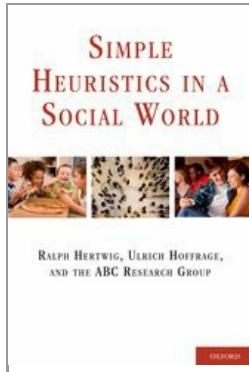University Press Scholarship Online

Oxford Scholarship Online

Simple Heuristics in a Social World

Ralph Hertwig, Ulrich Hoffrage, and ABC Research Group

The Evolutionary Rationality of Social Learning

Richard McElreath

Annika Wallin

Barbara Fasolo

[−] Abstract and Keywords

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 environment. This chapter reviews a slice of this literature, as well as explicitly analyze the evolution of social learning heuristics. The chapter outlines a family of social learning heuristics and analyze their evolutionary performance under two broadly different kinds of environmental variation. As each social learning heuristic also shapes a social environment as individuals use it, the chapter considers the population feedbacks of each heuristic as well. The analyses in this chapter are both ecological and game theoretic. This chapter's 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 the conclusions reflect an evolutionary rationality.

For example, heuristics that randomize their behavior can succeed where those that are consistent fail. Overall, however, the approach the chapter reviews here supports the general conclusion that social learning heuristics are likely to be multiple and subtly adapted to different physical, statistical, and social environments.

A common premise in 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, not only of 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 that 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. Effects like these reveal how hard it is to really control our thoughts, as well as the power that mere utterances have over us. Of course, people are savvy and do not robotically obey all suggestions or commands. 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 that 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.

(p.382) 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 information from their social environment to arrive at inferences about the reasons for and consequences of behavior. And animals other than humans also use social information (e.g., Bonner, 1980; Galef, 1992, 1996; Giraldeau & Caraco, 2000; Fragaszy & Perry, 2003; Laland & Galef, 2009; Price, Lambeth, Schapiro, & Whiten, 2009). Although the psychological mechanisms and diversity of social learning among, for example, baboons is not the same as that among humans, the savvy monkey also 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 will review a slice of this literature, as well as explicitly analyze the

evolution of social learning heuristics. A social learning heuristic is, for instance, to learn from the most successful individual in the surroundings (*payoff-biased learning*; see Table 14-1), or to pick up the behavior that appears to be most common in one's environment (*consensus learning*). The social learning heuristics discussed in this chapter are constructed with this evolutionary analysis in mind, but do contain assumptions about how information is searched for, when search stops, and how this information is combined into a decision (Gigerenzer, Todd, & the ABC Research Group, 1999). For instance, consensus learning is modeled by assuming that each individual randomly samples the behavior of three individuals surrounding it, and preferentially adopts the majority behavior.

In the chapter, we outline a family of social learning heuristics and analyze their evolutionary performance—their 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 changes the success rate of this heuristic (and that of others), 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 their long-term survival and reproduction in a population, not atomistic one-shot **(p.383)**

Table 14-1: Major Social Learning Heuristics from the Literature, with Other Names for the Same Strategies, and Citations for a Sample of Previous Evolutionary Analysis

Heuristic	Other names	Citations
Unbiased social learning	Linear social learning, random copying, imitation	Aoki, Wakano, & Feldman (2005); Boyd & Richerson (1985, 1995); Cavalli-Sforza & Feldman (1981); Mesoudi & Lycett (2009); Rogers (1988); Wakano, Aoki, & Feldman (2004)
Consensus learning	Conformity, conformist transmission, positive frequency dependent imitation, majority rule imitation	Boyd & Richerson (1985); Henrich & Boyd (1998, 2001); Lehmann & Feldman (2008); Mesoudi & Lycett (2009); 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

payoffs. As a result, some of our conclusions reflect an evolutionary rationality that is

sometimes counterintuitive. For example, heuristics that randomize their behavior can succeed where those that are consistent fail. Overall, however, the approach that 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 on bounded rationality (Simon, 1955b), evolutionary ecologists and anthropologists studying social learning have proposed that there exists a toolbox of contextually deployed heuristics, similar to the adaptive toolbox, that are suited to different ecological and social environments (reviews in Henrich & McElreath, 2003; Richerson & Boyd, 2005b). 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 **(p.384)** than others. Thus the expectation is that there are many inferential strategies that individuals can use to choose behavior. Although some of these strategies are more cognitively demanding and information-hungry than others, all are quite "bounded," compared to Bayesian justifications for social learning (Bikhchandani, Hirshleifer, & Welch, 1992; Boyd & Richerson, 2001, 2005b). 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, Lalive, Richerson, McElreath, & Lubell, 2008; McElreath et al., 2005, 2008; Mesoudi, 2008; Mesoudi & O'Brien, 2008; Mesoudi & Whiten, 2008).

