# THE EVOLUTIONARY RATIONALITY OF SOCIAL LEARNING

RICHARD MCELREATH, BARBARA FASOLO, AND ANNIKA WALLIN

## Introduction

A common premise surrounding magic is that words themselves have power. Speaking the right words in the right context is believed to create fantastic effects. Everything from Old Norse runes to magic in the Harry Potter books requires activation with words. This kind of belief is a feature of not only western myth and magic, but also of African (famously of Azande oracles: Evans-Pritchard 1937) and Asian (Daoist) traditions. Some healers in the Muslim world write in ink verses from the Koran, and then they wash the ink into a potion to be consumed. In Swahili, one can use the same word, *dawa*, to refer to both magical spells and the influence a charismatic speaker has over a crowd.

Why do so many peoples believe that words themselves are magical? These beliefs are not necessarily irrational. Every one of us, by speaking, can alter the minds of those within earshot. With the word *snake*, one can conjure a potentially terrifying image in the minds of others. Statements like these reveal how hard it is to really control our thoughts as well as the power mere utterances have over us. Of course people are savvy and do not robotically obey all suggestion or command. However, spoken opinion and advice is highly valued almost everywhere. The words of others, carrying information, can have powerful effects on our own behavior. The mere suggestion that something—like a measles vaccine—is dangerous can have huge effects on behavior. People and governments intuit this power and as a result attempt to control the words they themselves and others are exposed to. Words really are a kind of mind control, or at least mind influence. Their power can travel through the empty air and affect the behavior of masses of other people in powerful ways. They are like magic.

---

*Date*: July 20, 2010. Cite as: McElreath, R., Fasolo, B., & Wallin, A. (in press). The Evolutionary Rationality of Social Learning. In R. Hertwig, U. Hoffrage & the ABC Research Group (Eds.), *Simple Heuristics in a Social World.* Oxford University Press.

The psychology of humans is uniquely elaborated for this kind of "magical" influence. The capacity for language is only one way that social influence on behavior is truly baked into our nature. Observational learning of various kinds is equally powerful, as people combine theory and social information to arrive at inferences about the reasons for and consequences of behavior. But animals other than humans also use social information (e.g. Bonner 1980, Galef 1992, 1996, Price et al. 2009, Giraldeau & Caraco 2000, Fragaszy & Perry 2003, Laland & Galef 2009). While the psychological mechanisms and diversity of social learning heuristics among, for example, baboons is not the same as that among humans, the savvy monkey too uses information from its social environment. As a result, the field of evolutionary ecology has long been interested in the design and diversity of social learning heuristics, simple strategies that animals use to extract useful information from their social environments.

In this chapter, we review a slice of this literature, as well as explicitly analyze the evolution of social learning heuristics. We outline a family of social learning heuristics and analyze their evolutionary performance—ability to persist and replace other heuristics—under two broadly different kinds of environmental variation. As each social learning heuristic also shapes a social environment as individuals use it, we consider the population feedbacks of each heuristic, as well. Feedbacks occur when the behavior generated by a heuristic in turn change the success rate of a heuristic, a phenomenon sometimes called *frequency-dependence*. The analyses in this chapter are ecological—the performance of each heuristic is always in the context of a specific set of assumptions about the population structure and environment. They are also game theoretic—social learning heuristics use but also modify the social environment, inducing strong frequency-dependence. Our analyses are also explicitly evolutionary—heuristics succeed or fail depending upon long-term survival and reproduction in a population, not atomistic one-shot payoffs. As a result, some of our conclusions reflect an evolutionary rationality that is sometimes counter-intuitive. For example, heuristics that randomize their behavior can succeed where those that are consistent fail. Overall, however, the approach we review here supports the general conclusion that social learning heuristics are likely to be multiple and subtly adapted to different physical, statistical, and social environments.

## Social learning heuristics

In parallel to the literature in bounded rationality, evolutionary ecologists and anthropologists studying social learning have proposed that there exists a toolbox of contextually deployed heuristics that are suited to different ecological and social environments (reviews in Henrich & McElreath 2003, Richerson & Boyd 2005). The basic premise is that information about the world is costly to acquire and process (Boyd & Richerson 1985). So as a method of reducing information requirements and processing costs, natural selection favors strategies that leverage the specific correlations of specific environments in order to make locally-adaptive choices. Each heuristic in the toolbox is best deployed in a different circumstance, and some heuristics are more domain-general than others. Thus the expectation is an ecology of inferential strategies that individuals can use to choose behavior. While some of these strategies are more cognitively demanding and information hungry than others, all are quite "bounded," compared to Bayesian justifications for social learning (Boyd & Richerson 2005, 2001, Bikhchandani et al. 1992). Like other hypothesized heuristics, these social learning heuristics can be compared to laboratory behavior. In recent years, there has been a small industry of testing these models against dynamic learning data (Efferson et al. 2008, McElreath et al. 2008, 2005, Mesoudi & O'Brien 2008, Mesoudi & Whiten 2008, Mesoudi 2008).

In this section, we outline and begin to analyze a toolbox of social learning heuristics that evolutionary ecologists and evolutionary anthropologists have studied. The collection of heuristics we review is not complete. Many other heuristics could be nominated, and each heuristic we do nominate is in reality a family of heuristics. However, by constraining our discussion to the most commonly-discussed strategies, we have space to derive each from first (or at least basic) principles and, later, analyze the performance of several in different ecological circumstances.

Theory leads us to expect that people (and perhaps other animals) possess a toolbox of social learning heuristics. Our goal is to study the conditions, both in terms of physical and social environment, that favor different heuristics. In Table 1, we list several social learning heuristics from the literature, also listing aliases and a sample of relevant citations to previous work. In the remainder of this chapter we demonstrate the analysis of a few of these. We also present a new analysis of the evolution of social heuristics in time varying environments. The

**Box 1 Psychologist's guide to theoretical evolutionary ecology**

We provide here short definitions of some of the key evolutionary ecology concepts in the chapter. A complete introduction can be found in McElreath & Boyd (2007).

**Population:** All organisms of the same species who are linked by gene-flow, the possible exchange of genes across generations. Populations can be subdivided into smaller groups, in which case not all individuals will be able to interbreed in a given generation. Nevertheless, as long as sub-populations are linked by migration across generations, all individuals in the total population can in principle be linked by gene-flow. The population is the natural unit of evolution, as the frequencies of genes and behavior change over time among the individuals within it.

**Life cycle:** The sequence of events that happen between birth and death. These events, aggregated over many individuals in a population, induce selection on specific genetic or cultural variants.

**Strategies:** Heritable aspects of contingent behavior. Heuristics are strategies. Behavior is distinct from strategy, as the same strategy can produce different behavior in different contexts. In evolutionary models, strategies are what evolve, and the frequencies of different strategies, or the alleles that code for them, describe the state of the population.

**Fitness:** Typically, the expected number of copies of a specific strategy per individual in the next generation. Fitness depends upon survival and reproduction. Fitness concepts do however vary among models of evolutionary processes, because the goal is to define a quantity that will allow us to predict the population dynamics. Evolutionary ecologists attempt to understand what will evolve, and fitness is a tool in such an analyses.

**Dynamics:** Time evolution in a physical system. In evolutionary models, the dynamics are the time trends of the frequencies of different heritable strategies and behaviors. The frequencies at any time in the future depend upon the frequencies in the past. Evolutionary analysis is a branch of dynamical systems theory.

**Equilibrium:** A combination of strategies at the population level at which the dynamics of the population result in no change. Equilibria can be stable or unstable. The dynamics of a population return the population to a stable equilibrium, when the frequencies are changed slightly. In contrast, the dynamics lead the population away from an unstable equilibrium, when the frequencies are changed slightly. Stable equilibria are candidate end states of the evolutionary process.

**Invasion:** When a rare strategy can increase in numbers in a population, it can invade that population. A strategy that once common can repel rare invaders of all other strategies is an evolutionarily stable strategy.

**Geometric mean fitness** If we define fitness as the product of survival probability and mean reproduction rate, then geometric mean fitness is the geometric mean of different probable fitness values. Natural selection in many models maximizes geometric mean fitness, rather than average fitness, because natural selection is affected by both the mean and variance in fitness across generations.

Table 1. Major social learning heuristics from the literature, with other names for the same strategies and citations to a sample of previous evolutionary analysis.

| Heuristic | Other names | Citations |
|---|---|---|
| Unbiased social learning | Linear social learning, random copying, imitation | Boyd & Richerson (1995, 1985), Cavalli-Sforza & Feldman (1981), Rogers (1988), Mesoudi & Lycett (2009), Aoki et al. (2005), Wakano et al. (2004) |
| Consensus learning | Conformity, conformist transmission, positive frequency dependent imitation, majority rule imitation | Boyd & Richerson (1985), Mesoudi & Lycett (2009), Lehmann & Feldman (2008), Henrich & Boyd (2001, 1998), Wakano & Aoki (2007) |
| Payoff bias | Success bias, indirect bias | Boyd & Richerson (1985), Henrich (2001), Schlag (1998, 1999) |
| Prestige bias | Indirect bias | Boyd & Richerson (1985), Henrich & Gil-White (2001) |
| Kin bias | Vertical transmission, parent-child transmission | McElreath & Strimling (2008) |

dynamical systems approach common in evolutionary analysis may be unfamiliar to many readers, so we provide a quick guide in Box 1.

