. The questionnaire also asked about learning strategies. Individual learners were questioned about the events that led them to revise their choices. Social learners were asked if they thought the color producing the best payoff or the color producing the worst payoff was better social information. They were also asked to provide more detailed information about their assessment of the social information.

After the questionnaire, subjects received their total payoffs based on points summed over all six repetitions. The average payoff from actual play (i.e. not including show-up fees) was 31.66 CHF (24.69 USD; 20.26 EUR). Sessions lasted about 2 hours.

B Payoff-dependence, models and anti-models

Posit a reference technology, which can be thought of as an incumbent technology, called technology 0. The word “technology” is generic and includes anything, including behavior, that controls the rate at which inputs (e.g. time) yield outputs (e.g. food). Payoffs for technology 0 are a random variable, X , with realizations, x , distributed according to the p.d.f. $f(x)$ and the c.d.f. $F(x)$. The mean payoff is μ_0 . An additional technology, which can be thought of as a technological innovation, is also available in the population. Call it technology 1. It brings random payoffs, Y , with realizations y . With probability e technology 0 is optimal in the sense that the expected payoff from technology 1 when technology 0 is optimal is less than the expected payoff from technology 0: $E[Y] = \mu_{10} < E[X] = \mu_0$. With probability $1 - e$ technology 1 is optimal in that its expectation in this case, μ_{11} is the greater of the two: $E[Y] = \mu_{11} > E[X] = \mu_0$. If technology 1 is optimal, payoffs are distributed according to $g_{opt}(y)$ and $G_{opt}(y)$, which are the p.d.f. and c.d.f. respectively. If technology 0 is optimal, payoffs for technology 1 are distributed according to $g_{sub}(y)$ and $G_{sub}(y)$.

Assume a person samples N individuals randomly from the previous period and

ranks them according to their payoffs. Someone who imitates the best then chooses the same technology as the person in the sample who received the highest payoff in the previous period. Someone who does the opposite of the worst chooses the technology not chosen by the lowest paid person in the sample. Let $I_t \in \{0, 1, \dots, N\}$ be a random variable specifying the number of people in the sample, taken at period $t+1$, who choose technology 1 during period t . I_t is thus binomial,

$$P(I_t = i_t) = \binom{N}{i_t} (q_t)^{i_t} (1 - q_t)^{N - i_t}, \quad (1)$$

where q_t is the proportion of the population choosing technology 1 in period t . If $i_t = N$, which happens with probability $(q_t)^N$, someone who imitates the best chooses technology 1 with certainty and never chooses technology 0. In contrast, someone who does the opposite of the worst never chooses technology 1 but always chooses technology 0. If $i_t = 0$, which happens with probability $(1 - q_t)^N$, someone who imitates the best always chooses technology 0 and never technology 1, while someone who does the opposite of the worst always chooses technology 1 and never technology 0.

For the moment, assume $N > i_t > 0$, and thus a mix of choices exists in the sample. To simplify the notation below, temporarily drop the “ t ” subscript from i_t , but note that in what follows $i = i_t$. Designate the $N - i$ draws from the technology 0 distribution as X_1, \dots, X_{N-i} and their corresponding order statistics as $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(N-i)}$. Similarly the i draws from the technology 1 distribution are Y_1, \dots, Y_i , and the resulting order statistics are $Y_{(1)} \leq Y_{(2)} \leq \dots \leq Y_{(i)}$. Under a mix of choices the density over the lowest payoff from the technology 0 distribution is

$$f_{X_{(1)}}(x) = (N - i) \{1 - F(x)\}^{N-i-1} f(x),$$

while the density over the highest payoff from the same distribution is

$$f_{X_{(N-i)}}(x) = (N - i)\{F(x)\}^{N-i-1}f(x).$$

If technology 0 is optimal, the density over the lowest payoff from the sub-optimal technology 1 distribution is

$$g_{Y_{(1)}}^{sub}(y) = i\{1 - G_{sub}(y)\}^{i-1}g_{sub}(y),$$

while the density over the highest payoff in this case is

$$g_{Y_{(i)}}^{sub}(y) = i\{G_{sub}(y)\}^{i-1}g_{sub}(y).$$

If technology 1 is optimal, the density over the lowest payoff from the technology 1 distribution is