In this section, we will outline and begin to analyze a toolbox of social learning heuristics that evolutionary ecologists and evolutionary anthropologists have studied, using both empirical and analytical methods. The collection of heuristics that we review is not complete. Many other heuristics could be nominated, and each heuristic we do nominate is in reality a family of heuristics. By constraining our discussion to the most commonly discussed strategies, however, 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, in terms of both physical and social environments, that favor different heuristics. In Table 14-1, 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 will demonstrate the analysis of a few of these. We will also present a new analysis of the evolution of heuristics in time-varying environments. The dynamical systems approach common in evolutionary analysis may be unfamiliar to many readers, so we provide a quick guide in Box 14-1 to some major concepts.

The Environmental Challenge

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

Assume that 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 $1 + b \rangle$ 1. All other behavior leaves **(p.386)** fitness unchanged. In particular, let w_0 be an individual's fitness before behaving. Because there are a great many alternative choices, randomly guessing will not yield a fitness payoff much greater than w_0. Then those who choose optimally have fitness $w_0(1 + b)$, whereas those who do not, have fitness w_0. Because fitness does not depend upon how many other individuals also choose the same option, these payoffs are not directly frequency-dependent.

(p.385) Box 14-1: Readers' 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 and Boyd (2007).

Population: All organisms of the same species that 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 subpopulations 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 (DNA-sequences) 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 on 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 analysis.

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 evolutionary 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.

Individual Updating

The foil for all the social learning heuristics that 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 its survival multiplied by $c \in [0, 1]$. This means that the fitness of an individual updater is always $w_0(1 + b)c \, \rangle \, w_0$. 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, although biologically satisfying, adds little in terms of understanding. It changes none of the qualitative results that 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 (an individual to learn from). Much of the earliest evolutionary work on social learning has considered this strategy (Boyd & Richerson, 1985; Cavalli-Sforza & Feldman, 1981), and even more recent work continues to study its properties (Aoki, Wakano, & Feldman, 2005; Wakano, Aoki, & Feldman, 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 (see Table 14-1).

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 processes, not deviations from **(p.387)** normative standards. The word "bias" has been used in this way for some time in the evolutionary study of social learning (Boyd & Richerson, 1985, for example).

Let q ("quality" of social information) be the proportion of optimal behavior among possible targets (the individuals to learn from) 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 will acquire optimal behavior. Like other social learning heuristics, unbiased social learning actively shapes social environments itself—a heuristic that uses behavior as a cue and produces behavior that 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 that 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 learning (below), in our simple model of it, uses three targets, and payoff-biased 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 that 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, whereas the payoff-biased learner need only observe payoffs and then invest time observing one target. We could invent stories that favor consensus, as well. And although constructing and formalizing such stories is likely instructive, it is a sufficiently detailed project that we have not undertaken it in this chapter. But we do not wish to convey 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 our investigation.

Consensus Learning

An often-discussed category of social learning heuristics is 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 **(p.388)** as a cue to guide choice. This kind of strategy has been called *positive frequency dependence* and *conformist transmission*. We adopt the label "consensus" 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), and because the alternative "positive frequency dependence" is an unwieldy term.

Consensus learning can be most easily modeled by assuming that an individual samples three targets at random and preferentially adopts the most common behavior among them. In Box 14-2, we show how to use this definition to derive the expected probability that an individual using consensus learning will acquire optimal behavior:

(1) $\Pr(1) = q + q(1-q)(2q-1).$

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, because our goal is to **(p.389)** motivate a mode of analysis and to emphasize the differences among quite different social learning heuristics, rather than among variants of the same heuristics.