**The environmental challenge.** In order to make progress defining and analyzing the performance of different learning heuristics, we have to define the challenge the organism faces. Here we use an evolutionary framing of the common multi-armed bandit problem.

Assume that a each individual at some point in its life has to choose between a very large number of distinct behavioral options. These options could be timing of reproduction, patterns of paternal care, or any other set of mutually exclusive options. Only one of these options is optimal, producing a higher fitness benefit than all the others. We will assume that a single optimal behavior increases an individual's fitness by a factor $1 + b > 1$. All other behavior leaves fitness unchanged. In particular, let $w_0$ be an individual's fitness before behaving. Since there are a very large number of alternative choices, randomly guessing will not yield a fitness payoff much greater than $w_0$. Then those who choose optimally have fitness $w_0(1 + b)$, while those who do not have fitness $w_0$. Because fitness does not depend upon how many other individuals also choose the same option, these payoff are not directly frequency dependent.

**Individual updating.** The foil for all the social learning heuristics we consider here is a gloss *individual updating* heuristic. However the mechanism works in detail, we assume that individuals have the option of relying exclusively on their own experience, when deciding how to behave. We assume that individual updating requires sampling and processing effort, as well as potential trial-and-error. As a result, an organism that uses individual updating to learn optimal behavior pays a fitness cost by having their survival multiplied by $c \in [0, 1]$. This means the fitness of an individual updater is always $w_0(1 + b)c > w_0$. We assume that individual updating is always successful at identifying optimal behavior. We have analyzed the same set of heuristics, assuming that individual updating is successful only a fraction $s$ of the time. This assumption, while biologically satisfying, adds little in terms of understanding. It changes none of the qualitative results we will describe in later sections, while adding mathematical complexity.

**"Unbiased" social learning.** Probably the simplest kind of social learning is a strategy that randomly selects a single target to learn from. Much of the earliest evolutionary work on social learning has considered this strategy (Cavalli-Sforza & Feldman 1981, Boyd & Richerson 1985), and even more recent work continues to study its properties (Aoki et al. 2005, Wakano et al. 2004).

To formalize this heuristic, consider a strategy that, instead of trying to update individually, copies a random member of the previous generation. Such a strategy avoids the costs of learning, $c$. Social learning may entail costs, but they are assumed to be lower than those of individual updating. The unavoidable cost of social learning is that the payoff from such a heuristic depends upon the quality of available social information. We will refer to such a strategy as "unbiased" social learning. We use the word "unbiased" to describe this kind of social learning, although the word "bias" is problematic. We use the term to refer only to deviations from random, not from normative standards. The word "bias" has been used in this way for some time in the evolutionary study of social learning (Boyd and Richerson 1985, for example).

Let $q$ ("quality" of social information) be the proportion of optimal behavior among possible targets of social learning. Then the expected fitness of an unbiased social learner is $w_0(1+qb)$, wherein $b$ is discounted by the probability that the unbiased social learner acquires optimal behavior. Like other social learning heuristics, unbiased social learning actively shapes social environments itself—a heuristic that uses behavior is a cue and produces behavior will necessarily create feedbacks in the population of learners. As a result, a satisfactory analysis must be dynamic. We consider such an analysis in a later section.

Note that we assume no explicit sampling cost of social learning. Indeed, several of the social learning heuristics we consider in this chapter use the behavior of more than one target, and we have not considered explicit costs of sampling these targets either. Consensus social learning (below), in our simple model of it, uses three targets, and payoff biased social learning (later in this section) uses two targets. Does this mean that consensus is worse than payoff bias, when both are equal in all other regards? We think the answer to this question will depend upon details we have not modeled. Do other activities provide ample opportunity to sample targets for social learning or must individuals instead search them out and spend time observing their behavior? If the behavior in question is highly complex and requires time and practice to successfully transmit, like how to make an arrow, then consensus learning may entail higher behavioral costs than, say, payoff bias. This is because a consensus learner needs to observe the detailed technique of three (or more) individuals, while the payoff biased learner must only observe payoffs and then invest time observing one target. We could invent stories that favor consensus, as well. And while constructing and formalizing such stories is likely productive, it is a sufficiently detailed project that we have not undertaken it in this chapter. But we do not wish to send the message that sampling costs and sampling strategy—how many to sample and when to stop, for example—are uninteresting or unimportant questions. They are simply beyond the scope of this introduction.

**Consensus social learning.** An often discussed category of social learning heuristics are those that use the commonality of a behavior as a cue (Boyd & Richerson 1985, Henrich & Boyd 1998, Mesoudi & Lycett 2009, Wakano & Aoki 2007). When an individual can sample more than two targets, it is possible to use the frequency of observed behavior among the targets as a cue to guide choice. This kind of strategy has been called positive frequency dependence and conformist transmission. We adopt the label "consensus" social learning here, because "conformity" is a vague term that many people associate with social learning of any kind (as it is often used in psychology). The alternative "positive frequency dependence" is an unwieldy term.

Consensus social learning can be most easily modeled by assuming that an individual samples three targets at random and preferentially

**Box 2 Deriving consensus social learning**

We use a simple table to derive probabilities of acquiring optimal (1) and non-optimal (0) behavior, using a consensus social learning heuristic.

| Observed behavior | Pr(Obs) | Pr(1) | Pr(0) |
|---|---|---|---|
| 1 1 1 | $q^3$ | 1 | 0 |
| 1 1 0 | $3q^2(1-q)$ | $2/3+D$ | $1/3-D$ |
| 1 0 0 | $3q(1-q)^2$ | $1/3-D$ | $2/3+D$ |
| 0 0 0 | $(1-q)^3$ | 0 | 1 |

$D < 1/3$ is the strength of the preference for consensus. The columns are, in order from left to right: the vector of observed behavior from a sample of three targets, where 1 indicates optimal behavior and 0 any non-optimal behavior; the probability of sampling that vector; the probability of acquiring optimal behavior under the heuristic, given that sample; and the probability of acquiring non-optimal behavior. First, multiply each probability of the observed vector of behavior by the probability of acquiring optimal behavior, Pr(1). Then, add together all the products from each row. In this case, $q^3 \times 1 + 3q^2(1-q) \times (2/3+D) + 3q(1-q)^2 \times (1/3-D) + (1-q)^3 \times 0$ simplifies to Expression 1 in the main text, assuming for simplicity, that $D = 1/3$. We are making the simplifying assumption in this derivation that all non-optimal behavior is categorized together. As long as most immigrants come from a single or few neighboring patches, this will not be a bad approximation. Thus we consider these results to hold for structured populations with nearest-neighbor migration. When it is a bad approximation, however, it is a conservative estimate that biases our analysis against consensus learning, not in favor of it.

adopts the most common behavior among them. In Box 2, we show how to use this definition to derive the expected probability that an individual using consensus social learning will acquire optimal behavior:

$$\text{(1)} \qquad \Pr(1) = q + q(1-q)(2q-1).$$

Boyd and Richerson (1985) have considered a number of generalizations of this heuristic, including different weights given to each target, as well as correlations among the behavior of the targets. We will ignore these complications in this chapter, as our goal is to motivate a mode of analysis and to emphasize the differences among quite different social learning heuristics, rather than among variants of the same heuristics.

**Payoff biased social learning.** Another often analyzed category of social learning heuristic is payoff guided learning (Boyd & Richerson 1985, Schlag 1998, 1999, Stahl 2000). By comparing observable payoffs—health, surviving offspring, or even more domain-specific measures of success—among targets, an individual can use differences in

**Box 3 Deriving payoff biased social learning**

Again we use a table to derive probabilities of acquiring optimal (1) and non-optimal (0) behavior, this time using payoff biased social learning.

| Actual behavior | Pr(actual) | Observed payoffs | Pr(obs) | Pr(1) | Pr(0) |
|---|---|---|---|---|---|
| 1 1 | $q^2$ | | | 1 | 0 |
| 1 0 | $2q(1-q)$ | 1 0 | $(1-x)^2$ | 1 | 0 |
| | | 0 0 | $x(1-x)$ | 1/2 | 1/2 |
| | | 1 1 | $x(1-x)$ | 1/2 | 1/2 |
| | | 0 1 | $x^2$ | 0 | 1 |
| 0 0 | $(1-q)^2$ | | | 0 | 1 |

In this table, $q$ is the frequency of optimal behavior among targets and $x$ is the chance of incorrectly judging the payoff of a target (or similarly, $1-x$ is the correlation between behavior and observed payoffs). The individual using payoff biased social learning samples two targets at random and assesses their payoffs. The individual copies the behavior of the target with the higher observed payoff, unless both observed payoffs are the same, in which case one target is copied at random.

payoff as a guide for learning. This kind of heuristic generates a dynamic often called "the replicator dynamic" in evolutionary game theory (Gintis 2000). This dynamic is very similar to that of natural selection, and is often used as a boundedly-rational assumption in social evolutionary models (McElreath et al. 2003, e.g.) and even epidemiology (Bauch 2005).

