CHICAGO JOURNALS

When Natural Selection Favors Imitation of Parents

Richard McElreath and **Pontus Strimling**

Department of Anthropology and Graduate Groups in Ecology, Animal Behavior, and Population Biology, University of California, Davis, CA 95616, U.S.A. (mcelreath@ucdavis.edu)/Department of Mathematics and Physics, Mälardalen University, SE-721 23 Västerås, Sweden. 24 VII 07

It is commonly assumed that parents are important sources of socially learned behavior and beliefs. However, the empirical evidence that parents are cultural models is ambiguous, and debates continue over their importance. A formal theory that examines the evolution of psychological tendencies to imitate parents (vertical transmission) and to imitate non-parent adults (oblique transmission) in stochastic fluctuating environments points to forces that sometimes make vertical transmission adaptive, but oblique transmission recovers more quickly from rapid environmental change. These results suggest that neither mode of transmission should be expected to dominate the other across all domains. Vertical transmission may be preferred when (1) learned behavior affects fertility rather than survival to adulthood, (2) the relevant environment is stable, or (3) selection is strong. For those interested in the evolution of social learning in diverse taxa, these models provide predictions for use in comparative studies.

Given the tremendous attention paid to parents and parenting in popular culture, one might think that the science of parents' social influence had been worked out long ago. In contrast to the situation in the genetic arena, where the fact that every child has exactly two biological parents who contribute approximately equal amounts of hereditary material has led to powerful deductions about behavior and evolution, in the cultural arena surprisingly little is known about how much behavior and belief children acquire from their parents via social learning. The resolution of this issue is very important, because different patterns of social learning generate different patterns of adult behavior and even very different cultural evolutionary dynamics (Cavalli-Sforza and Feldman 1981; Boyd and Richerson 1985; Barth 1990). Debates ranging from the causes of crime to the nature of intelligence may hinge on an accurate understanding of how much and when chil-

dren preferentially learn from their parents as opposed to peers or nonparent adults.

Some anthropologists have claimed evidence of the importance of transmission of culture from parents to children (vertical transmission), at least in some domains (Cavalli-Sforza et al. 1982; Hewlett and Cavalli-Sforza 1986; Guglielmino et al. 1995; Hewlett, De Silvestri, and Guglielmino 2002). Each of these studies is valuable, but each has flaws that make it easy for skeptics to doubt its inferences. A classic study of correlations among college students, their parents, and their friends by Cavalli-Sforza et al. (1982) is sometimes cited as providing evidence of vertical transmission, but the observed correlations may arise from correlations between parents and unmeasured nonparents, as well as simply from the genetic heritability of traits. Also, parent-offspring correlations observed in young children may not persist when the children are older and have been exposed to many other cultural models. In one of the best empirical studies of the problem, Aunger (2000) studied the transmission of food taboos in the Ituri Forest. The analysis suggests that, while initial taboos are acquired from parents, later horizontal adult transmission has a huge effect on the resulting pattern of variation.

Aunger (2000) further argued that self-report of parental influence often reflects a normative reporting bias. A unique study by Hewlett and Cavalli-Sforza (1986) is also cited as presenting evidence of vertical transmission (although the paper itself argues for some complexity in the process). The key problem there is that the only evidence of learning from parents is self-report, and behavior varies so little in the community that it would probably be impossible to determine which learning pattern generated adult behavior. Even accurate self-report in this case would not mean that people would not eagerly imitate nonparents to acquire new techniques or technologies.

Population-level studies that claim evidence of learning from parents (Guglielmino et al. 1995; Hewlett et al. 2002) suffer from the limitation that some important domains, such as language, are known to be eagerly learned from nonparents and show strong in-group correlations through time (the "group effect" mentioned by Hewlett et al. 2002)—exactly the pattern used to infer transmission from parents. Furthermore, pure learning from parents does not lead to conservation of between-group variation unless there is no migration among groups. Exactly as in the genetic case, small amounts of migration are sufficient to unify subpopulations with respect to drift (Wright 1931). Therefore, simple parent-child social learning cannot explain the finding that some elements of culture vary mainly among groups rather than within them (Henrich and Boyd 1998).

While the above studies do not manage to distinguish between the competing hypotheses, they do show that observed patterns of variation are often *consistent* with vertical social learning—an important condition for investing in future re-

search. However, these patterns are also consistent with many forms of oblique and horizontal transmission. There is good reason to worry over merely consistent results. Studies that claim evidence of vertical cultural transmission usually do not control for the fact that parents and offspring share half of their genes. Once genetic sources of parent-child correlation in, for example, personality characteristics are accounted for, very little to no correlation remains to be explained (for a review of this literature, see Harris 1999). On the basis of such evidence, some developmental psychologists have made strong arguments that parents are not, in fact, a child's pre-ferred targets for social learning (Harris 1999). They argue instead that children are eager imitators of peers and pres-tigious adults. At minimum, these arguments should caution us against simply interpreting the existence of parent-child correlations as evidence of vertical cultural transmission.