Box 14-2: Deriving Consensus Learning

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

Observed behavior	Pr(Obs)	Pr(1)	Pr(0)
111	q3	1	0
110	3q2(1-q)	2/3+D	1/3-D
100	3q(1-q)2	1/3-D	2/3+D
000	(1-q)3	0	1

$0 \langle D$ 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 Equation 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 one or a 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.

Payoff-Biased Learning

Another often-analyzed category of social learning heuristic is payoff-biased learning (Boyd & Richerson, 1985; Schlag, 1998, 1999; Stahl, 2000). By comparing observable payoffs—health, having surviving offspring, or even more domain-specific measures of success—among targets, an individual can use differences in 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, Boyd, & Richerson, 2003) and even epidemiology (Bauch, 2005).

A simple model of payoff-biased learning assumes that individuals sample two targets and preferentially adopt the behavior of the target with the higher observed payoff. This means that if one of the targets is observed to be more successful, then it is more likely that the individual will behave as this target does. 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 14-3 that this heuristic leads to a chance

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

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. For instance, the dynamics of academic writing will change depending on whether the number of publications, or the impact factor of these publications, is seen as the best cue to academic success. When this is the case, unanticipated social processes become possible, such as the runaway exaggeration of preferences for traits that are judged as successful (Boyd & Richerson, 1985).

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 learning, and payoff-biased 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: (a) a spatially variable environment, in which different behavior is optimal in different places,

and (b) a temporally variable environment, in which different behavior is optimal at different times. A spatially variable environment could, for instance, be an environment where prey is more readily available in certain areas than in others. In contrast, we can think of a "temporally variable" environment as farmland where crops do well or poorly depending on seasonal variation. We also consider the interaction of these two kinds of variation (Figure 14-1).
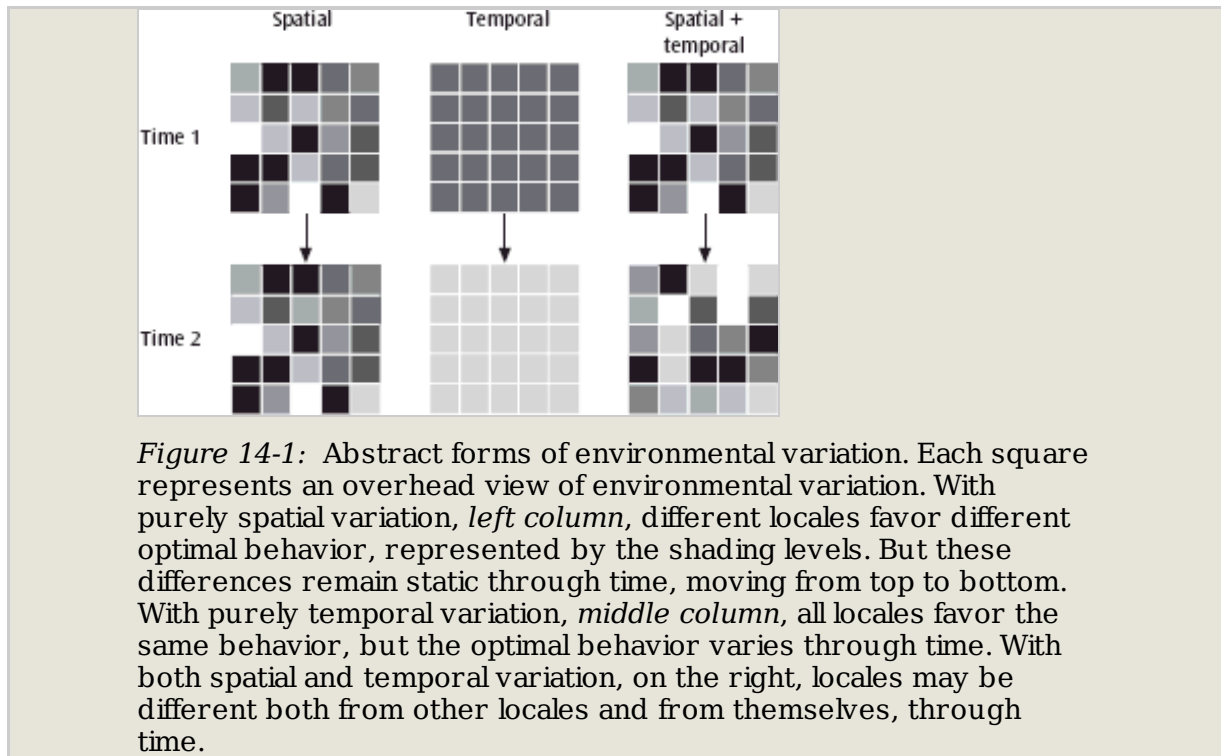