A simple model of payoff biased social learning assumes that individuals sample two targets and preferentially adopt the behavior of the target with the higher observed payoff. We assume that there is a chance $x$ that the individual can correctly judge the payoff of a target to be high or low. Another interpretation is that $x$ is the chance that a target's observable payoff is uncorrelated with behavior. Using these assumptions, we show in Box 3 that this heuristic leads to a chance:

$$\Pr(1) = q + q(1-q)(1-x)$$

of acquiring optimal behavior.

A more general model of payoff-bias allows for the aspect of the target to be judged as "success" to itself be socially transmitted. When this is the case, unaniticipated social processes become possible, such as the run-away exaggeration of preferences for traits that are judged as successful (Boyd and Richerson 1985).

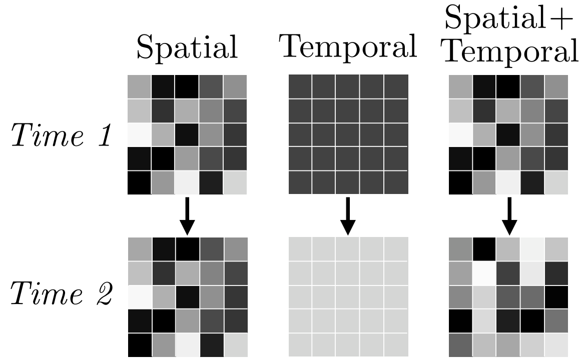


Figure 1. Abstract forms of environmental variation. Each square represents an overhead view of environmental variation. With purely spatial variation, left column, different locales favor different optimal behavior, represented by the shading levels. But these differences remain static through time, moving from top to bottom. With purely temporal variation, middle column, all locales favor the same behavior, but the optimal behavior varies through time. With both spatial and temporal variation, on the right, locales may be different both from other locales and from themselves, through time.

## Ecological variation and social learning

Given the definitions of heuristics in the previous section, we now turn to analyzing the evolutionary dynamics of these four strategies—individual updating, unbiased social learning, consensus social learning, and payoff biased social learning—both alone and in competition. We will assume that each heuristic is a heritable strategy and study their population dynamics. We consider how these heuristics perform in two statistical environments: (1) a spatially variable environment, in which different behavior is optimal in different places, and (2) a temporally variable environment, in which different behavior is optimal at different times. We also consider the interaction of these two kinds of variation (Figure 1).

The reason for focusing on environmental *variation*, the rates at which the environment changes spatially and temporally, is that "learning" as it has long been studied in evolutionary ecology has identified ecological variation as a prime selection pressure favoring both individual learning (phenotypic plasticity, as it is often called) and social learning (Levins 1968, Boyd & Richerson 1988). In a perfectly stationary environment, genetic adaptation (canalization) does a fine job of adapting the organism, without any of the cognitive overhead and potential for error that arises from using information during development to alter behavior. Thus evolutionary ecologists still consider the nature of environmental variation to be a key factor in the evolution, maintenance, and design of learning (see e.g. Dunlap & Stephens 2009, Lande 2009).

Our goal in this section is to describe some conditions under which each social learning heuristic is well-adapted. No single heuristic can succeed in all circumstances. To some extent, all social learning heuristics depend upon some kind of individual updating, for example. Additionally, the differences among social learning strategies generate different dynamics for the quality of the social environment. Because our analysis is explicitly evolutionary, it will turn out that good heuristics are those which can live well with themselves. Social heuristics tend to shape the social environments that they rely upon for information. Thus the precise ways in which the physical and social environments interact play a large role in determining the long-term evolutionary success of a heuristic.

**Spatial variation in the environment.** In this section we consider the first of two stereotyped forms of environmental variation, spatial variation in optimal behavior. We assume that the environment is sub-divided into a large number of distinct patches, each with a unique optimal behavior. Optimal behavior within each patch is forever the same. However different patches never have the same optimal behavior. Within each patch, a large sub-population of organisms follows the life-cycle: (1) birth, (2) learning, (3) behavior, (4) migration, (5) reproduction, (6) death. Individuals are born naive and must use some strategy to acquire behavior. If behavior is optimal for the local patch, then the individual's fitness is multiplied by the factor $1 + b > 1$. Otherwise, fitness is unchanged. A proportion $m$ of the local population emigrates to other patches and an equally sized group immigrates from other patches. Generations overlap only long enough for newly born naive individuals to possibly learn from the previous generation of adults. Because of migration, some of the adults available to learn from are immigrants, all of whom possess non-optimal behavior for their new patch. While fitness is assigned in natal patches, we assume that adults continue to display their behavior after migration, and so naive individuals run the risk of learning from immigrants. Additionally, we assume that naive individuals cannot tell who is and is not a native of their local patch. While such cues might be available in many circumstances, they are certainly not always available.

*Individual updating.* The expected fitness of an individual updater is:

$$w(\mathrm{I}) = w_0(1+b)c,$$

where $0 < c < 1$ is a multiplicative cost to survival or reproduction. Provided that $(1+b)c > 1$, individual updating will be the best-adapted heuristic, whenever the quality of social information in the local patch, $q$, is equal to zero. However, since individual updating quickly increases the frequency of optimal behavior in the local patch, this heuristic quickly generates a social environment favorable to one social learning heuristic or another.

*Unbiased social learning.*

> Précis. *While individual updaters generate locally adaptive behavior that social learners can exploit, mixing among patches erodes this information. Therefore, unbiased social learning can invade a population using individual updating, provided mixing among patches is not too strong. Unbiased social learning can never completely replace individual updating, however. Thus when unbiased social learning can invade, there will exist a stable mix of individual updating and unbiased social learning in the population.*

We now consider when unbiased social learning (U) can out-perform individual updating. In generation $t$, the expected fitness of an individual using unbiased social learning is:

$$w(\mathrm{U})_t = w_0(1+q_t b),$$

where $q_t$ is the frequency of optimal behavior among targets in the current generation $t$. To compute the expected fitness across generations, we need an expression for the average amount of optimal behavior in the population. In Box 4, we show how to estimate this expression.

We use this expression to prove how selection increases and decreases frequencies of these two heuristics. As has been shown many times (see Rogers 1988 for a clear example), neither individual updating nor unbiased social learning can exclude one another under all circumstances, and so models of this kind predict that both will co-exist, in the absence of other heuristics. A stable proportion of individual updaters, $\hat{p}$, is found where:

$$w(\mathrm{I}) = w(\mathrm{U})|_{p=\hat{p}},$$

$$\hat{p} = \frac{m\big((1+b)c-1\big)}{(1-m)(1+b)(1-c)}. \tag{2}$$

**Box 4 The steady state amount of optimal behavior under a mix of individual updating and unbiased social learning**

To compute the expected fitness across generations, we need to study the dynamics of $q$. The frequency of optimal behavior among targets at time $t$ is defined by the recursion:

$$q_t = (1 - m)\big(p_{t-1} + (1 - p_{t-1})q_{t-1}\big) + m(0),$$

where $p_{t-1}$ is the proportion of the local population comprising individual updaters, in the previous generation. To understand this recursion, first consider that a proportion $p_{t-1}$ of targets are individual updaters. If a social learning targets one of these, then it is certain to acquire optimal behavior (before migration). If instead a social learner targets another social learner, which happens $1 - p_{t-1}$ of the time, there is a chance $q_{t-1}$ of acquiring optimal behavior, because that is the chance each social learner in the previous generation had of acquiring optimal behavior. Finally, only a proportion $1 - m$ of the local group remains to be potentially targets of learning. The proportion $m$ that immigrates possesses only non-optimal behavior, however it was learned. If we assume that natural selection of the frequencies of learning heuristics is slow relative to the dynamics of $q$, then we can treat $p_t$ as a constant $p$ in the expression above and set $q_t = q_{t-1} = \hat{q}$ and solve for the expected proportion of optimal behavior among targets:

$$\hat{q} = \frac{(1 - m)p}{1 - (1 - m)(1 - p)}. \tag{3}$$

Numerical works shows that this fast-slow dynamics approximation is very accurate, unless selection (proportional to $b$) is very strong.