In summary, while claims are sometimes made that vertical cultural transmission is strong (Shennan and Steele 1999; La-land, Odling-Smee, and Feldman 2000), we find the evidence ambiguous. In light of this empirical ambiguity, it would be helpful to have a family of models to aid in understanding the adaptive design of vertical social learning. When would natural selection, if it had its way, favor learning from parents instead of from nonparents? There has been little analysis of this problem. Cavalli-Sforza and Feldman (1981) examined the cultural dynamics that result from vertical transmission, but they did not model when natural selection would favor such a learning strategy (see also Feldman and Zhivotovsky 1992). Most formal models of the evolution of social learning have assumed either oblique or vertical transmission but not both (e.g., Rogers 1988; Boyd and Richerson 1995, 1996; Feld-man, Aoki, and Kumm 1996; Wakano, Aoki, and Feldman 2004; Aoki, Wakano, and Feldman 2005). These models there-fore do not help us understand when vertical social learning, relative to oblique, may be adaptive. Takahasi (1999) analyzed models in which individuals could learn vertically or obliquely/horizontally and found vertical transmission to be favored in many circumstances. However, none of these mod-els included environmental fluctuation. Such models may overestimate the importance of vertical social learning. Since models of the evolution of social learning suggest that en-vironmental fluctuation favors social learning (Boyd and Richerson 1985) and since vertical transmission is analogous to adding an extra chromosome, it seems likely that envi-ronmental fluctuation would select against vertical social learning just as it selects against direct genetic adaptation (Wakano et al. 2004).

This paper is the first to develop and analyze models of the simultaneous evolution of vertical and oblique social learning under stochastic environmental fluctuation. We con-struct two models. In both models, individuals can learn for themselves (individual learning), learn from a parent (vertical social learning), or learn from a nonparent adult (oblique social learning). The mechanism of learning is controlled by a single genetic locus, and learned behavior affects fitness in one of two ways. In the first model, learned behavior affects fertility. In the second model, learned behavior affects survival to adulthood. We show that, in model 1, stable environments and strong selection favor vertical social learning. The reason is that mere existence can serve as a cue to a naive individual that its parent did something right. However, when the en-vironment fluctuates more frequently, oblique social learning is advantaged, because it allows novel adaptive behavior to be acquired more quickly. In model 2, in which selection acts on survival to adulthood rather than fertility, selection does not favor vertical social learning unless that learning is in-herently more accurate.

These results suggest that neither mode of transmission should be expected to dominate the other across all domains. Those who believe that vertical or oblique transmission does dominate will need both better evidence and additional mod-els to demonstrate that dominance. At the end of this paper, we suggest additional relevant theory and discuss predictions derived from our models.

The Models

We constructed two models that differ only in the way se-lection operates. In the first, selection operates on fertility by increasing or decreasing the number of expected offspring. In the second, selection acts on viability by leaving family size unchanged but altering the probability that a child survives to reproduce. Population biologists often distinguish between these two modes of selection because they may favor different adaptations. Most elements of the two models are identical, and therefore we explain the common structure before in-troducing the specifics of each.

Basic Model Structure

We begin with a model structure similar to one used to explore the evolution of oblique transmission (Rogers 1988). Then we add an allele that leads an individual to learn from its parents. In an environment that can have any of an infinite number of states, at any one time a single behavior is optimal for the current state of the environment. All other behavior leaves fitness unchanged from a baseline w_0. In each gener-ation, after offspring are born but before they have a chance to acquire behavior the state of the environment can change, rendering all previous behavior nonoptimal. The state of the environment in any generation t is given by u_t. If the envi-ronment has just changed to a new state, then $u_t = 1$. This happens with probability u. Otherwise, $u_t = 0$. All individuals in the population experience the same environment and can-not detect when a change occurs.

The population is very large, and generations are discrete and nearly nonoverlapping. Young individuals can observe and learn from adults, but no individual survives for two generations. An individual's learning strategy is determined by a single allele, and inheritance is haploid and sexual, under

random mating. Individual learners (I) ignore the previous generation and explore the environment themselves. They always acquire currently optimal behavior but also pay a cost of learning. Oblique social learners (S) acquire the behavior of a random member of the previous generation. Vertical social learners (V) copy one of their two parents with equal probability. Thus, we assume that social learners who acquire adaptive behavior have the highest fitness, A, followed by individual learners, who have fitness B. Social learners who imitate someone with maladaptive behavior have the lowest fitness, C. Mating is random, and the reproduction of an individual is the product of its own fitness (A, B, or C) and that of its mate. We analyzed an equivalent additive model and derived all the same conclusions. However, since fitness effects must be multiplicative in model 2, making them multiplicative in both models makes it easier to compare the results.

To model the possibility that imitating parents is easier and more accurate than imitating nonparents, we reduce the odds of successful oblique social learning by a factor $1 - k$. This assumption arises from the intuition that parents may have stronger motivation than nonparents to teach adaptive behavior and be more available to imitate. When $k = 0$, oblique and vertical social learning are equally accurate. When $k = 0.1$, imitating a nonparent is, on average, 10% less accurate than imitating a parent.

Under the above assumptions, we can construct a mating table that specifies the probabilities of occurrence of each type of family and the expected number of offspring of each family type (table 1). From this table, it is possible to write recursions for the frequencies of five types of individuals in the population after mating and learning: individual learners with optimal behavior (i'), oblique learners with and without optimal behavior (s_1' and s_0', respectively), and vertical learners with and without optimal behavior (v_1' and v_0', respectively). These recursions are built from table 1 by multiplying the probability of each mating (col. 3) by the fertility of that mating (col. 4) and the proportion of offspring of a given type from that mating (cols. 5–9). Each row yields a product of three columns of this kind, and the products are added to produce a total number of offspring of a given type in the next generation. This expression is finally divided by the average fitness (the sum products of cols. 3 and 4) to produce a complete recursion for the frequency of the type.

In model 1, we assume that average family size is the product of the fertilities of the two parents (see table 1). Baseline individual fertility is w_0. We let adaptive behavior increase an individual's fertility by a factor $b > 1$. Individual learners have their fertility reduced by a factor $c < 1$, such that their fitness is multiplied by $bc > 1$. Thus, in this model, $A = w_0 b$, $B = w_0 bc$, and $C = w_0$.