$$g_{Y_{(1)}}^{opt}(y) = i\{1 - G_{opt}(y)\}^{i-1}g_{opt}(y),$$

and the density over the highest payoff is

$$g_{Y_{(i)}}^{opt}(y) = i\{G_{opt}(y)\}^{i-1}g_{opt}(y).$$

Consequently, if technology 0 is optimal, the *ex ante* conditional probability that the minimum payoff among the N individuals sampled comes from the technology 0 distribution follows. To simplify notation below, denote this probability $h_{00}(i)$, which is only relevant when $\mu_0 > E[Y] = \mu_{10}$:

$$h_{00}(i) = P(Y_{(1)} > X_{(1)} \mid N > i > 0) = 1 - \int_{-\infty}^{\infty} \int_{-\infty}^x f_{X_{(1)}}(x)g_{Y_{(1)}}^{sub}(y) dydx.$$

The *ex ante* conditional probability that the minimum payoff comes from the technology

0 distribution when technology 1 is optimal (i.e. $E[Y] = \mu_{11} > \mu_0$) is

$$h_{01}(i) = P(Y_{(1)} > X_{(1)} \mid N > i > 0) = 1 - \int_{-\infty}^{\infty} \int_{-\infty}^x f_{X_{(1)}}(x) g_{Y_{(1)}}^{opt}(y) dy dx.$$

In a similar fashion, the *ex ante* conditional probability that the highest payoff in the sample comes from the technology 1 distribution when technology 0 is optimal (i.e. $\mu_0 > E[Y] = \mu_{10}$), denoted $k_{10}(i)$, is

$$k_{10}(i) = P(Y_{(i)} > X_{(N-i)} \mid N > i > 0) = 1 - \int_{-\infty}^{\infty} \int_{-\infty}^x f_{X_{(N-i)}}(x) g_{Y_{(i)}}^{sub}(y) dy dx.$$

The analogous *ex ante* conditional probability that the highest payoff comes from the optimal technology 1 (i.e. $E[Y] = \mu_{11} > \mu_0$) is

$$k_{11}(i) = P(Y_{(i)} > X_{(N-i)} \mid N > i > 0) = 1 - \int_{-\infty}^{\infty} \int_{-\infty}^x f_{X_{(N-i)}}(x) g_{Y_{(i)}}^{opt}(y) dy dx.$$

Taking account of (1) and the fact that technology 1 may or may not be optimal, the probability that imitating the best (IB) identifies the optimum (opt) is

$$P(\text{opt} \mid \text{IB}) = \sum_{i_t=1}^{N-1} \binom{N}{i_t} (q_t)^{i_t} (1 - q_t)^{N-i_t} \{e(1 - k_{10}(i_t)) + (1 - e)k_{11}(i_t)\} + e(1 - q_t)^N + (1 - e)(q_t)^N. \quad (2)$$

Figure S1 shows this probability as a function of q_t for various values of e when the technology 0 distribution is $N(0, 1)$, the sub-optimal technology 1 distribution is $N(-1, 1)$, and the optimal technology 1 distribution is $N(1, 1)$.

The analogous probability for doing the opposite of the worst (DOW) is

$$P(\text{opt} \mid \text{DOW}) = \sum_{i_t=1}^{N-1} \binom{N}{i_t} (q_t)^{i_t} (1 - q_t)^{N-i_t} \{e(1 - h_{00}(i_t)) + (1 - e)h_{01}(i_t)\} + e(q_t)^N + (1 - e)(1 - q_t)^N. \quad (3)$$

Figure S2 shows this probability as a function of q_t for various values of e and the same distributions used for imitating the best immediately above.

With respect to the present experiment, $e = 0$, and so we define technology 1 as always optimal. For this specific case, the probabilities specified by (2) and (3) reduce to

$$P(\text{opt} \mid \text{IB}) = \sum_{i_t=1}^{N-1} \binom{N}{i_t} (q_t)^{i_t} (1 - q_t)^{N-i_t} k_{11}(i_t) + (q_t)^N \quad (4)$$

and

$$P(\text{opt} \mid \text{DOW}) = \sum_{i_t=1}^{N-1} \binom{N}{i_t} (q_t)^{i_t} (1 - q_t)^{N-i_t} h_{01}(i_t) + (1 - q_t)^N. \quad (5)$$