Inspecting the partial derivatives of the right-hand side shows that increasing migration, increasing the value of optimal behavior, and decreasing the cost of individual updating all increase the equilibrium frequency of individual updating ($\partial\hat{p}/\partial m > 0$, $\partial\hat{p}/\partial b > 0$, and $\partial\hat{p}/\partial c > 0$, $\forall b > 0, c \in [0, 1], m \in [0, 1], (1 + b)c > 1$).

These results tell us that, if migration is too common, then unbiased social learning cannot invade a population of individual updaters, because too often the behavior available to copy is appropriate for a different patch. However the amount of migration unbiased social learning can tolerate depends upon the costs and benefits of learning. Increasing migration, increasing the value of optimal behavior, and decreasing the cost of individual updating all increase the equilibrium frequency of individual updating and decrease the frequency of social learning.

*Consensus social learning.*

> Précis. *Consensus social learning yields higher fitness and replaces unbiased social learning, provided mixing between patches is not so strong as to make the expected local proportion of optimal behavior below one-half. If*

**Box 5 Condition for consensus social learning to invade a mixed population of individual updating and unbiased social learning**

The expected fitness of a rare consensus learner (C) in generation $t$ is:

$$w(\mathrm{C})_t = w_0(1 + b(q_t + q_t(1 - q_t)(2q_t - 1)),$$

where the factor $q_t + q_t(1-q_t)(2q_t - 1)$ was derived in Box 2. The invader faces a value of $q_t = \hat{q}$, reached under the joint action of both individual updaters and unbiased social learners. But regardless of the value of $q$, for consensus learning to do better than either common heuristic, all that is required is that:

$$w(\mathrm{C})_t > w_0(1 + q_t b) \implies q_t > \frac{1}{2}.$$

Consensus social learning is favored in any generation in which the expected proportion of optimal behavior among targets is greater than one-half. Substituting in the expression for $\hat{q}$, this condition simplifies to $m < \frac{\hat{p}}{1+\hat{p}}$. So as long as migration isn't so strong as to flood local adaptive learning, which happens at a rate $\hat{p}$, consensus learning can invade a mix of individual updating and social learning. Since $\hat{p}$ is a function of $m, b, c$, we can substitute in the expression for $\hat{p}$ derived in the previous section. Doing so results in condition 4 in the main text. If consensus learning can invade, it will always exclude unbiased social learning. Sometimes consensus learning can also exclude individual updating. If consensus learning is common, then the expected proportion of locally optimal behavior is:

$$\hat{q} = (1 - m)\big(\hat{q} + \hat{q}(1 - \hat{q})(2\hat{q} - 1)\big) = \frac{3}{4} + \frac{\sqrt{(1 - m)(1 - 9m)}}{4(1 - m)}.$$

This expression is hard to interpret directly, but for small $m$ (such that $m^2 \approx 0$), it is approximately $1 - m$, which shows that migration tends to reduce the proportion of locally optimal behavior as you might expect. Using the exact expression, consensus learning can exclude individual updating when $w_0(1 + b\hat{q}) > w_0(1 + b)c$ and $c > (1 + b/2)/(1 + b)$, which is satisfied when both $m \leq 1/9$ and $1/2 < c < 3/4$.

*mixing is sufficiently weak and individual updating sufficiently costly, then consensus social learning can actually out-compete both individual updating and unbiased social learning.*

When can a consensus learning heuristic invade a population of individual updaters and social learners? We derived above that, when the environment varies spatially, the population will approach a stationary proportion of individual updaters, unless migration is very powerful relative to the value of optimal behavior, in which case individual updating will dominate. At the stationary mix of both heuristics, the expected fitness of both individual updating and unbiased social learning is $w_0(1+b)c$. For consensus learning to invade, it only has to achieve greater fitness than this.

In Box 5, we prove that the condition for consensus social learning to invade a population of individual updating and unbiased social learning is:

$$c > \frac{1 + b/2}{1 + b}. \tag{4}$$

This is easier to satisfy as $c$ increases. This means that consensus learning can invade, provided that individual updating is sufficiently cheap (remember: high $c$ means cheap updating). If $c$ is too small (too costly), then there won't be enough individual updating at equilibrium to keep the average frequency of optimal behavior ($q$) above one-half.

Consensus learning will exclude and replace simple social learning in this environment, whenever it can invade. Perhaps counter-intuitively, if the rate of mixing is low enough, consensus learning can also exclude even individual updating, which simple social learning can never do. We prove this also in Box 5. Provided migration is weak enough and individual updating is sufficiently expensive, but not too expensive, consensus learning can dominate the population entirely. There is an intermediate range of individual updating costs that allow consensus to dominate a population.

The exact result here depends critically on the precise model of consensus learning. However, the qualitative result is likely quite general. Consensus learning is a non-linear form of social learning. As a consequence, it can actively transform the frequency of behavior from one generation to the next. It is a form of "inference," to speak casually. When mixing is weak, this inferential process can substitute for costly individual updating, because the increase in locally optimal behavior consensus learning generates each generation will balance the loss from immigration.

*Payoff bias.*

> Précis. *Payoff biased social learning relies upon the observable consequences of previous choice. As a result, the lower the correlation between observable success and optimal behavior in the relevant domain, the lower the benefit payoff biased learning provides. Payoff biased social learning can, like consensus social learning, replace both unbiased social learning and individual updating, under the right conditions. If migration is weak enough and error in judging payoffs great enough, then consensus social learning can out-compete payoff biased learning.*

Payoff bias can always invade and replace unbiased social learning. The condition for payoff bias to invade a population of individual updaters and unbiased social learners is:

$$w_0(1 + b(q_t + q_t(1 - q_t)(1 - x))) > w_0(1 + bq_t),$$

The above simplifies to $x < 1$ for all $q_t \in [0, 1]$, and so payoff biased learning dominates unbiased social learning whenever there is any correlation between observable success and the behavior of interest.

Like consensus learning, payoff bias is non-linear and actively changes the frequency of adaptive behavior from one generation to the next. Also like consensus learning, this means payoff bias can sometimes exclude individual updating, in this case provided:

$$x < 1 - \frac{m}{1 - m} \cdot \frac{b}{b - (1 - m)((1 + b)c - 1)}. \tag{5}$$

So as long as migration is not too strong and cues of payoffs are sufficiently accurate, it is possible for payoff bias to completely exclude individual updating.

Finally, consensus social learning can sometimes invade and replace payoff biased social learning. Consensus can invade a pure population of payoff biased social learning and replace it, provided:

$$m < \frac{x(1 - x)}{x(1 - x) + 2}.$$

When $x$ is very small, payoff biased learning is highly accurate and therefore unless migration is also very weak, consensus learning lacks a sufficient advantage.

*Summary.* All of the above results explore the properties of unbiased, consensus, and payoff bias social learning when the environment varies spatially. We showed that unbiased social learning can never completely replace individual updating of some kind, because unbiased learning does not transform the frequency of optimal behavior. As a result, it does nothing to modify its social environment for its own good. In contrast, both consensus bias and payoff bias actively transform the frequency of optimal behavior across generations, increasing it slightly above its previous value (assuming $q > 1/2, x < 1$). As a result, both consensus and payoff bias can completely exclude individual updating, provided the force of mixing, $m$, isn't too strong and individual updating is sufficiently costly.

We think such conditions are actually quite rare (and as we show in the next sections, completely absent under our model of temporal environmental variation), but they do reveal a fundamental property of these non-linear social learning heuristics: they actively modify their social environment, and this process can substitute for individual updating, in the right kinds of environments. But each heuristic modifies the social environment using different cues, and therefore they behave differently in different environments. A symptom of this fact is that either consensus or payoff bias can dominate the other, depending upon the amount of mixing among patches ($m$) and the amount of error in judging payoffs ($x$).

**Temporal variation.** When the environment varies through time, instead of across space, many of the same principles that we discovered above hold true. However, there are important differences between temporal variation and spatial variation. Under purely spatial variation in optimal behavior, individuals do well to avoid learning from immigrants to their local patch. However, since the locally optimal behavior does not change over time, a reliable store of locally adaptive culture can accumulate (as long as mixing isn't too strong). Indeed, we showed that both consensus and payoff bias can even exclude individual updating, maintaining optimal behavior at high frequency, even though neither uses individual experience with the environment.

When the environment varies through time, the nature of the problem is subtly different. Now the optimal behavior in each patch will eventually change. When it does, all previously learned behavior is rendered non-optimal. As a result, all social learning heuristics are at a disadvantage, just after a change in the environment. Specifically, we will assume that the environment no longer varies spatially—all patches favor the same behavior. However, there is a chance $u$ in each generation that all patches switch to favoring a new behavior. Since there is a very large number of alternative behaviors, all previously learned behavior is then non-optimal.