In model 2, we assume that average family size is a constant, w_0, by setting all individual fertilities to this value (see table 1). We let adaptive behavior increase an individual's chances of surviving to adulthood by a factor $b > 1$. Individual learners

have their survival chances reduced by a factor $c < 1$, such that their fitness is multiplied by $bc > 1$. Thus, in this model, $A = B = C = w_0$. To construct recursions that account for differences in survival to adulthood, we modify the recursions that arise from table 1. Each recursion is multiplied by the viability factors of that type:

$$i'' = i' \frac{hbc}{\bar{w}'},$$

$$s_1'' = s_1' \frac{hb}{\bar{w}'},$$

$$s_0'' = s_0' \frac{h}{\bar{w}'},$$

$$v_1'' = v_1' \frac{hb}{\bar{w}'},$$

$$v_0'' = 1 - i'' - s_1'' - s_0'' - v_1'',$$

where h is the baseline chance of survival to adulthood and $\bar{w}' = i'hbc + (s_1' + v_1')hb + (s_0' + v_0')h$ is the average fitness after survival is accounted for. The factor h divides out in each case and never affects differential fitness.

Stationary Distributions and Stability Conditions

The recursions for each model are difficult to analyze, because this is a stochastic system and therefore the state variables will never reach equilibrium. However, it is possible to take the expectation over environments of each recursion and solve for steady-state values of the expected state variables because the system does eventually reach a steady state at which each state value (i, s_1, s_0, and v_1) is drawn from a stationary distribution (fig. 1). If we know the expected value of this distribution, we can still understand a great deal about how the system behaves.

A more formal explanation of this process may be found in the CA+ online supplements. For each recursion, we derived what we call an "expected-value recursion" by taking the expectation over states of the environment (changed or unchanged). This yielded four new recursions in terms of u, not u_r. We then solved for possible stable values of these expected values to produce means of the stationary distribution of the state variables $E(\hat{i})$, $E(\hat{s}_1)$, $E(\hat{s}_0)$, $E(\hat{v}_1)$. With the steady-state frequencies of types in hand, we conducted linear stability analysis (see McElreath and Boyd 2007, chap. 7, for examples) at these expected values in order to derive conditions under which each characteristic steady state ("equilibrium" of the expected values) is stable (supplement A). We then verified the cogency of these deductions with stochastic simulations.

The equilibria and their stability conditions are summarized in table 2. There are four equilibria in each model. Only one is stable for any given combination of the parameters, meaning that initial conditions do not matter provided that

Table 1. Mating Table for Models 1 and 2

Parent 1	Parent 2	Frequency	Fertility	I	Probabilities of Offspring			
					S1	S0	V1	V0
I	I	i^2	B^2	1	0	0	0	0
I	S1	$2is_1$	BA	1/2	$(1/2)r'$	$(1/2)(1-r')$	0	0
I	S0	$2is_0$	BC	1/2	$(1/2)r'$	$(1/2)(1-r')$	0	0
I	V1	$2iv_1$	BA	1/2	0	0	$(1/2)r_v(2)$	$(1/2)[1-r_v(2)]$
I	V0	$2iv_0$	BC	1/2	0	0	$(1/2)r_v(1)$	$(1/2)[1-r_v(1)]$
S1	S1	s_1^2	A^2	0	r'	$1-r'$	0	0
S1	S0	$2s_1s_0$	AC	0	r'	$1-r'$	0	0
S1	V1	$2s_1v_1$	A^2	0	$(1/2)r'$	$(1/2)(1-r')$	$(1/2)r_v(2)$	$(1/2)[1-r_v(2)]$
S1	V0	$2s_1v_0$	AC	0	$(1/2)r'$	$(1/2)(1-r')$	$(1/2)r_v(1)$	$(1/2)[1-r_v(1)]$
S0	S0	s_0^2	C^2	0	r'	$1-r'$	0	0
S0	V1	$2s_0v_1$	CA	0	$(1/2)r'$	$(1/2)(1-r')$	$(1/2)r_v(1)$	$(1/2)[1-r_v(1)]$
S0	V0	$2s_0v_0$	C^2	0	$(1/2)r'$	$(1/2)(1-r')$	0	1/2
V1	V1	v_1^2	A^2	0	0	0	$r_v(2)$	$1-r_v(2)$
V1	V0	$2v_1v_0$	AC	0	0	0	$r_v(1)$	$1-r_v(1)$
V0	V0	v_0^2	C^2	0	0	0	0	1

Note: I, S, and V represent individual, oblique, and vertical learners, respectively. S1 and V1 are social learners who acquired adaptive behavior and S0 and V0 learners who acquired maladaptive behavior. The symbols i, s_1, s_0, v_1, and v_0 correspond to the frequencies of each type in the population. The frequency of currently adaptive behavior learned by individuals in the previous generation is given by $r' = (1-u_t)(1-k)(i+s_1+v_1)$. The function $r_v(x) = (1-u_t)(1/2)x$ gives the probability of acquiring adaptive behavior via vertical social learning, when x of an individual's parents acquired adaptive behavior in the previous generation.

all strategies are initially present at some frequency. At equilibrium 1, individual learning can be the only strategy. At equilibrium 2, both individual learning and oblique social learning can coexist, but vertical social learning is excluded. At equilibrium 3, both individual learning and vertical social learning can coexist, but oblique social learning is excluded. At equilibrium 4, vertical social learning alone may persist in the long run. The expected values of the state variables at each equilibrium are summarized in table 3.