(p.390) Box 14-3: Deriving Payoff-Biased Learning

Again we use a table to derive probabilities of acquiring optimal (1) and non-optimal (0) behavior, this time using payoff-biased 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	$x(1-x)$	½	½
		1 1	$x(1-x)$	½	½
		0 1	x^2	0	1
0 0	$(1-q)^2$			0	

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 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.

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 (Boyd & Richerson, 1988; Levins, 1968). 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). **(p.391)**

Figure 14-1: 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.

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 that can "live well with themselves." Such 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 will consider what happens when what is optimal behavior varies with spatial location. We assume that the environment is subdivided into a large number of distinct patches, each with a unique optimal behavior; that is, for each patch, there is a different optimal behavior, and only one optimal behavior. Optimal behavior within each patch is forever the same. However, different **(p.392)** patches never have the same optimal behavior. Within each patch, a large subpopulation of organisms follows the life cycle: (1) birth, (2) learning, (3) behavior, (4) migration, (5) reproduction, (6) death. Individuals are born naïve and must use some strategy to acquire behavior. If that behavior is optimal for the local patch, then the individual's fitness is multiplied by the factor $1 + b \rangle 1$. Otherwise, fitness is unchanged. A proportion, m, of the local population emigrates to other patches, and an equal-sized group immigrates from other patches. Generations overlap only long enough for newly born naïve 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. Although fitness is assigned in natal patches, we assume that adults continue to display their behavior after migration, and so naïve individuals run the risk of learning from immigrants. Additionally, we assume that naïve individuals cannot tell who is and is not a native of their local patch. Even though such cues might be available in many circumstances, they are certainly not always available. We now determine the expected fitness of an organism using each of the four learning heuristics.

Individual Updating

The expected fitness of an individual updater is:

(3) $w(I) = w_0(1 + b)c,$

where $0 \langle c \langle 1$ is a multiplicative cost to survival or reproduction. Provided that $(1 + b)c \rangle 1$, individual updating will be the best-adapted heuristic, whenever the quality of social information in the local patch, q, is equal to zero. However, because 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**: *Although 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 that 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 be a stable mix of individual updating and unbiased social learning in the population.*

(p.393) We now consider when unbiased social learning (U) can outperform individual updating. In generation t, the expected fitness of an individual using unbiased social learning is:

(4) $w(U)_t = w_0(1 + q_t b),$

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 14-4, 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 each other under all circumstances, so models of this kind predict that both **(p.394)** will coexist, in the absence of other heuristics. A stable proportion of individual updaters $\hat{p}$ is found where:

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

$$\text{(7)} \quad \hat{p} = \frac{m((1+b)c - 1)}{(1-m)(1+b)(1-c)}.$$

Box 14-4: 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, it is defined by the recursion:

$$\text{(5)} \quad q_t = (1-m)(p_{t-1} + (1 - p_{t-1})q_{t-1}) + m(0),$$

where p_{t-1} is the proportion p_{t-1} of the local population comprising individual updaters, in the previous generation. To understand this recursion, first consider that a proportion 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 of q_{t-1} 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 a potential target 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 social 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:

$$\text{(6)} \quad \hat{q} = \frac{(1-m)p}{1 - (1-m)(1-p)}.$$

Numerical work 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: ($\partial\hat{p}/\partial m \rangle 0,, \partial\hat{p}/\partial b \rangle 0$ and $and \partial\hat{p}/\partial c \rangle 0,,$ ” $b \rangle 0,, c \in [0,, 1],, m \in [0,, 1],, (1+b)c \rangle 1)$.

