

The coevolution of genes, innovation and culture in human evolution

Richard McElreath^{*,1}

¹Department of Anthropology, University of Utah, Salt Lake City 84103, USA

Final Draft
April 17, 2009

Contents

1	Introduction	2
2	The gene-culture framework	6
2.1	What does it mean to say that culture is an inheritance system?	8
2.2	Is extra-genetic “inheritance” common?	10
3	Simultaneous evolution of innovation and social learning	11
3.1	What is learning for?	13
3.2	Adding cumulative social learning	15
3.3	Joint dynamics of innovation and cumulative social learning .	19
3.4	How much cumulative culture?	22
3.5	When social learning enhances innovation	23
4	Where did culture come from?	25
4.1	Evolving cultural evolution	25
4.2	Evolvability as a side effect	27

*Corresponding author. mcelreath@ucdavis.edu

1 Introduction

Wheat is one of humanity’s great inventions. Coming in great variety, locally adapted to microclimates, it converts energy into a form people can use to make more people. It is also nearly wholly dependent upon people for its survival—like other grains, wheat seed does not easily break from the grass, but instead stays firm, stuck to a “tough” rachis, waiting for the farmer or machine to remove them all at once. A sensible wild grain instead disperses seed to the wind. Other domestic species are similar; they have partly out-sourced their reproduction to humans, in exchange for security. Domestic cattle—whatever you think of their intelligence—have done quite well, compared to their extinct wild ancestors.

All of the human species’ domesticated inventions represent a transfer of information. Information about the environment—how to extract energy from it, how to resist drought, how to make harvest easier for a human farmer—makes its way into wheat’s genes, during artificial selection. Generations of human farmers have shuttled information about the world, the world’s pests, and their own preferences into the plant’s genome, creating a technology that “knows” about the world we live in. Of course the farmer must have some extra information, in order to profit from wheat. But a tremendous amount of information must be contained in the plant’s genome, and it accumulated there over many generations.

When I say “information” here, I mean that if we knew how to interpret it, the sequence of bases in wheat DNA would tell us new things about the environment and how to adapt to it. However, there are more formal definitions of information that suggest that natural selection accomplishes a similar feat, building information about the environment into our own genome (Frank 2009). The fit the Darwin observed between organisms and their environments reflects this flow of information. Each organism embodies a recent step in a long chain of information commerce, between the environment and the population of genes. This commerce is not without friction, of course. Sexual reproduction and recombination interfere directly with selection’s ability to describe the environment. But over time, natural selection manages to adapt organisms to the environment, by differentially favoring alleles.

Of course real environments fluctuate and vary. The planet we live in now is quite different than that of Pleistocene, and spatial variation from pole to pole is at least as great as temporal change. As a result, some of the informa-

tion that organisms accumulate about the environment is meta-information, information about information in the environment. Whenever a seed assays available moisture and postpones germination as a result, the plant is employing this kind of accumulated meta-information. In less philosophical language, it learns.

Learning is a kind of phenotypic plasticity, a condition in which information in the genome teaches the organism to respond to information in the environment. Instead of natural selection building in a direct description of the environment, there is instead meta-information about variation in the environment. This meta-information might be an array of phenotypes that are triggered as the organism receives information from the environment, during its development. This is what the seed does, when it “decides” whether or not to germinate. The meta-information can also be an exploration strategy, however, so that instead of the organism’s genome containing information about, say, the location of a water source, the genome contains information about how to find a water source. This information is still relevant to only some environments, and is therefore information about these environments, because different exploration strategies—different ways of learning—are better or worse under different conditions.

These are two kinds of information about the environment: information built into the genome by generations of natural selection and information we acquire during our individual lifetimes, interpreted in light of information in the genome. But there is also an important third sort of information that some species make use of. Much of the knowledge that most farmers employ to manage wheat accumulated over many generations, but it isn’t contained in anyone’s genome, at least not in any simple sense. Instead, farmers inherit each generation the accumulated culture of farming. This information is fit to the environment, just as other human traditions can exhibit amazing adaptation. But no individual in the course of his or her lifetime could accumulate it. Instead, it has taken many generations to develop, in a way similar to how information over many generations accumulated in wheat’s genome. In the case of wheat, humans built the information directly into the plant. In the case of other elements of culture, humans built the information into human brains, and later books and other forms of storage that human brains can access. This information is also often meta-information, providing strategies for solving specific problems as well as strategies for learning in itself.