Individual learners must exist at some frequency in order for the population to track the environment. Nevertheless, equilibrium 4 is possible because these are the expected values of the state variables, not constant values. At this last equilibrium, rare individual learners are favored immediately after a change in the environment (when $u_t = 1$), but selection then purges them again. In the long run, the expected contribution to descendant generations of any given individual learning allele is zero, even though the presence of these alleles is important for the dynamics of the system and the continued dominance of vertical social learning.

The equilibrium dynamics for both models are easy to visualize by plotting each stability condition as a function of c. Then the regions of u and c that produce the different equilibrium outcomes can be plotted in a two-dimensional space for fixed values of b and k. In each case, individual learning is favored when the environment is unstable and learning is inexpensive, and vertical social learning is favored when the environment is stable. Oblique social learning and vertical social learning do not coexist at a deterministic steady state, although fluctuating selection may maintain both in the population. It is important to note that vertical social learning

never invades in model 2 if $k = 0$. Instead, it is neutrally stable with oblique social learning. However, fluctuating selection immediately after a change in the environment favors oblique social learning (as we explain in detail below), suppressing the frequency of vertical social learning and effectively allowing oblique learning to dominate in the long run. Even when $k > 0$, vertical social learning has a smaller domain of attraction in model 2 than in model 1.
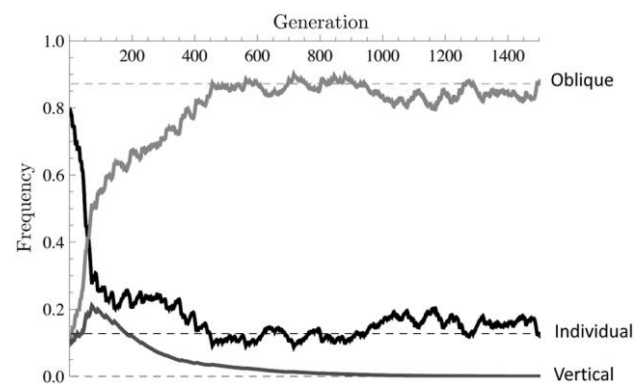

Figure 1. Forward simulation of the true stochastic dynamics of model 1; $b = 1.1$, $c = 0.93$, $k = 0$, $u = 0.3$. *Solid lines*, temporal dynamics of the frequencies of individual learning (*black*), oblique social learning (*light gray*), and vertical social learning (*dark gray*). *Dashed lines*, expected values of the stationary distributions of each. After an initial period of adjustment, the population forever fluctuates around these expected values, and we analyze formal expressions for these (table 3) to deduce the effects of the parameters on the behavior of the model.

Table 2. Stability Conditions for Models 1 and 2

| Equilibrium | Stability Conditions | |
	Model 1	Model 2
1	$u > 1 - \frac{bc-1}{(b-1)(1-k)}$	$u > 1 - \frac{bc-1}{(b-1)(1-k)}$
2	$u > 1 - \frac{2bc(1-k)-1}{(2b-1)(1-k)}$	$u > 1 - c\left(2 - \frac{1}{1-k}\right)$
3	$u > 1 - c$	$u > 1 - c, k > 0$
4	$u < 1 - c$	$u < 1 - c, k > 0$

Note: Derived from linear stability analysis of expected-value recursions.

Because the actual dynamics of the system are stochastic, we also conducted simulations with discrete stochastic environmental fluctuations instead of the expected values used in the formal analysis above (supplement B). This allowed us to verify the cogency of the analysis. When we plotted the average frequency of each type of learning—individual, oblique social, and vertical social—over the last 1,000 generations of 4,000-generation-long simulations, with initial frequencies set equal across all five types, the qualitative outcomes were the same as the analytical solutions. Therefore, while we have been unable to work with exact analytical descriptions of the stochastic dynamics (this is not surprising; see Gillespie 1994, chap. 4), stochastic simulations suggest that the qualitative insights of the expected-value analysis are accurate. However, the analytic results tend to overestimate the range of parameter values that lead to the invasion of vertical social learning, especially when selection is strong ($b \gg 1$).

Discussion: When Does Vertical Social Learning Evolve?

Our analyses show that vertical social learning evolves in more stable environments, lending support to earlier suggestions that this may be true (Laland, Richerson, and Boyd 1996; Henrich and Boyd 1998; Hewlett and Lamb 2002). To understand why decreasing amounts of environmental change lead from individual learning to oblique and then to vertical social learning, it is helpful to understand what effect favors vertical social learning at all. Two antagonistic forces affect whether oblique or vertical social learning is adaptive. One is that vertical social learning benefits from a simple heuristic arising from selection on fertility. The second is that vertical social learning does not recover as quickly as oblique social learning from environmental change.

What Favors Vertical Transmission?

The main benefit of vertical social learning arises from the fact that the existence of a child indicates, on average, that its parents behaved correctly. Figure 2 illustrates how this can result in a greater probability of acquiring adaptive behavior through imitating a parent rather than a nonparent adult,

provided that individual learners are not too common and that the general amount of adaptive behavior in the population is not too low (supplement C). This advantage of vertical social learning arises because, at least in model 1, parents with adaptive behavior produce more offspring. All other things being equal, this favors adopting the parents' strategy. On average, families in which there is the greatest chance of acquiring adaptive behavior through imitating parents will have more children. Over multiple generations, this effect multiplies itself as the lineage increases in size. Thus, the thought experiment in figure 2 underrepresents the magnitude of the advantage. When selection is strong (i.e., b is large), this effect is of course stronger, and the domain of attraction for vertical social learning grows. No similar effect occurs in model 2, where family size does not depend on learned behavior.
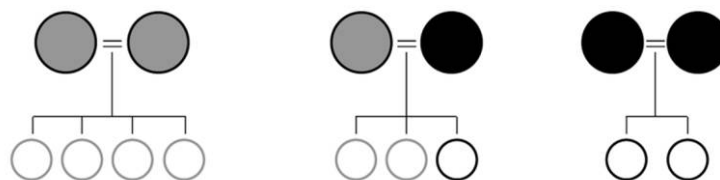
What Favors Oblique Transmission?