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 that unbiased social learning can tolerate depends on 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 Learning

Précis: *Consensus learning yields higher fitness and replaces unbiased social learning, provided that mixing between patches is not so strong as to make the expected local proportion of optimal behavior fall below one-half. If mixing is sufficiently weak and individual updating sufficiently costly, then consensus 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 14-5, we prove that the condition for consensus learning to invade a population of individual updating and unbiased social learning is:

(8)
$$c > \frac{1 + b/2}{1 + b}.$$

(p.395) Box 14-5: Condition for Consensus 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:

(9) $w(C)_t = w_0[1 + b(q_t + q_t(1 - q_t)(2q_t - 1))],$

where the factor $q_t + q_t(1 - q_t)(2q_t - 1)$ was derived in Box 14-2. The invader faces a value of, 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:

(10) $w(C)_t > w_0(1 + q_t b) \Rightarrow q_t > 1/2.$

Consensus 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, this condition simplifies to $m \langle \frac{\hat{p}}{1 + \hat{p}}$. So as long as migration is not so strong as to flood local adaptive learning, which happens at a rate, consensus learning can invade a mix of individual updating and social learning. Because is a function of *m*, *b, c*, we can substitute in the expression for 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:

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

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 one might expect. Using the exact expression, consensus learning can exclude individual updating when $w_0(1 + b\hat{q}) \rangle w_0(1 + b)c$ and $c \rangle (1 + b/2)/(1 + b)$, which is satisfied when both and $1/2 \langle c \langle 3/4$.

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 will not 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 exclude even individual updating, which simple social learning can never do. We prove this also in Box 14-5. Provided migration **(p.396)** is weak enough and individual updating is expensive enough (but not too expensive), consensus learning can dominate the population entirely. There is an intermediate range of individual updating costs that allows consensus to dominate a population.

The exact result here depends critically on the precise model of consensus learning. However, the qualitative result is likely to be quite general. Consensus learning is a nonlinear 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 that consensus learning generates in each generation will balance the loss from immigration.

Payoff-Biased Learning

Précis: *Payoff-biased learning relies on 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 of payoff-biased learning. Payoff-biased learning can, like consensus learning and under the right conditions, replace both unbiased social learning and individual updating. If migration is weak enough and error in judging payoffs great enough, then consensus 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:

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

The above simplifies to $x < 1$ for all $q_t \in [0, 1]$, 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 nonlinear and actively changes the frequency of adaptive behavior from one generation to the next. Also like consensus learning, this means that payoff bias can sometimes exclude individual updating; in this case, provided that:

(13) $x < 1 - \dfrac{m}{1 - m} * \dfrac{b}{b - (1 - m)((1 + b)c - 1)}.$

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.

(p.397) Finally, consensus learning can sometimes invade and replace payoff-biased learning. Consensus can invade a pure population of payoff-biased learning and replace it, provided:

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

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 aforementioned results explored the properties of unbiased, consensus, and payoff-biased learning when the environment varies spatially. We have shown 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 that migration is not too common and individual updating is sufficiently costly.

We think that such conditions are actually quite rare (and as we show in the next sections, completely absent from our model of temporal environmental variation), but they do reveal a fundamental property of these nonlinear 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 principles that we reported 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. But because the locally optimal behavior does not change over time, a reliable store of locally adaptive culture can accumulate (as **(p.398)** long as mixing is not too strong). Indeed, we have shown 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, previously learned behavior may no longer be 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. Because there is a very large number of alternative behaviors, previously learned behavior may no longer be 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 a 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 over pure heuristics. In the case 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 is ever reduced to a very small number of individuals, then it will take a long time to recover. Therefore, its strategy has to both do well and avoid bottlenecks in order for the lineage to grow quickly and sustain its **(p.399)** numbers. When the environment varies temporally, selection favors heuristics that attend to both mean fitness and variance in fitness.

To apply this idea to our social learning heuristics, consider again the previously mentioned basic unbiased social learning model. We showed that there is a stable mix of individual updaters and social learners in this model, lying at:

(15) $$\hat{p} - \frac{m((1 + b)c - 1)}{(1 - m)(1 + b)(1 - c)},$$

where 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 14-6 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 is not 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 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:

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

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

(17) $$\log[\bar{r}] = \sum_{i=1}^{\infty} p_i \log w_i,$$

for n environmental states, where $\displaystyle\sum_{i=1}^{n} p_i = 1$.