In this chapter, my aim to is to provide both an introduction to and

example of theory development within the evolution of this kind of accumulated cultural information and the genetic information that makes its accumulation possible. Unlike other apes, humans rely upon accumulated bodies of adaptive information—culture—that do not reside in the genome, but nevertheless do depend upon information in the genome for their continuity and pattern of evolution. In order to understand why humans are the only ape to have crossed this “gap” and become so committed to and dependent upon socially-transmitted complex adaptations, we must understand both the genetical origins of the psychology that makes cultural evolution possible—the information-about-information that resides in our genomes—and the behavioral origins of accumulated socially-transmitted information-about-information, culture.

How did cultural evolution evolve? Our goal is to ultimately understand why the human genome, contrasted with those of other apes, has learned to learn about the environment in a way that generates complex behavioral adaptations that rival the complexity of those produced directly by natural selection. Addressing this question brings up a number of puzzles. I focus on two closely related concepts.

Accumulated culture is a poor guide to the origins of accumulation.

First, complex cultural adaptations like boats and agriculture appear obviously worth learning. However, when cultural abilities were first evolving, these fancy accumulated bodies of information must not have yet existed. Explaining the origins of domesticated wheat cannot reference contemporary highly-domesticated varieties, but must instead reference wild varieties with their inconvenient wild characteristics. Similarly, we cannot explain the origin of human culture with reference to contemporary cultural adaptations. The information that has to be built into the genome, in order for a child or adult to acquire complex accumulated culture, would seem to pre-require the complex culture. How can culture get started, if there is little culture yet to acquire (Boyd and Richerson 1996)?

Culture makes humans evolvable, but that is not why culture evolved.

Second, while culture makes human societies highly evolvable—they can quickly (in genetic time) generate complex local adaptations to a large number of novel circumstances—this evolvability is not easy to understand as the original reason for the evolution of cultural learning. Quite sim-

ply, the first cultural learners would not have benefitted from the evolvability. The long-term population-level benefits of complex accumulated culture seem obvious. Once fancy human culture existed, it allowed us to adapt to every terrestrial environment (except Antarctica) and accumulate powerful bodies of information like quantum mechanics. The fitness of any contemporary human has been buoyed by many past generations of accumulated extraction of information from the environment and transmission and curation of this information. In this way cumulative culture generates important and powerful group advantages. But these group benefits are not easily understood as the reasons for humans' evolving cultural abilities. Instead, evolutionary ecologists would rather attempt to explain the origins of cultural capacities with individual benefits accruing to individual learners. Thus instead of marveling at the adaptedness of accumulated human culture, our task is to understand how psychological abilities driven by individual selective advantage can build group benefits as a by-product (Boyd and Richerson 1985).

The approach I take in this chapter is to first review the gene-culture or dual-inheritance approach to human evolution. I present relevant pre-existing theory on the problem of getting cumulative culture started. I use this theory then to introduce a new model that considers the simultaneous co-evolution of all three kinds of information transfer that I have outlined here: (1) the incorporation of meta-information—information about how to use information adaptively—into the genome through natural selection, (2) the adaptive use of direct environmental information through individual learning, and (3) the accumulation and transmission of environmental information across generations, outside of the genome. I will show that allowing for all three of these dynamics simultaneously illuminates one potential path across the cultural “gap,” whereby cumulative culture can get started, despite our first puzzle above. In the process, however, the evolutionary dynamics seem to cover their own tracks, hiding the initial changes that make the crossing possible. Finally, by allowing social learning to enhance innovation in ways other than merely allowing one to start where others left off, the model can produce levels of behavioral adaptation much greater than are possible if innovation and cumulative social learning are considered orthogonal psychological abilities.