Whether alleles leading to vertical social learning spread in the population in the long run will also depend, however, on the rate of environmental change. One way to see vertical social learning is as a separate, nongenetic chromosome. Behaviors acquired in this way will recombine with alleles at the same rate as alleles on different chromosomes— 50%, which is substantial. However, oblique social learning has an even larger recombination rate (fig. 3). Through imitating nonparents in a large population, naive individuals essentially immediately reach equilibrium covariance between their alleles and learned behavior. The main advantage of this is that it affords them better access to individual learners, who are generating new adaptive information. Vertical social learners will eventually arrive at the same or a similar chance of acquiring adaptive behavior, but that will take a number of generations. Because an individual is a vertical social learner, at least one of its parents was also a vertical social learner. Imitating a parent then entails the likelihood of imitating an

Table 3. Steady-State Expected Frequencies of Individual Learners and Adaptive Behavior among Social Learners for Each Equilibrium

| Equilibrium | Model 1 | | Model 2 | |
	$E(\hat{i})$	$\hat{\alpha}$	$E(\hat{i})$	$\hat{\alpha}$
1	1		1	
2	$\frac{(bc-1)[u-k(1-u)]}{b(1-c)(1-u)(1-k)}$	$\frac{bc-1}{b-1}$	$\frac{(bc-1)[c-(1-k)(1-u)]}{(1-c)(1-u)(1-k)}$	$\frac{bc-1}{c(b-1)}$
3	$\frac{2(bc-1)[u-(1-c)]}{(1-c)(1-u)}$	$\frac{bc-1}{b-1}$	$\frac{2(bc-1)[u-(1-c)]}{(1-c)(1-u)}$	$\frac{bc-1}{c(b-1)}$
4	0	$\frac{b(1-u)-1}{b-1}$	0	$\frac{b(1-u)-1}{(b-1)(1-u)}$

Note: Since there are at most two learning types expected at steady state, the frequency of the other learning allele is always $1 - E(\hat{i})$. Where $\hat{\alpha}$ is the frequency of adaptive behavior among social learners, $\hat{\alpha} = E(\hat{s}_1)/[1 - E(\hat{i})]$ for oblique social learning and $\hat{\alpha} = E(\hat{v}_1)/[1 - E(\hat{i})]$ for vertical social learning.

Vertical social learning:

$$\frac{\bigcirc\bigcirc\bigcirc\bigcirc}{\bigcirc\bigcirc\bigcirc\bigcirc+\bigcirc\bigcirc\bigcirc\mathbf{\bigcirc}+\mathbf{\bigcirc\bigcirc}}(1) + \frac{\bigcirc\bigcirc\mathbf{\bigcirc}}{\bigcirc\bigcirc\bigcirc\bigcirc+\bigcirc\bigcirc\bigcirc\mathbf{\bigcirc}+\mathbf{\bigcirc\bigcirc}}(1/2)$$

$$+ \frac{\mathbf{\bigcirc\bigcirc}}{\bigcirc\bigcirc\bigcirc\bigcirc+\bigcirc\bigcirc\bigcirc\mathbf{\bigcirc}+\mathbf{\bigcirc\bigcirc}}(0) = (4 + 3/2)/9 = 0.611$$

Oblique social learning:

$$\frac{\bigcirc\bigcirc\bigcirc\bigcirc}{\bigcirc\bigcirc\bigcirc\bigcirc+\bigcirc\bigcirc\bigcirc\mathbf{\bigcirc}+\mathbf{\bigcirc\bigcirc}}(1/4) + \frac{\bigcirc\bigcirc\mathbf{\bigcirc}}{\bigcirc\bigcirc\bigcirc\bigcirc+\bigcirc\bigcirc\bigcirc\mathbf{\bigcirc}+\mathbf{\bigcirc\bigcirc}}(2/4)$$

$$+ \frac{\mathbf{\bigcirc\bigcirc}}{\bigcirc\bigcirc\bigcirc\bigcirc+\bigcirc\bigcirc\bigcirc\mathbf{\bigcirc}+\mathbf{\bigcirc\bigcirc}}(3/4) = 4/9 = 0.444$$
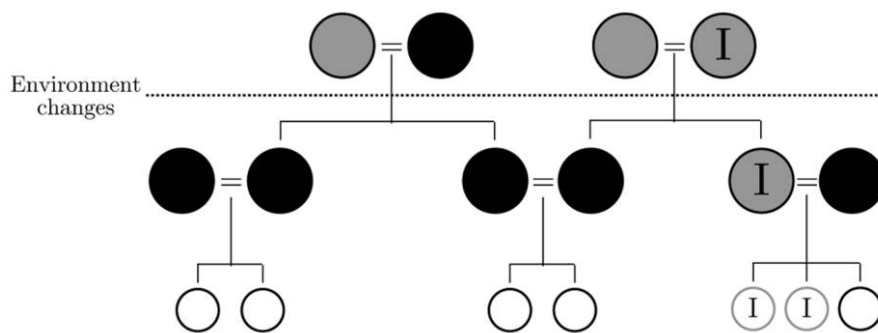
Figure 2. How imitating parents can increase the odds of acquiring adaptive behavior. Filled circles at top represent three mated pairs: *gray*, adaptive behavior; *black*, maladaptive behavior. Pairs with more adaptive behavior produce more offspring. The chance of acquiring adaptive behavior when imitating a parent (either, at random) is computed by first computing the chance that a given child comes from a type of pairing and then multiplying each of those probabilities by the odds of acquiring adaptive behavior, which is the proportion of parents who have adaptive behavior, in that pair. The chance of acquiring adaptive behavior via oblique transmission is instead computed by using the proportion of nonparents who have the adaptive behavior. (In a large population, this will simply be the population average frequency.) Because pairs with more adaptive behavior produce more offspring, it makes sense to imitate parents. The fact that one is a child essentially becomes a cue that, on average, one's parents behaved correctly. This example is informative only when the environment does not change too rapidly and individual learners are not too common.