(p.400) Box 14-6: 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:

(18) $\quad w(IU) = f w_0 [1 + b] c + (1 - f) w_0 (1 + \hat{q} b),$

where f is the fraction of the time that 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 $\partial w(IU)/\partial f|_{f=f^*} = 0$ for f^*. This yields:

(19) $\quad f^* = \dfrac{m((1+b)c - 1)}{(1-m)(1+b)(1-c)},$

which is the same expression as Equation 15 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.

Let us take the geometric logic above and apply it to our models of social learning. We will assume again that there is a large number of alternative behaviors, but instead of purely spatial variation, we will now assume that there is purely temporal variation. With 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 14-7, we demonstrate how to derive long-term growth rates under temporal environmental variation, in this case using the example of the purely unbiased social learning strategy, U.

Once we have a geometric fitness expression for a heuristic, we can analyze its evolutionary dynamics. Unfortunately, the form of this expression 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 14-8, we demonstrate how to construct weak-selection approximations for this model.

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 $r(U)$ in Box 14-7, the long-run growth of this mixed heuristic is:

(20)
$$r(\mathrm{IU}) = \sum_{t=0}^{\infty} u(1-u)^t \log \left[f w_0 (1+b)c + (1-f) w_0 (1+bq(t)) \right]$$

(p.401) Box 14-7: 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:

(21)
$$q(t) = \sum_{i=1}^{t} p(1-p)^{i-1} = 1 - (1-p)^t.$$

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 that the environment changed one generation ago ($t = 1$). In general, there is a chance 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:

(22)
$$r(\mathrm{U}) = u \log w_0] + \sum_{t=1}^{\infty} u(1-u)^t \log w_0 (1 + bq(t))].$$

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 that 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 . But a proportion $q(t)$ of social learners will get lucky and choose a target with optimal behavior, whereas the rest will not. Thus the probability $q(t)$ is inside the logarithm.

Using this expression, we prove in Box 14-9 that the mixed heuristic that randomly uses individual and social updating will invade and ultimately replace a population of purely

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 purely 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 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 outperforms 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.

(p.402) Box 14-8: Weak Selection Approximation

Weak selection applies when b and $(1 - c)$ are small such that terms of orderand 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 will only be exactly valid for choices that have modest effects on fitness. However, numerical work, as well as the simulations we show in a later section, confirms that the qualitative conclusions we reach here are general to strong selection. To apply the weak selection approximation, we use a Taylor series expansion of $r(U)$ around $b = 0$, $c = 1$, and keep the linear terms in b, c. This gives us:

$$(23) \quad r(U) \approx \log[w_0] + \frac{(1-u)bp}{1-(1-p)(1-u)}.$$

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

$$(24) \quad \bar{p} = \frac{u(b-(1-c))}{(1-u)(1-c)}.$$

This is the same as Equation 15, 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.

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 on 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 autocorrelated—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). Although we see no reason why the conclusions here should not extend, qualitatively, to autocorrelated environments, we do believe that it is a problem worth modeling.

(p.403) Box 14-9: 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:

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

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

$r(IU)|_{p=\bar{p}} > r(I)$ where is defined by Equation 24 in Box 14-8. 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. Because both I and U have the same fitness at $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 $f = f^*$ that no other value of f can invade. When IU is common, the proportion of optimal behavior is now given by:

(26) $$q(t) = \sum^t f^*(1-f^*)^{-1},$$

where f^* is the common chance a IU individual updates individually. The evolutionarily stable value of f^* is found where $\partial r(IU)/\partial f|_{f=f^*} = 0$. Again using a weak selection approximation and solving the above for f^* yields:

$$(27) \quad f^* = \frac{u((1+b)c - 1 + u((1-b)(1-c_i - b))}{(1-c_i^i(1-u)^2}.$$

By substituting the value of f^* back into r(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 $r(IU)|_{f=p=f^*} \rangle r(I)$ and when $r(IU)|_{f=p=f^*} \rangle r(U)|_{p=f^*}$. Both of these conditions are true for all $b \rangle 0$, $0 \langle c \langle 1$, $0 \langle u \langle 1$, so the mixed heuristic can both invade a population of pure heuristics as well as resist invasion by either pure heuristic.