This kind of environmental variation also forces us to contend with what evolutionary ecologists call geometric mean fitness (see Orr 2009 for a recent review and comparison of different evolutionary concepts of fitness). When environments vary through time, even rare catastrophe can mean the end of a lineage. As a result, selection may favor risk-averse strategies that are adapted to statistical environments, instead of current environments (Gillespie 1974, Levins 1968).

*Temporal variation may favor randomized or mixed strategies.*

> Précis. *A mixed heuristic is one in which individuals use two or more heuristics different proportions of the time. When the environment varies purely across space,*

*selection does not clearly favor either pure social learning heuristics or mixed social learning heuristics. When the environment varies through time, however, selection favors mixed heuristics over pure ones. In the case of of unbiased social learning and individual updating, selection favors the mixed heuristic over both pure strategies.*

One important result of temporal variation is that a strategy that mixes individual updating and social learning will often out-compete both pure strategies. In the spatial variation case, it makes no obvious difference whether individuals randomly update for themselves or learn socially. But when the environment varies through time, natural selection tends to favor "bet hedging" strategies that engage in adaptive randomization of behavior. The mathematics can be opaque at first, but grasping the cause is easy: survival and reproduction are multiplicative processes. As a result, if a lineage ever is reduced to a very small number of individuals, then it will take it a long time to recover. Therefore a strategy has to both do well and avoid bottlenecks in order to grow quickly and sustain its numbers. When selection varies temporally, selection favors heuristics that attend both mean fitness as well as variance in fitness.

To apply this idea to our social learning heuristics, consider again the basic individual updating versus random target social learning model from before. We showed that there is a stable mix individual updaters and social learners in this model, lying at:

$$\hat{p} = \frac{m\big((1+b)c-1\big)}{(1-m)(1+b)(1-c)}, \tag{6}$$

where $\hat{p}$ is the stable fraction of individual updaters. The average fitness of both individual updaters and social learners at this proportion is the same, $w_0(1+b)c$.

Another way to combine these heuristics is internally, within individuals. Suppose that there is a *mixed* strategy, IU, that uses individual updating a proportion $f$ of the time and unbiased social learning $1-f$ of the time. We prove in Box 6 that natural selection does not distinguish among mixed and pure heuristics in this model. In general, when environmental variation is purely spatial, selection does not clearly distinguish between pure and mixed heuristics. (Although, in very small populations, even this isn't true—see Bergstrom & Godfrey-Smith 1998).

However, when the environment varies temporally, the answer changes. Now, because of the impact of temporal fluctuations in fitness, the environment ends up favoring the mixed heuristic that randomizes its use

**Box 6 Spatial variation does not favor either mixed or pure heuristics**
Consider an alternative "mixed" strategy that internally randomizes between individual updating (I) and unbiased social learning (U). This heuristic's expected fitness is:

$$w(\mathrm{IU}) = f w_0(1+b)c + (1-f)w_0(1+\hat{q}b),$$

where $f$ is the fraction of the time the individual randomly uses individual updating. In order to deduce what value of $f$ natural selection would favor, we find the value of $f$ that maximizes the fitness of the strategy, by solving $\partial w(\mathrm{IU})/\partial f|_{f=f^\star} = 0$ for $f^\star$. This yields:

$$f^\star = \frac{m\big((1+b)c-1\big)}{(1-m)(1+b)(1-c)},$$

which is the same expression as (6), and therefore expected fitness at this optimal value of $f$ is also $w_0(1+b)c$, the same as either pure strategy at equilibrium.

of individual and social updating. This is a result of selection in temporally fluctuating environments depending upon geometric mean fitness, rather than arithmetic mean fitness (see Cooper & Kaplan 1982, Philippi & Seger 1989, for reviews). In general, if the environment varies temporally between two states, each with probability $p_1$ and $p_2$ respectively, then the long-term growth rate is:

$$\bar{r} = w_1^{p_1} w_2^{p_2}.$$

This is in fact the *geometric mean fitness* of the strategy. Evolutionary ecologists usually work with the natural logarithm of this average, $\log[\bar{r}] = p_1 \log[w_1] + p_2 \log[w_2]$.

In the case of analyzing social learning, the state of environment is the time since the environment last changed, and this could be anything from one generation ago to an infinity of generations ago. This might seem daunting at first, but it is really just an application of the logic above, extrapolating from two states of the environment to an infinity of states. The kind of fitness expression we seek is $\log[\bar{r}] = \sum_{i=1}^{n} p_i \log[w_i]$, for $n$ environmental states, where $\sum_{i=1}^{n} p_i = 1$.

Let's take the geometric logic above and apply it to our models of social learning. We will assume again that there are a large number of alternative behaviors, but instead of pure spatial variation, we will now assume that there is purely temporal variation. Each generation, there is a chance $u$ that the environment changes and makes another random behavior optimal, rendering all previously learned behavior non-optimal. Let $p$ again be the frequency of individual updating in the population. In Box 7, we demonstrate how to derive long-run growth rates under temporal environmental variation, in this case using the example of the pure unbiased social learning strategy, U.

**Box 7 Log-geometric growth rate of unbiased social learning**

Let $p$ again be the frequency of individual updating in the population. One generation after a change in the environment, the proportion of adaptive behavior is $q(1) = p$, because individual updaters have had one generation to pump new knowledge into the society. After one more generation without another change, $q(2) = p + (1 - p)q(1) = p + (1 - p)p = 1 - (1 - p)^2$. Then $q(3) = p + (1 - p)q(2) = 1 - (1 - p)^3$. This series continues and implies that, if the environment changed $t > 0$ generations ago, the expected chance of acquiring optimal behavior via social learning is:

$$q(t) = \sum_{i=1}^{t} p(1-p)^{i-1} = 1 - (1-p)^t. \tag{7}$$

We can compute the expected value of $q(t)$ for any generation $t$, if we are willing again to assume that $p$ changes slowly, relative to $q$. In that case, in any given generation, there is a chance $u$ that the environment changed in the most recent generation ($t = 0$), and therefore only those who updated individually have optimal behavior. There is a chance $u(1 - u)$ that the environment changed one generation ago ($t = 1$). In general, there is a chance $u(1 - u)^t$ that the environment last changed $t$ generations ago. Using the definition of log geometric mean fitness, we build the growth rate of an unbiased social learner (U) by multiplying each probability of each environmental state by the log-fitness in that state:

$$r(\mathrm{U}) = u \log[w_0] + \sum_{t=1}^{\infty} u(1-u)^t \log[w_0(1 + bq(t))].$$

The above can be motivated in the following way. Expected fitness is the product of each fitness raised to the probability of its occurrence, so the expected log-fitness is the sum of each log-fitness multiplied by the probability it occurs. If the environment has just changed, which happens $u$ of the time, then the individual receives $w_0$. The other possibility is that the environment has not changed in the last $t$ generations, yielding a chance $q(t)$ of fitness $w_0(1 + b)$ and a chance $1 - q(t)$ of fitness $w_0$, for each individual using this heuristic. That is, every social learner experiences the same $q(t)$ in any generation $t$, and the value of $q(t)$ is determined by the probabilities $u(1 - u)^t$. But a proportion $q(t)$ of social learners will get lucky and choose a target with optimal behavior, while the rest will not. Thus the probability $q(t)$ is inside the logarithm.

Once we have a geometric fitness expression for a heuristic, we can analyze its evolutionary dynamics. Unfortunately, the form of this expression in this case makes the mathematics intractable. There are no algebraic methods for closing such an infinite sum, in which the power $t$ is both outside and inside the logarithm. We can make progress, however, by constructing an approximation that is valid for *weak selection.* In Box 8, we demonstrate how to construct weak selection approximations for this model.

**Box 8 Weak selection approximation**

Weak selection applies when $b$ and $c$ are small such that terms of order $b^2$ and $(1-c)^2$ and greater are approximately zero. The use of weak selection approximations is common in evolutionary ecology, because it often makes otherwise intractable problems analytically solvable. One must keep in mind, however, that our conclusions from here out will only be exactly valid for choices that have modest effects on fitness. Numerical work, as well as the simulations we show in a later section, confirm that the qualitative conclusions we reach here are general to strong selection, however. To apply the weak selection approximation, we use a Taylor series expansion of $r(\mathrm{U})$ around $b = 0, c = 1$ and keep the linear terms in $b, c$. This gives us:

$$r(\mathrm{U}) \approx \log[w_0] + \frac{(1-u)bp}{1-(1-p)(1-u)}.$$

We wish to compare this expression to the weak-selection approximation of the growth rate of individual updating, which by the same method is $r(\mathrm{I}) \approx \log[w_0] + b - (1-c)$. Selection will adjust $p$ until $r(\mathrm{U}) = r(\mathrm{I})$, which implies an expected long-run value of $p$:

$$\bar{p} = \frac{u(b-(1-c))}{(1-u)(1-c)}. \tag{8}$$

This is the same as Expression 6, once we apply the weak selection approximation and let $m = u$. We will need this result in order to compare the two pure heuristics to the mixed heuristic in the temporal variation context.