imitator, which is not good unless most behavior is already adaptive. This difference in recombination rate means that oblique learners will tend to acquire adaptive behavior more quickly after a change in the environment unless selection is very strong ($b > 2c(1 - k)$) (supplements C, D). Eventually, as new adaptive behavior filters through the population, vertical social learning catches up with and may even surpass oblique social learning, but if the environment changes again first, vertical social learners lose, on average.

These factors explain why, when $k = 0$ in model 2 (selection on viability), selection never favors vertical social learning. Vertical social learning has no value in that case because all families are the same size. Thus, a major advantage to vertical social learning that arises in model 1, in which se-

lection is on fertility, never appears here. If all families are the same size, the fact that one exists contains no information about the adaptedness of one's parents' behavior. Oblique transmission's advantage when the environment changes still exists in this model. Thus, provided that $k = 0$, selection always favors oblique social learning if it favors any social learning at all. If $k > 0$, however, vertical social learning can invade just because it is more accurate than oblique social learning. Oblique social learning is advantaged when the environment changes, and if it changes often enough, oblique social learning can displace vertical.

To summarize, these models offer us three lessons about the evolutionary logic of these learning strategies. First, vertical social learning is very similar to innate genetic adapta-

Vertical social learning:

$$\frac{\bigcirc}{\bigcirc + \bigcirc\bigcirc + \bigcirc\bigcirc}(1/2) + \frac{\bigcirc\bigcirc + \bigcirc\bigcirc}{\bigcirc + \bigcirc\bigcirc + \bigcirc\bigcirc}(0) = 1/10$$

Oblique social learning:

$$\frac{\bigcirc}{\bigcirc + \bigcirc\bigcirc + \bigcirc\bigcirc}(0) + \frac{\bigcirc\bigcirc + \bigcirc\bigcirc}{\bigcirc + \bigcirc\bigcirc + \bigcirc\bigcirc}(1/4) = 2/10$$

Figure 3. How imitating nonparents can increase the odds of acquiring adaptive behavior. Symbols are as in figure 2, but we also indicate individual learners with *I*. Just after a change in the environment—indicated by the dotted line—all social learners acquire maladaptive behavior. A single individual learner, however, acquires newly adaptive behavior. The offspring of these individuals now face the decision (or natural selection does) of how they should acquire behavior. Two of these offspring are individual learners, like one of their parents. The rest decide (or natural selection does) between vertical and oblique social learning. Vertical social learners will acquire adaptive behavior 1/10 of the time. Oblique social learners will acquire adaptive behavior 2/10 of the time. Oblique transmission has the advantage in this example because the recent change in the environment has made adaptive behavior rare. Individual learners are the best source of adaptive behavior now, and, since social learning is heritable (being a social learner means that at least one parent is one), social learners will have better access to individual learners by learning from nonparents. This increased "recombination" rate between genes and cultural variants favors oblique social learning. However, if the environment does not change frequently enough, this advantage will be ephemeral.

tion: it tends to do well when behavioral flexibility is of little value. It is distinct from genetic adaptation, however, in that it does not rely on mutation to generate new variants. Even very rare individual learners, as in equilibrium 4, will produce "mutations" that are "biased" in an adaptive way that genetic mutations are not. Being able to acquire these positively biased new variants means that vertical social learning may still be able to adapt more quickly than innate genetic adaptation. Second, an additional advantage of imitating parents may be that they are easier to imitate, and this advantage can operate even when selection does not favor vertical social learning for any structural reason. Third, in contrast to vertical social learning, oblique social learning, combined with individual

learning, is a robust solution to even quite large temporal environmental stochasticity. When the environment changes rapidly, behavioral heritability is the enemy. Oblique social learning, by increasing the rate of recombination between genes and learned behavior, reduces this heritability and allows more rapid adaptation.

Limits of Our Models

We have necessarily made a number of simplifying assumptions in these analyses, and not just to make the model more tractable. Simple models are designed to help us understand the joint action of a specific number of forces. The assump-

tions are strategic, just as controlling for confounding variables in experiments is strategic. Nevertheless, we could modify the model in a number of ways. We conclude by discussing two modifications that represent other hypotheses about what factors would change the balance of value between vertical and oblique cultural transmission.

All of the social learning in the two models described above is "linear": learners accurately replicate the frequency of adaptive behavior within the category of people they learn from. There are other social learning strategies, however, that are nonlinear, and nonlinear strategies have their advantages. Parents are a very limited sample of the population. Models of success-biased imitation and majority-rule conformity suggest that both are adaptive learning heuristics in the proper contexts (Boyd and Richerson 1985; Henrich and Boyd 1998). There is also empirical evidence that humans employ both (Henrich and Gil-White 2001; McElreath et al. 2005). Using either depends on learning from nonparents.