In the case of the experiment, the payoff distribution for the sub-optimal technology 0 was $N(30, 12)$, and the distribution for the optimal technology 1 was $N(38, 12)$. Both distributions were truncated at 0 and 68, and thus the standard deviations were slightly less than 12. Apart from some minor disturbances associated with truncation, both distributions were roughly symmetric. In addition, social learners did not *sample* from a larger population. Instead perfectly accurate social information was available in every period. As a consequence, only the conditional probabilities $h_{01}(i)$ and $k_{11}(i)$ are needed to derive a theoretical prediction. For the specific truncated payoff distributions used in the experiment and groups of 5 individual learners, Figure S3, which is the same as Figure 1 in the main text, shows these probability functions. This graph is a theoretical prediction of what the aggregate data should look like if 1) social learners always imitated the best when they accessed the best color and if 2) social learners who accessed the worst color in a given period always did the opposite of the worst. The predictions are conditional in two ways. First, the prediction is conditioned on whether a social learner chose to access the worst or best color in a period. Second, the prediction is conditioned on the number of individual learners choosing the optimal technology, which was information unavailable to social learners.

As Figure S3 shows, the two heuristics have symmetric value when the two distribu-

tions are both normal with equal variances. If one of the distributions is notably skewed or has a larger variance, this changes matters considerably, but here we focus on the case with normal distributions and equal variances. The values of the two heuristics are symmetric in that doing the opposite of the worst has value when optimal choices are at low frequencies in exactly the same way that imitating the best has value when optimal choices are at high frequencies. A central question then centers around the forces controlling the frequency of optimality among the social group (i.e. q_t).

In the above model, we adopted a sampling interpretation. Each individual samples N others from the population. The important question then concerns what controls the distribution of technologies in the population for this distribution will affect the statistical properties of multiple samples. For example, imagine a situation in which technology 0 is an incumbent technology that has been available for many years, and technology 1 is a recent innovation that produces higher payoffs on average, a fact not known with certainty to the individuals in the population. If many individuals are conservative with respect to adopting new technologies, and perhaps they also exhibit a tendency to use the most common technology, then such forces will initially keep the optimal technology at low frequency in the population. In this case samples will be biased toward the sub-optimal technology and imitating the best will tend to perpetuate this phenomenon. Doing the opposite of the worst, however, will tend to identify the optimal innovation and initially have the opposite effect. If the new optimal technology also brings yields with a lower variance, this fact will further reduce the value of imitating successful behaviors relative to avoiding unsuccessful behaviors. Importantly, however, in a dynamical setting with feedbacks, when the same individuals are learning both individually and socially, maintaining the advantage of doing the opposite of the worst will be very difficult. As it increases optimality, it simultaneously reduces its own value. Imitating success does not have this problem.

As a different approach, the sampling interpretation is not essential. If the social

group is fixed and individual learners and social learners are separate, as we exogenously imposed in the experiment, the resulting model is the same as in (4) and (5). The interpretation is different, however, in that q_t is the probability each individual learner chooses the optimal technology in the fixed social group. What does this mean for the dynamics of social learning in the experiment? If social learners believe that individual learners are initially impartial about which color is best ($q_t \approx 0.5$), imitating the best and doing the opposite of the worst are roughly equivalent in value. If social learners assume individual learners learn through time ($\forall t, q_{t+1} \geq q_t$ with a strict inequality for some t), imitating the best becomes better than doing the opposite of the worst. The speed of this process depends on the effectiveness, real or assumed, of individual learning.

Lastly, imitating the best and doing the opposite of the worst are not the only possible relations between social information and the subsequent choices of social learners. The social learner could also do the opposite of the best (DOB) and imitate the worst (IW). The probabilities of identifying the optimal technology under these two social learning heuristics are also shown in Figure S3. Specifically, because only two colors are available, if imitating the best identifies the sub-optimal technology, doing the opposite of the best must identify the optimal technology. Thus the following must be true.

$$P(\text{opt} \mid \text{DOB}) = P(\text{sub-opt} \mid \text{IB}) = 1 - P(\text{opt} \mid \text{IB}) \quad (6)$$

Similarly, it must also be true that

$$P(\text{opt} \mid \text{IW}) = P(\text{sub-opt} \mid \text{DOW}) = 1 - P(\text{opt} \mid \text{DOW}). \quad (7)$$

The equations mean that one can also compare the value of, say, imitating the worst versus doing the opposite of the worst by inspecting the space below and above the relevant curve in Figure S3.