(p.404) Consensus Learning Less Favored Under Temporal Variation

Précis: *When the environment varies through time, consensus learning does much worse than it does under spatial variation. It can never exclude individual updating, because every time the environment changes, $q \langle 1/2$ 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 have already demonstrated that consensus learning can in fact exclude both simple social learning and individual updating, provided that the rate of mixing between locales is sufficiently low. Under purely temporal variation, consensus learning can never exclude the other social 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 nonadaptive behavior. As a result, consensus learning depends on 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 nonlinearities of consensus learning is analytically difficult. But the results are easy to visualize in simulation from the fitness definitions. (The short simulation code can be obtained from the first author.) Figure 14-2 plots the proportions of consensus learning, unbiased social learning, and individual updating through time, for both purely spatial and purely temporal environmental variation. In the absence of temporal variation in optimal behavior, consensus learning can actually exclude both individual updating and unbiased social learning (Panel A). However, under purely 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 (Panel B).

A small amount of spatial variation and mixing can go a long way towards helping consensus learning, however (Figure 14-2, Panel C). Whereas 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 learning. After a change in the environment, consensus social learners suffer reduced fitness, declining in frequency as individual updating increases in frequency (see the time series in Panel C). 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, whereas unbiased **(p.405)**
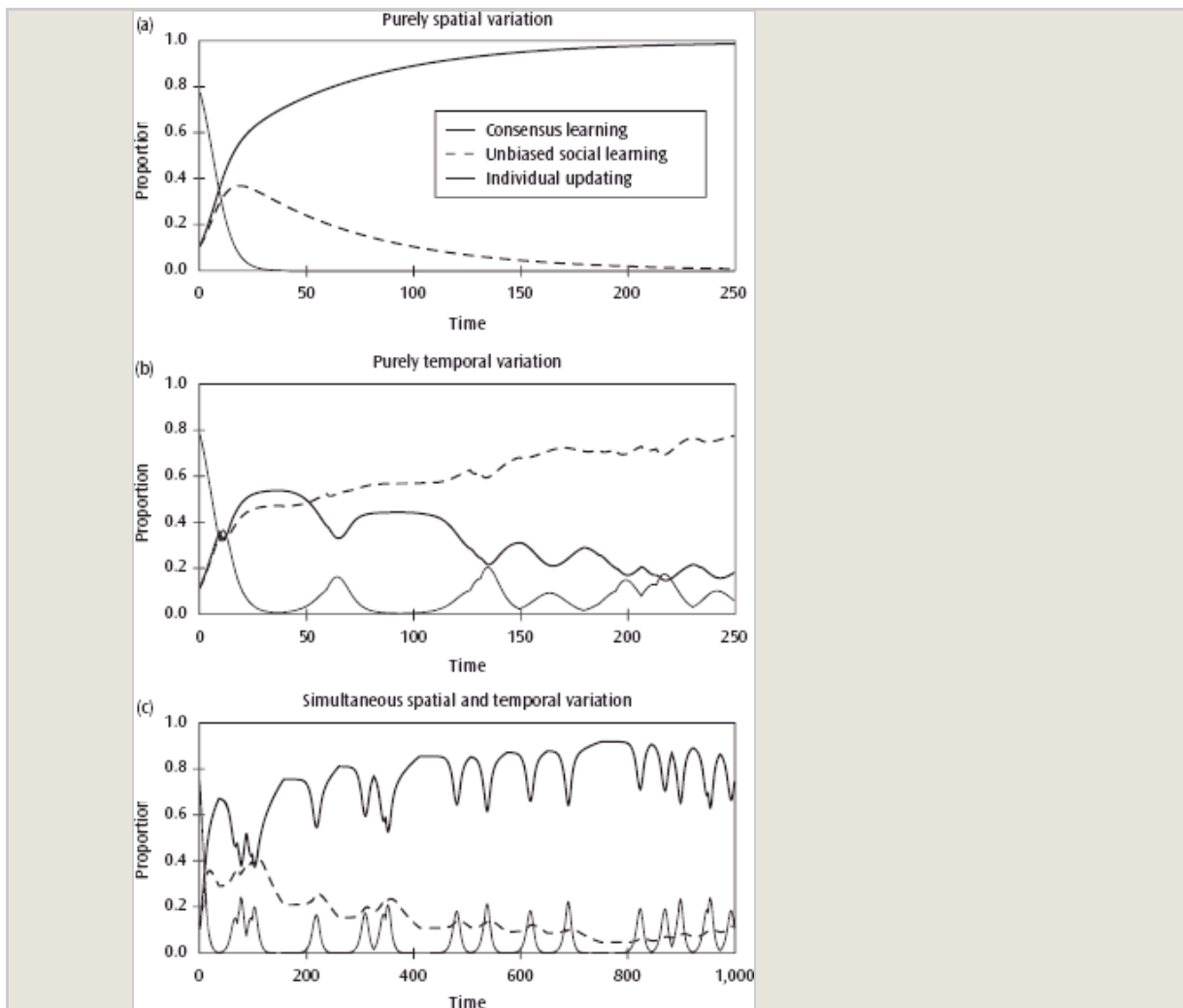