The models here could be modified to allow oblique learners to compare the apparent success of candidate adults and preferentially imitate the adult with the highest observed fertility (in model 1) or health (in model 2). For majority-rule conformity, individuals could sample n adults (possibly including parents) and preferentially copy the most common behavior among them. Adding these strategies to the models would likely increase the range of parameters under which oblique social learning of some kind evolves. Parents may be part of one's sample, but as neighborhoods increase in size, parents will have smaller and smaller effects on final learned behavior.

We qualify the above speculation by noting that oblique social learning may encourage the evolution of "parasitic" beliefs that are good at getting themselves copied but harmful to their carriers. For example, Richerson and Boyd (2005) suggest that some modern vocations, such as being a scientist, compromise reproduction but expose followers to wider audiences, thereby recruiting more members despite fitness losses. In a basic sense, any belief that manages to channel some of the believer's energy into spreading the belief rather than pursuing fitness gains might be parasitic. Vertical social learning might be a good idea because it is partial inoculation against the evolution of such parasitic beliefs, just as vertical transmission of infection may reduce parasite virulence (Anderson and May 1982; Ewald 1987). One could incorporate this idea into the models in this paper by allowing both adaptive and maladaptive behavior to mutate so that both reduce a carrier's fitness by a factor $p < 1$ (for "parasite") while increasing the chance that its carrier will become a cultural model. One would then need to track four types (adaptive-normal, maladaptive-normal, adaptive-parasite, and maladaptive-parasite) for each learning strategy. This would greatly multiply the dimensionality of the system. Chances are that it would be impossible to derive the kind of analytical results we have presented here, but numerical analysis would still be possible and might be quite revealing.

Is Vertical Culture More Adaptive?

An important limitation of our models is common to many formal models of this kind. Socially learned behavior in our models is parasitic: a population consisting of oblique-social and individual learners, sustaining substantial culturally transmitted behavior, has the same average fitness as a population consisting entirely of individual learners (Rogers 1988). This is a dissatisfying feature of the literature, because presumably something about human culture raises mean fitness (Boyd and Richerson 1995). One reviewer asked us to inquire of our models whether vertical social learning changed this result. Indeed it does, but only in the extreme case. From table 3, we can calculate the mean fitness at steady state in each equilibrium. In model 1, for the first three equilibria, the average fitness is $w_0^2 b^2 c^2$. This replicates Rogers's result that oblique social learning does not increase adaptedness. For the fourth equilibrium, however, average fitness is given by $w_0^2 b^2 (1 - u)^2$. This is greater than $w_0^2 b^2 c^2$ whenever $u < 1 - c$, which is the condition for this equilibrium to be stable. Likewise, in model 2, mean fitness after viability selection is the same in equilibria 1–3 but higher in equilibrium 4 whenever $b(1 - u) > bc$, which is true whenever the equilibrium is stable ($k > 0$ is also required, in this case). Thus, when vertical social learning excludes the other learning strategies, mean fitness does increase. Individual learners do fluctuate in and out of the population at the fourth equilibrium. Thus, the population gets the benefit of their innovations immediately after rare changes in the environment but does not pay much cost of sustaining them in the long run.

This result echoes a previously modeled way for social learning to increase mean fitness (Boyd and Richerson 1995). If individual learning is activated more when cheap or effective, then social learning can increase adaptedness. Therefore it is not necessarily anything special about vertical social learning that allows it to increase mean fitness while oblique cannot. Instead, any pattern that falls back on social learning when individual learning is costly or unnecessary would produce the same result. For example, if individuals could guess to some degree when the environment had changed and learn individually, they would replicate the qualitative pattern at this fourth equilibrium. The simplest extension of the models in this paper that would address the adaptedness question directly would be to incorporate a chance of detecting nonadaptive behavior and activating individual learning in such a case (as in Boyd and Richerson 1995, 1996; Enquist, Eriksson, and Ghirlanda 2007). Our guess is that oblique social learning would then show the same effect on mean fitness as vertical social learning.

Predictions

Models of this sort are often most useful in helping to verify the logic of arguments or to clarify the relationships among concepts. Sometimes we are lucky and manage to extract

qualitative predictions from them as well. In this case, our models imply that vertical social learning should be more common (1) in behavioral domains that affect fertility rather than survival, (2) when the behavior's effective environment is stable, and (3) when selection is strong.

The first prediction will be upheld if there are stronger parent-offspring correlations in direct child-care behavior than in adult self-directed welfare behavior. For example, we might predict that mothers will preferentially imitate their own mothers when learning to care for their own children while they will not show a preference for a parent when learning about foraging techniques.

The second prediction will be upheld if the young take note of rapid changes in the world and pay less attention to parents and other elders as a result (see Mead 1978). In more stable times, preferential imitation of parents will be more common. The relevant environmental instability may also differ across domains. For example, knowledge of how to navigate by the stars is likely to remain useful over long time periods. A rare volcanic eruption may make another navigation technique adaptive, but the higher recombination rate of oblique social learning will be less useful in such a domain unless new refinements of navigation by constellations are constantly cropping up. In contrast, knowledge of where to hunt and fish could become outdated very quickly as prey species move or are depleted. In this case, the higher recombination rate of oblique social learning could be highly adaptive.

Third, strong selection in model 1 increases the range of parameters that favor vertical social learning. Strong selection means greater variance in family size, and existence becomes a stronger cue that one's parents did something right. In model 2, that mechanism cannot operate, and b does not appear at all in the condition for vertical social learning to invade. We might predict that behaviors that are life-or-death matters for one's successful reproduction are more likely to be learned from parents than from nonparents.