References

Fischbacher, U. (1999). *z-Tree: toolbox for readymade economic experiments*. IEW Working paper 21, University of Zuerich.

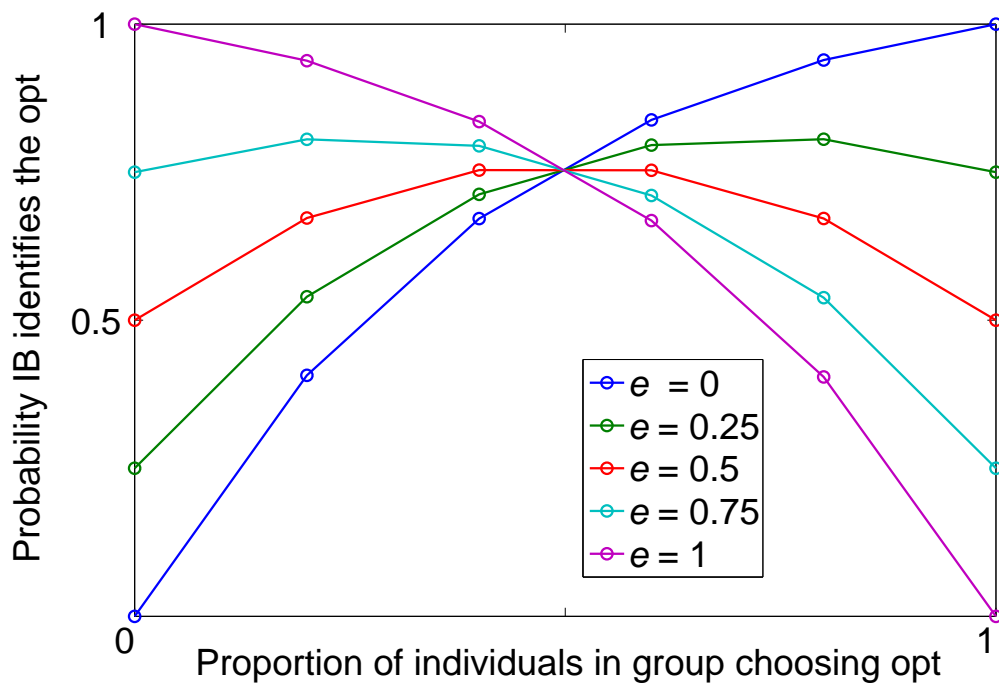


Figure S1. The probability that imitating the best identifies the optimal technology, $P(\text{opt} \mid \text{IB})$, as a function of the proportion of individuals in the group choosing the optimal technology, q_t , for 5 different values of e . In this case, technology 0 brings payoffs distributed according to $N(0, 1)$, while the sub-optimal technology 1 distribution is $N(-1, 1)$, and the optimal technology 1 distribution is $N(1, 1)$.