Finally we are ready to again consider, now in the context of temporal variation, the family of alternative heuristics that randomize their updating, using individual updating (I) a proportion $f$ of the time and unbiased social updating (U) a proportion $1-f$ of the time. Using the same logic that allows writing the expression $r(\mathrm{U})$ in Box 7, the long-run growth of this mixed heuristic is:

$$r(\mathrm{IU}) = \sum_{t=0}^{\infty} u(1-u)^t \log[f w_0(1+b)c + (1-f)w_0(1+bq(t))].$$

Using this expression, we prove in Box 9 that the mixed heuristic that randomly uses individual and social updating both invades and replaces a population of pure individual and social updaters. This was not the case for purely spatial environmental variation. The reason for the different result owes to bet-hedging (Philippi & Seger 1989) against the small payoff to social updaters, soon after a change in the environment. Because the mixed heuristic spreads its bets over two different portfolio options—individual updating and social updating—it experiences reduced risk of ruin, like pure individual updating, but also reaps higher rewards when the quality of socially learned behavior is high, like unbiased social learning. This is mathematically homologous to

**Box 9 Evolutionary dynamics of IU under weak selection**

Assuming the IU heuristic is rare and that $q(t)$ is therefore determined only by the proportions of pure individual and social updaters, $p$, and again that selection is weak, the above closes to:

$$r(\mathrm{IU}) \approx \frac{b\big(p(1-u)+uf(1-(1-c)(1-f))\big)}{1-(1-p)(1-u)} + (1-c)f + \log[w_0].$$

By comparing the above to the growth rates for the pure strategies, it turns out that this mixed strategy can invade the stable mix of pure strategies, for any value of $f$. The condition for a rare IU individual to invade the mix of I and U is $r(\mathrm{IU})|_{p=\bar{p}} > r(\mathrm{I})$, where $\bar{p}$ is defined by Expression 8 in Box 8. This condition is satisfied for all $b > 0, 0 < c < 1, 0 < u < 1$, which means the mixed strategy can always invade a population of pure heuristics. This kind of result is typical of game theoretic solutions of this kind. The value of $f$ does not matter for invasion, because whatever the value of $f$, the first mixed strategy individual will behaviorally simulate either a pure I or a pure U individual. Since both I and U have the same fitness at $p = \bar{p}$, it makes no difference which of these heuristics is realized. The value of $f$ will matter, however, as IU increases in frequency. Once common, it turns out that the mixed heuristic IU is also always stable against invasion by pure I and U. To prove this we need to calculate the optimal value of $f = f^\star$ that no other value of $f$ can invade. When IU is common, the proportion of optimal behavior is now given by $q(t) = \sum_{i=1}^{t} f^\star(1-f^\star)^{i-1}$, where $f^\star$ is the common chance a IU individual updates individually. The evolutionarily stable value of $f^\star$ is found where $\partial r(\mathrm{IU})/\partial f|_{f=f^\star} = 0$. Again using a weak selection approximation and solving the above for $f^\star$ yields:

$$f^\star = \frac{u\big((1+b)c - 1 + u((1-b)(1-c)-b)\big)}{(1-c)(1-u)^2}. \tag{9}$$

By substituting the value of $f^\star$ back into $r(\mathrm{IU})$, one can derive the growth rate of the mixed strategy when it is common and using the optimal value of $f$. We ask when $r(\mathrm{IU})|_{f=p=f^\star} > r(\mathrm{I})$ and when $r(\mathrm{IU})|_{f=p=f^\star} > r(\mathrm{U})|_{p=f^\star}$. Both of these conditions are true for all $b > 0, 0 < c < 1, 0 < u < 1$, and so the mixed heuristic can both invade a population of pure heuristics as well as resist invasion by either pure heuristic.

human investors spreading risk over multiple stocks, because even if the chance of any individual stock losing most of its value is low, if all assets are placed in a single stock, eventually all of one's assets will lose most of their value. Similarly, the mixed heuristic out-performs both pure heuristics, because this strategy is never entirely ruined by a change in the environment, but neither does it entirely forgo the gains to social learning that accrue when the environment remains stable. Both pure heuristics instead take risks by betting entirely on one kind of event or the other.

The important lesson here is that temporal variation strongly favors a fundamentally different way of combining heuristics, mixing them within individuals rather than among individuals. In this way, the physical environment, whether variation is spatial or temporal, favors different social learning strategies and combinations of those strategies. This in turn leads us to make different predictions about the kinds of heuristics that will be adapted to different domains of behavior, depending in part upon the relative strengths of spatial and temporal variation.

It should be noted, however, that the model in this section is still just a model—it is limited to fairly particular assumptions about the environment and the heuristics. The temporal variation here is not auto-correlated—if the environment has just changed, it is no more or less likely to change again. Real environments, ecological measurements suggest, tend to include a good amount of autocorrelation, as evidenced by "red" noise in their time series (Whitehead & Richerson 2009). While we see no reason that the conclusions here should not extend, qualitatively, to autocorrelated environments, we do believe that it is a problem worth modeling.

*Consensus learning less favored under temporal variation.*

> Précis. *When the environment varies through time, consensus social learning does much worse than it does under spatial variation. It can never exclude individual updating, because every time the environment changes, $q < 1/2$ for at least one generation. As a result, consensus learners do badly compared to unbiased learners. However, adding in some spatial variation as well helps consensus learning recover.*

In the case of purely spatial variation, we already demonstrated that consensus social learning can in fact exclude both simple social learning and individual updating, provided the rate of mixing between locales is sufficiently low. Under purely temporal variation, consensus learning can never exclude the other learning heuristics. Just after a change in the optimal behavior, all previously-learned behavior is non-optimal. Therefore inferring behavior from the majority will lead to stabilizing non-adaptive behavior. As a result, consensus learning depends upon some other heuristic—or mix of heuristics—to increase the frequency of newly-optimal behavior, after a change in the environment.

The mathematics of this case are complex, because accounting for geometric fitness effects and the non-linearities of consensus learning is analytically difficult. But the results are easy to visualize in simulation from the fitness definitions. (The short simulation code is contained in the short appendix at the end of this chapter.) Figure 2 plots the

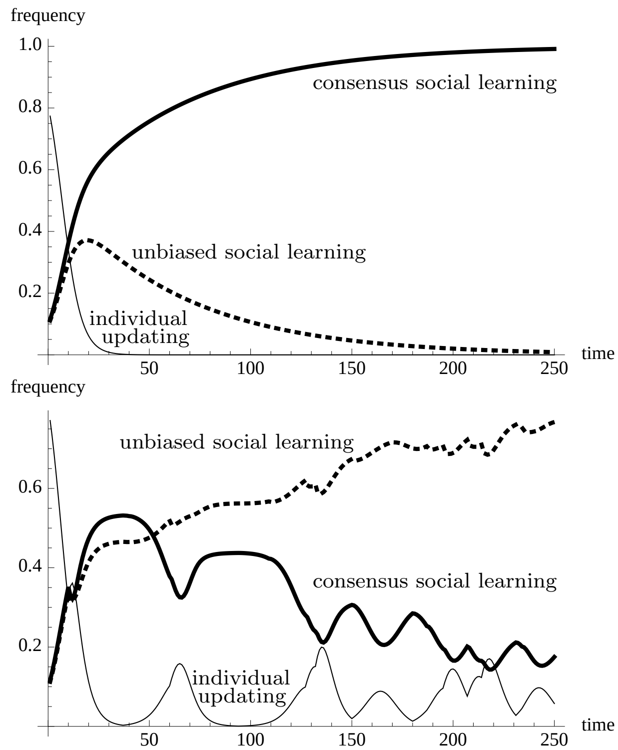


FIGURE 2. Stochastic simulations of the evolution of consensus learning, under either spatial or temporal environmental variation. The thin solid trend in each panel is the proportion of the population that uses individual updating. The thick dotted and thick solid curves are unbiased social learning and consensus learning, respectively. In both panels, $b = 0.5, c = 0.8$. Initial proportions for individual, unbiased, and consensus: $0.8, 0.1, 0.1$. Top panel: Purely spatial variation, $m = 0.05, u = 0$. Consensus learning evolves to exclude both individual and unbiased social learning. Bottom panel: Purely temporal variation, $m = 0, u = 0.05$. Consensus learning now does much worse, owing to its inability to increase the frequency of rare, novel adaptive behavior.

proportions of individual updating (thin solid trend), unbiased social learning (thick dotted trend), and consensus learning (thick solid trend) through time, for both pure spatial and pure temporal environmental variation. In the absence of temporal variation in optimal behavior, consensus learning can actually exclude both individual updating and