All of these predictions could be translated to help in the study of diverse taxa. Many researchers now agree that social learning of some kind is taxonomically broad (Fragaszy and Perry 2003). Models like these should be able to make predictions about which species will show which kinds of social learning, as well as how social structure may influence nongenetic evolution. For example, a species that lives in mainly parent-offspring social groups should adapt to environmental change more slowly than one that lives in larger social groups with many nonkin. A species that evolved in a highly unpredictable environment should be more likely to attend to nonparents than one that evolved in a highly stable one.

Acknowledgments

References Cited

Anderson, R. M., and R. M. May. 1982. Coevolution of hosts and parasites. *Parasitology* 85:411–26.

Aoki, K., J. Y. Wakano, and M. W. Feldman. 2005. The emergence of social learning in a temporally changing environment: A theoretical model. *Current Anthropology* 46: 334–40.

Aunger, R. 2000. The life history of culture learning in a face-to-face society. *Ethos* 28:1–38.

Barth, F. 1990. The guru and the conjurer: Transactions in knowledge and the shaping of culture in Southeast Asia and Melanesia. *Man: Journal of the Royal Anthropological Institute*, n.s., 25:640–53.

Boyd, R., and P. J. Richerson. 1985. *Culture and the evolutionary process.* Chicago: University of Chicago Press.

———. 1995. Why does culture increase human adaptability? *Ethology and Sociobiology* 16:125–43.

———. 1996. Why culture is common, but cultural evolution is rare. *Proceedings of the British Academy* 88:77–93.

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

Cavalli-Sforza, L. L., M. W. Feldman, K. H. Chen, and S. M. Dornbusch. 1982. Theory and observation in cultural transmission. *Science* 218:19–27.

Enquist, M., K. Eriksson, and S. Ghirlanda. 2007. Critical social learning: A solution to Rogers' paradox of nonadaptive culture. *American Anthropologist* 109:727–34.

Ewald, P. W. 1987. Transmission modes and evolution of the parasitism-mutualism continuum. *Annals of the New York Academy of Sciences* 503:295–306.

Feldman, M. W., K. Aoki, and J. Kumm. 1996. Individual versus social learning: Evolutionary analysis in a fluctuating environment. *Anthropological Science* 104:209–32.

Feldman, M. W., and L. A. Zhivotovsky. 1992. Gene-culture co-evolution: Toward a general theory of vertical transmission. *Proceedings of the National Academy of Sciences, U.S.A.* 89:11935–38.

Fragaszy, D., and S. Perry, eds. 2003. *The biology of traditions: Models and evidence.* Cambridge: Cambridge University Press.

Gillespie, J. H. 1994. *The causes of molecular evolution.* Oxford: Oxford University Press.

Guglielmino, C. R., C. Viganotti, B. Hewlett, and L. Cavalli-Sforza. 1995. Cultural variation in Africa: Role of mechanisms of transmission and adaptation. *Proceedings of the National Academy of Sciences, U.S.A.* 92:7585–89.

Harris, J. R. 1999. *The nurture assumption: Why children turn out the way they do.* New York: Free Press.

Henrich, J., and R. Boyd. 1998. The evolution of conformist

transmission and between-group differences. *Evolution and Human Behavior* 19:215–42.

Henrich, J., and F. Gil-White. 2001. The evolution of prestige: Freely conferred status as a mechanism for enhancing the benefits of cultural transmission. *Evolution and Human Behavior* 22:1–32.

Hewlett, B., and L. Cavalli-Sforza. 1986. Cultural transmission among Aka Pygmies. *American Anthropologist* 88: 922–34.

Hewlett, B., and M. E. Lamb. 2002. Integrating evolution, culture and developmental psychology: Explaining caregiver-infant proximity and responsiveness in central Africa and the USA. In *Between culture and biology: Perspectives on ontogenetic development*, ed. H. Keller, Y. H. Poortinga, and A. Schölmerich, 241–69. Cambridge: Cambridge University Press.

Hewlett, B. S., A. De Silvestri, and C. R. Guglielmino. 2002. Semes and genes in Africa. *Current Anthropology* 43:313–21.

Laland, K. N., J. Odling-Smee, and M. W. Feldman. 2000. Niche construction, biological evolution, and cultural change. *Behavioral and Brain Sciences* 23:131–46.

Laland, K. N., P. J. Richerson, and R. Boyd. 1996. Developing a theory of animal social learning. In *Social learning in animals: The roots of culture*, ed. C. M. Heyes and B. G. Galef Jr., 129–54. San Diego: Academic Press.

McElreath, R., and R. Boyd. 2007. *Mathematical models of social evolution: A guide for the perplexed.* Chicago: University of Chicago Press.

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

Mead, M. 1978. *Culture and commitment: The new relationships between the generations in the 1970s.* New York: Columbia University Press.

Richerson, P. J., and R. Boyd. 2005. *Not by genes alone: How culture transformed human evolution.* Chicago: University of Chicago Press.

Rogers, A. R. 1988. Does biology constrain culture? *American Anthropologist* 90:819–31.

Shennan, S. J., and J. Steele. 1999. Cultural learning in hominids: A behavioural ecological approach. In *Mammalian social learning: Comparative and ecological perspectives*, ed. H. O. Box and K. R. Gibson, 367–88. Cambridge: Cambridge University Press.

Takahasi, K. 1999. Theoretical aspects of the mode of transmission in cultural inheritance. *Theoretical Population Biology* 55:208–25.

Wakano, J. T., K. Aoki, and M. W. Feldman. 2004. Evolution of social learning: A mathematical analysis. *Theoretical Population Biology* 66:249–58.

Wright, S. 1931. Evolution in Mendelian populations. *Genetics* 16:97–159.