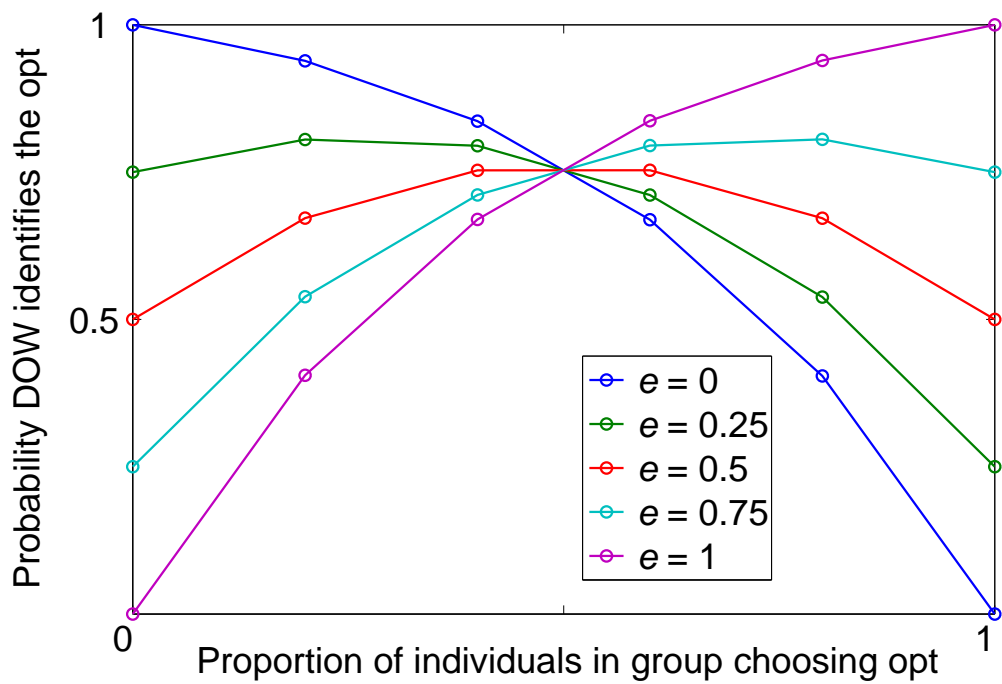


Figure S2. The probability that doing the opposite of the worst identifies the optimal technology, $P(\text{opt} \mid \text{DOW})$, as a function of the proportion of individuals in the group choosing the optimal technology, q_t , for 5 different values of e . The payoff distributions are the same as in the Figure above.

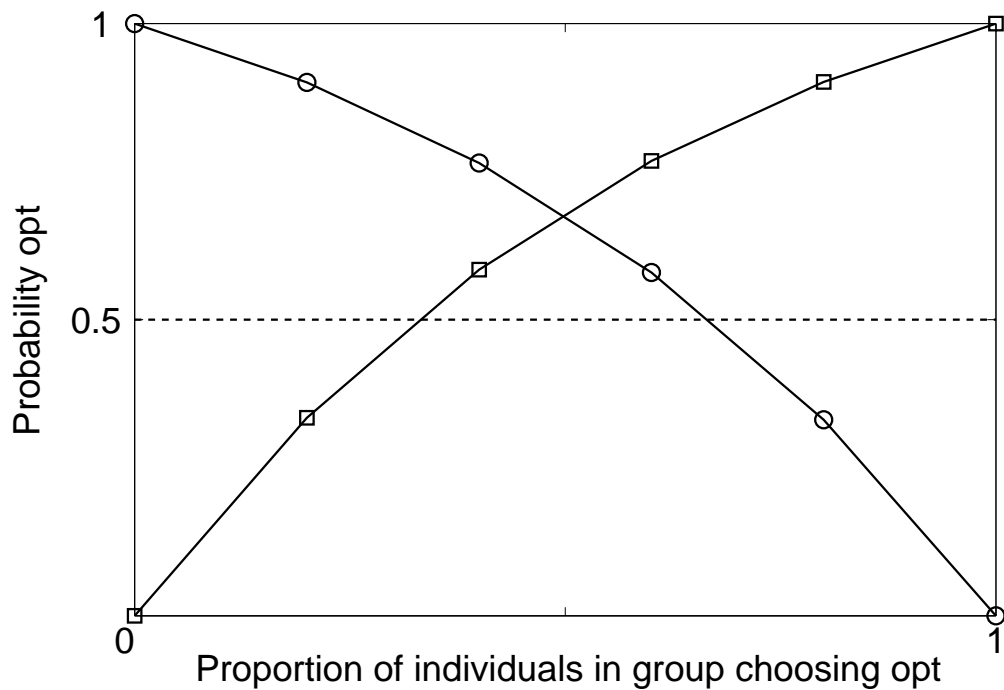


Figure S3. The line with squares shows the probability that imitating the best identifies the optimal color under the specific payoff distributions used in the experiments. The curve is a function of the proportion of individual learners choosing the optimal color. Note that the residual probability (i.e. from the curve upward) is the probability that imitating the best identifies the sub-optimal color and thus the probability that doing the opposite of the best identifies the optimal color. The line with circles shows the analogous probability that doing the opposite of the worst identifies the optimal color. The residual probability in this case is the probability that doing the opposite of the worst identifies the sub-optimal color and thus the probability that imitating the worst identifies the optimal color. As a reference against which to judge the statistical biases the different types of social learning create, a constant probability of 0.5 is also shown.