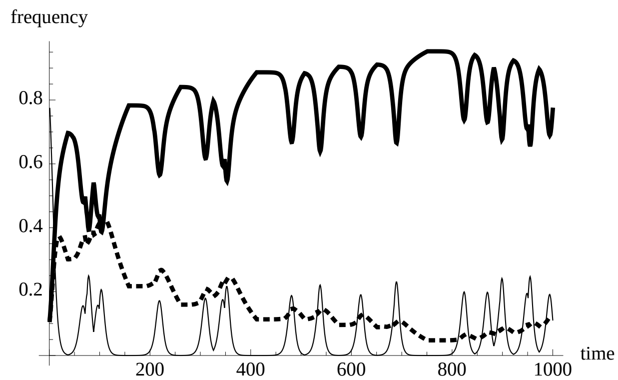


Figure 3. Stochastic simulations of the evolution of consensus learning, under simultaneous spatial and temporal environmental variation. The thin solid trend is the proportion of the population that uses individual updating. The thick dotted and thick solid curves are unbiased social learning and consensus learning, respectively. Same parameter values and initial proportions as in Figure 2. Consensus social learning can out-compete unbiased social learning, even in the presence of temporal variation, provided there is enough mixing among spatially variable patches. Here, $u = m = 0.05$.

unbiased social learning (top panel). However, under temporal variation in the environment, consensus learning does quite poorly, owing to its drop in frequency each time the environment shifts from one optimal behavior to another. A mix of individual updating and unbiased social learning evolves.

A small amount of spatial variation and mixing can go a long way towards helping consensus social learning, however (Figure 3). While temporal variation hurts consensus learning much more than it hurts unbiased social learning, spatial variation and mixing hurts unbiased learning more than it hurts consensus social learning. After a change in the environment, consensus social learners suffer reduced fitness, declining in frequency as individual updating increases in frequency (see time series in Figure 3). But once the local frequency of optimal behavior has increased, unbiased social learners have no particular advantage over consensus social learners. Meanwhile, consensus social learners avoid learning from immigrants with behavior adapted to other patches, while unbiased social learners do not. Reintroducing mixing among spatially variable patches provides a constant environmental challenge that partially resuscitates consensus learning. Therefore it is not a valid conclusion that consensus learning is poorly adapted whenever environments vary through time. Instead, we should conclude that temporal variation works against consensus learning, while mixing and spatial variation work for it. If either force is strong enough, it can eclipse the other.

We caution the reader to note that the same principle will hold in the case of consensus social learning in temporally varying environments that holds for unbiased social learning: a mixed individual updating and consensus strategy will do better than a mix of pure strategies. We do not belabor this point here, because we know of no additional intuitions to be acquired from the analysis. But we do not want readers to conclude that mixed randomizing heuristics would not be favored for consensus learning as they would be for unbiased social learning. Indeed, the bet-hedging will arguably be stronger in the case of consensus learning, because the effects of a recent change in the environment are harsher for consensus learning than they are for unbiased social learning. At the same time, since consensus learning can drive the proportion of optimal behavior both down (when $q < 1/2$) as well as upwards (when $q > 1/2$), the dynamics may be much more complex and interesting.

*Summary.* In this section, we analyzed the effects of temporal environmental variation on unbiased and consensus social learning heuristics. Temporal variation requires a different approach to calculating evolutionarily-relevant payoffs, because if a strategy is reduced to zero numbers in any generation, then the strategy is dead forever. This "bottleneck" effect can have important consequences for the evolutionary rationality of heuristics. This principle lead us to two main results.

First, temporal variation can favor internally mixed heuristics, when purely spatial variation does not. The reason is that temporal variation favors bet-hedging heuristics that spread risk across alternative behavioral strategies. In this case, a mixed strategy that randomly deploys individual updating and unbiased social updating always replaces a population of pure individual updating and unbiased social learning strategies, when there is purely temporal variation in the environment.

Second, consensus social learning is disadvantaged under temporal variation. The reason is that, just after a change in the environment, all learned behavior is non-optimal. As a consequence, the majority behavior provides an invalid cue to optimality, in this context. Once some other heuristic or set of heuristics have again increased optimal behavior in the population, consensus can do well, but lost fitness during the transition can cause it to be out-competed by other learning heuristics. This does not happen under purely spatial variation, because a constant stream of immigrants actually provides an environmental challenge that consensus learning is well-suited to, provided that mixing is not so strong as to make the majority of local behavior non-optimal. Simultaneously combining spatial and temporal variation shows that consensus learning can be profitable when temporal variation is present, provided that there is enough spatial mixing and spatial variation.

## Conclusion

We have tried to show how the way in which optimal behavior varies across space or through time favors different social learning heuristics. We are skeptical that there will be any learning strategy, social or not, that is best in all contexts. Instead the type of analysis in this chapter suggests that either over evolutionary or developmental time, individuals acquire strategies that exploit patterns in specific environments. In this way, the tradition in evolutionary ecology of studying cognitive adaptation via social learning is quite similar to the tradition in bounded rationality. And like some analyses in bounded rationality, the environments in this chapter are statistical. Instead of adapting to a single state of the world, the theoretical organisms in our thought experiments adapt to a statistical world in which randomness and variation present survival challenges. Successful heuristics are the ones that out-reproduce competitors over many generations of learning and choice, sometimes hedging their bets against unpredictable bad times. In any particular generation, a social learning heuristic can appear nonsensical. It is in the long run, across the full distribution of environmental dynamics, that the evolutionary rationality of each heuristic appears.

The breadth of issues relevant to the evolution of social learning is huge. We have focused on the nature of environmental variation, because this topic has long been central to the study of learning, social or not, in evolutionary ecology (Levins 1968). Indeed, fluctuating selection has turned out to be central to broad debates that touch upon most corners of evolutionary biology (Gillespie 1994). Organisms are not adapted to a static world of stationary challenges, but rather a mosaic world that varies across space and fluctuates through time. A satisfactory account of the design of heuristics will include consideration of this fact, even if analyzing static decision problems is often a necessary step.

The precise kind of variation affected our conclusions. This result reinforces the message that consideration of a stochastic world will have an important role to play in the study of heuristics, social or otherwise. In some cases, even scholars studying the evolution of social learning in fluctuating environments have missed the importance of the precise assumptions about the nature of the statistical environment. Wakano & Aoki (2007) analyzed a model of the evolution of consensus social learning in a temporally varying environment and find that they reach different conclusions than Henrich & Boyd (1998), who studied the evolution of consensus learning under simultaneous spatial and temporal variation. As we showed in this chapter, temporal variation selects against consensus social learning in a way that spatial variation does not. Although Wakano & Aoki (2007) note the different assumptions about the nature of environment, they decide without analysis that the divergent assumptions have no role in explaining their divergent results. They instead speculate that Henrich and Boyd did not run their simulations to convergence. Explicitly testing the different performance of consensus learning under both models of environmental variation would have shed more light on the issue. Whitehead & Richerson (2009) use simulations to demonstrate that indeed some kinds of temporal variation are worse for consensus learning than others, serving to re-emphasize the importance of exactly what we assume in the statistical model of the environment.

More broadly, the analysis of simple social learning strategies strongly suggests that some kind of social learning will be adaptive, unless environments are extremely unpredictable. While the things people say and do are not always locally adaptive, the very action of a toolbox of social and individual updating heuristics can help construct social environments in which it is worth attending to the beliefs of others. These thought experiments therefore suggest one reason that people are so powerfully influenced by mere words.

## Appendix: Simulation code

This code runs in Mathematica version 7, although it should run without modification in earlier versions as well. By changing the values of the parameters $u$ and $m$, one can replicate the results in the figures in the main text.

```
p = 0.8; pc = 0.1; q = 0; b = 0.5; c = 0.8; s = 1;
u = 0.0;
```

```
m = 0.05;
hp = {}; hpc = {}; hq = {}; w0 = 1;
For[t = 0, t < 1000, t++,
  r = RandomReal[];
  q = (1 - m)*
    If[r < u, 0,
     p s + (1 - p - pc) q + pc (q + q (1 - q) (2 q - 1))];
  wi = w0 (1 + s b) c;
  wu = w0 (1 + b q);
  wc = w0 (1 + b (q + q (1 - q) (2 q - 1)));
  wbar = p wi + (1 - p - pc) wu + pc wc;
  pp = N[p wi/wbar];
  ppc = N[pc wc/wbar];
  AppendTo[hp, pp];
  AppendTo[hpc, ppc];
  AppendTo[hq, q];
  p = pp;
  pc = ppc;
  ];
Clear[pp, p, q, b, c, s, u, w0, f, pc, ppc, m];
ListPlot[{hp, hpc, 1 - hp - hpc}, Joined -> True,
 PlotStyle -> {{Black}, {Black, Thick}, {Black, Thick, Dotted}},
 AxesLabel -> {time, frequency}]
```

## References


Aoki, K., Wakano, J., & Feldman, M. W. (2005). The emergence of social learning in a temporally changing environment: A theoretical model. *Current Anthropology*, 46:334–340.