Figure 14-2: Stochastic simulations of the evolution of three learning strategies, under either spatial (Panel A), temporal (Panel B), or simultaneous spatial and temporal environmental variation (Panel C). In all three conditions, the initial proportions for consensus learning, unbiased social learning, and individual updating are 0.1; 0.1, and 0.8, respectively, and $b = 0.5$ and $c = 0.8$, respectively. Additionally, for purely spatial variation, $m = 0.05$ and $u = 0$; for purely temporal variation, $m = 0$ and $u = 0.05$; and for simultaneous spatial and temporal environmental variation, $u = m = 0.05$.

(p.406) 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, whereas mixing and spatial variation work for it. If either force is strong enough, it can eclipse the other.

The same principle applies in the case of consensus learning in temporally varying environments as 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 one should not 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, because consensus learning can drive the proportion of optimal behavior both downwards (when $q \langle 1/2$) as well as upwards (when $q \rangle 1/2$), the dynamics may be much more complex and interesting.

Summary

In this section, we have analyzed the effects of temporal environmental variation on unbiased and consensus 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 leads 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 purely individual updating and unbiased social learning strategies, when there is purely temporal variation in the environment.

Second, consensus 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 has again increased optimal **(p.407)** behavior in the population, consensus can do well, but lost fitness during the transition can cause it to be out-competed by other social learning heuristics. This does not happen under purely spatial variation, because a constant stream of immigrants actually provides an environmental challenge to which consensus learning is well suited, 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.

Conclusions

We have analyzed the long-term success of various social learning heuristics. Specifically, we have examined their ability to persist and to replace other heuristics, and we have done this in two broadly different kinds of environments: environments in which the optimal behavior varies across space, or through time. Because each social learning

heuristic also shapes its environment as individuals use it, our analysis has been at the same time ecological, game-theoretic, and evolutionary: The performance of each social learning heuristic depends on assumptions about the environment and population structure in which it is used. The use of a particular social learning heuristic will affect the success of this and other social learning heuristics over time. Therefore, our analysis is directed at the long-term survival and reproduction of each social learning heuristic.

According to our analysis, temporal and spatial variation 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 over either 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.

(p.408) 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 to 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 involved affects 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 and Aoki (2007) analyzed a model of the evolution of consensus learning in a temporally varying environment and found that they reached different conclusions from those of Henrich and Boyd (1998), who studied the evolution of consensus learning under simultaneous spatial and temporal variation. As we have shown, temporal variation selects against consensus learning in a way that spatial variation does not. Although Wakano and Aoki acknowledged the different assumptions about the nature of the environment, they decided without analysis that the divergent assumptions had no role in explaining their

divergent results. They instead speculated 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 and Richerson (2009) used simulations to demonstrate that, indeed, some kinds of temporal variation are worse for consensus learning than others, serving to reemphasize 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. Although 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 worthwhile to attend to the beliefs of others. The thought experiments therefore suggest one reason why people are so powerfully influenced by mere words.

Access brought to you by: Max-Planck Society