Bauch, C. T. (2005). Imitation dynamics predict vaccinating behaviour. *Proc. R. Soc. B*, 272:1669–1675.

Bergstrom, C. T. & Godfrey-Smith, P. (1998). Pure versus mixed strategists: the evolution of behavioral heterogeneity in individuals and populations. *Biology and Philosophy*, 13:205–231.

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

Bonner, J. T. (1980). *The evolution of culture in animals.* Princeton University Press, Princeton.

Boyd, R. & Richerson, P. J. (1985). *Culture and the Evolutionary Process.* Univ Chicago Press, Chicago.

Boyd, R. & Richerson, P. J. (1988). An evolutionary model of social learning: the effects of spatial and temporal variation. In Zentall,

T. R. & B. G. Galef, J., editors, *Social Learning: Psychological and Biological Approaches*, pages 29–47. Erlbaum, Hillsdale, New Jersey.

Boyd, R. & Richerson, P. J. (1995). Why does culture increase human adaptability? *Ethology and Sociobiology*, 16:125–143.

Boyd, R. & Richerson, P. J. (2001). Norms and bounded rationality. In Gigerenzer, G. & Selten, R., editors, *The Adaptive Tool Box*, pages 281–296. MIT Press, Cambridge MA.

Boyd, R. & Richerson, P. J. (2005). Rationality, imitation, and tradition. In *The Origin and Evolution of Cultures*, chapter 18, pages 379–396. Oxford University Press, Oxford.

Cavalli-Sforza, L. L. & Feldman, M. W. (1981). *Cultural transmission and evolution: a quantitative approach*. Princeton University Press, Princeton.

Cooper, W. S. & Kaplan, R. H. (1982). Adaptive "coin-flipping": a decision-theoretic examination of natural selection for random individual variation. *Journal of Theoretical Biology*, 94:135–151.

Dunlap, A. S. & Stephens, D. W. (2009). Components of change in the evolution of learning and unlearned preference. *Proc. R. Soc. B*, doi:10.1098/rspb.2009.0602.

Efferson, C., Lalive, R., Richerson, P. J., McElreath, R., & Lubell, M. (2008). Conformists and mavericks: The empirics of frequency-dependent cultural transmission. *Evolution and Human Behavior*, 29:56–64.

Evans-Pritchard, E. E. (1937). *Witchcraft, Oracles and Magic Among the Azande*. Oxford University Press, Oxford.

Fragaszy, D. & Perry, S., editors (2003). *The Biology of Traditions: Models and Evidence*. Cambridge University Press.

Galef, B. G. (1992). The question of animal culture. *Human Nature*, 3(2):157–178.

Galef, B.G., J. (1996). Social enhancement of food preferences in norway rats: A brief review. In *Social Learning and Imitation: The Roots of Culture*.

Gillespie, J. H. (1974). Natural selection for within-generation variance in offspring number. *Genetics*, 76:601–606.

Gillespie, J. H. (1994). *The Causes of Molecular Evolution*. Oxford University Press, Oxford.

Gintis, H. (2000). *Game Theory Evolving*. Princeton University Press, Princeton.

Giraldeau, L.-A. & Caraco, T. (2000). *Social foraging theory*. Princeton University Press, Princeton.

Henrich, J. (2001). Cultural transmission and the diffusion of innovations: Adoption dynamics indicate that biased cultural transmission is the predominate force in behavioral change and much of sociocultural evolution. *American Anthropologist*, 103:992–1013.

Henrich, J. & Boyd, R. (1998). The evolution of conformist transmission and the emergence of between-group differences. *Evolution and Human Behavior*, 19:215–242.

Henrich, J. & Boyd, R. (2001). Why people punish defectors: Weak conformist transmission can stabilize costly enforcement of norms in cooperative dilemmas. *Journal of Theoretical Biology*, 208:79–89.

Henrich, J. & Gil-White, F. (2001). The evolution of prestige: Freely conferred status as a mechanism for enhancing the benefits of cultural transmission. *Evolution and Human Behavior*, 22:165–196.

Henrich, J. & McElreath, R. (2003). The evolution of cultural evolution. *Evolutionary Anthropology*, 12:123–135.

Laland, R. N. & Galef, B., editors (2009). *The question of animal culture*. Harvard University Press, Cambridge, MA.

Lande, R. (2009). Adaptation to an extraordinary environment by evolution of phenotypic plasticity and genetic assimilation. *Journal of Evolutionary Biology*, 22:1435.

Lehmann, L. & Feldman, M. W. (2008). The co-evolution of culturally inherited altruistic helping and cultural transmission under random group formation. *Theoretical Population Biology*, 73:506–516.

Levins, R. (1968). *Evolution in Changing Environments*. Princeton University Press, Princeton.

McElreath, R., Bell, A. V., Efferson, C., Lubell, M., Richerson, P. J., & Waring, T. (2008). Beyond existence and aiming outside the laboratory: Estimating frequency-dependent and payoff-biased social learning strategies. *Philosophical Transactions of the Royal Society B*, 363:3515–3528.

McElreath, R. & Boyd, R. (2007). *Mathematical Models of Social Evolution: A Guide for the Perplexed*. Univ of Chicago Press, Chicago.

McElreath, R., Boyd, R., & Richerson, P. J. (2003). Shared norms and the evolution of ethnic markers. *Current Anthropology*, 44:122–129.

McElreath, R., Lubell, M., Richerson, P. J., Waring, T., Baum, W., Edsten, E., Efferson, C., & Paciotti, B. (2005). Applying evolutionary models to the laboratory study of social learning. *Evolution and Human Behavior*, 26:483–508.

McElreath, R. & Strimling, P. (2008). When natural selection favors imitation of parents. *Current Anthropology*, 49:307–316.

Mesoudi, A. (2008). An experimental simulation of the "copy-successful-individuals" cultural learning strategy: Adaptive landscapes, producer-scrounger dynamics and informational access costs. *Evolution and Human Behavior*, 29:350–363.

Mesoudi, A. & Lycett, S. J. (2009). Random copying, frequency-dependent copying and culture change. *Evolution and Human Behavior*, 30:41–48.

Mesoudi, A. & O'Brien, M. (2008). The cultural transmission of great basin projectile point technology i: An experimental simulation. *American Antiquity*, 73:3–28.

Mesoudi, A. & Whiten, A. (2008). The multiple roles of cultural transmission experiments in understanding human cultural evolution. *Philosophical Transactions of the Royal Society B*, 363:3489–3501.

Orr, H. A. (2009). Fitness and its role in evolutionary genetics. *Nature Reviews Genetics*, doi:10.1038/nrg2603.

Philippi, T. & Seger, J. (1989). Hedging one's evolutionary bets, revisited. *Trends in Ecology and Evolution*, 4(2):41–44.

Price, E. E., Lambeth, S. P., Schapiro, S. J., & Whiten, A. (2009). A potent effect of observational learning on chimpanzee tool construction. *Proc. R. Soc. B*, doi: 10.1098/rspb.2009.0640.

Richerson, P. J. & Boyd, R. (2005). *Not by Genes Alone: How culture transformed human evolution.* University of Chicago Press, Chicago.

Rogers, A. R. (1988). Does biology constrain culture? *American Anthropologist*, 90:819–831.

Schlag, K. H. (1998). Why imitate, and if so, how? *Journal of Economic Theory*, 78(1):130–156. Studies; Experimental English.

Schlag, K. H. (1999). Which one should I imitate? *Journal of Mathematical Economics*, 31(4):493–522. Studies; Experimental English.

Stahl, D. O. (2000). Rule learning in symmetric normal-form games: Theory and evidence. *Games and Economic Behavior*, 32:105–138.

Wakano, J. & Aoki, K. (2007). Do social learning and conformist bias coevolve? henrich and boyd revisited. *Theoretical Population Biology*, 72:504–512.

Wakano, J., Aoki, K., & Feldman, M. (2004). Evolution of social learning: a mathematical analysis. *Theoretical Population Biology*, 66:249–258.

Whitehead, H. & Richerson, P. (2009). The evolution of conformist social learning can cause population collapse in realistically variable environments. *Evolution and Human Behavior*, 30:261–273.

McElreath: mcelreath@ucdavis.edu. Department of Anthropology and Graduate Groups in Ecology, Population Biology and Animal Behavior, University of California, Davis CA 95616, USA.

Fasolo: Department of Management (Operational Research Group), London School of Economics and Political Science, G313 Houghton Street, London WC2A 2AE UK.

Wallin: Department of Philosophy, Lund University, Kungshuset, Lundagård, 222 22 Lund SE.