2 The gene-culture framework

One way to view the development of the body of theory known as gene-culture or dual-inheritance theory (Boyd and Richerson 1985, Cavalli-Sforza and Feldman 1981, Durham 1991) is to think of it as a revolutionary idea that suggests that human evolution cannot be understood in the same purely-genetical way that works for other organisms. One could view it this way, but I do not.

In evolutionary ecology, there is a long and successful tradition of modeling organic evolution as an interaction between information contained in genes, the structure of the population, and the state of the environment. In this framework, the only heritable variation is contained in genes, and so evolution is described as change in this variation over time.

Or is it? In the simplest models, it is true that the only evolving aspects of the population are genes. But in slightly more complex models that consider population structure, gene frequencies cease to be the only evolving information. Now the distribution of genes, age and sex structure, and local population densities can all evolve and exert very strong influences on the future changes in gene frequencies. These systems cannot be reduced to gene-only descriptions—additional information about the population and environment is needed to understand and predict change. Applied mathematicians are keenly aware of this fact, because we must define “state variables” for each of these evolving bits of information.

Routinely, genes are not the only state variables, even in culture-free models. Gene-culture models therefore dwell well within this successful tradition in population biology. There is not necessarily anything unusual about theorizing and modeling the idea that extra-genetic information is required to adequately describe a population, even when our focus is genetical evolution.

If gene-culture models are special in any way, it is that the extra-genetic information directly influences individual phenotype in a way very similar to the action of genes. The basic issue is to identify the minimal requirements for representing evolution of phenotype in a species. For example, we could construct a very simple genetic model in which the change (Δ) in the frequency of an allele, p , is a function of environmental state, E . This system would have a single recursion:

$$\Delta p = F(p, E),$$

where the function $F(p, E)$ is to be specified depending upon what model

of adaptation to the environment we might choose. It might be that E has little effect on individuals with different alleles, or it might be that E favors one over the others. It might be that E is fluctuating, so that selection favors different alleles at different times. The change might depend upon p itself, as it does in the example of sickle-cell anemia and other cases of overdominance. But nowhere do we allow in such a system for E itself to evolve in response to p .

The scientific question is whether such models are sufficient to model the evolution of a given organism's phenotype. For example, some moths imprint on the plants they graze upon as larvae. When they are adults, they will seek out these same plants to lay their eggs. The information about which plant to seek is not coded in the genome, even though the strategy for imprinting is. In this case, if we only know genotypes and the state of the environment, we could not predict the behavior of organisms in the next time period. We also need to know the distribution of imprinted memories among moths. In such a case, we need at least one more equation:

$$\begin{aligned}\Delta p &= F(p, q, E), \\ \Delta q &= G(p, q, E),\end{aligned}$$

where q is the frequency of some learned variant (an imprinted plant, say), and $G(p, q, E)$ a function telling us how learning responds to environment, E , and its own previous state, q , and the frequency of an allele, p .

This all sounds rather complex. And it can be. However, when important parts of phenotype are acquired during development and depend upon previous phenotypes, some system like this is useful for understanding how the organism evolves. Unless we think existing behavior could be predicted solely from knowing the environment and the distribution of genes, at some point evolutionary models may incorporate the dynamics of behavioral inheritance. No heroic assumptions are required for behavioral inheritance to exist: if portions of phenotype depend upon the phenotypes of other individuals, then weak or strong inheritance of behavior can exist. In the long run, in a given model, it might turn out that behavioral dynamics have little effect on the outcome. In others, it will make a huge difference.

Cultural evolutionary models (as well as niche construction models, see Odling-Smee et al. 2003) can model just the non-genetic behavioral dynamics, as if q above did not depend upon p , as well as joint dynamics of a coupled gene-culture system (Boyd and Richerson 1985, Cavalli-Sforza and Feldman

1981, Durham 1991). In each case, however, the structure of the model is decided by the question of interest. There are no global models that encompass all questions about evolution. This is why we call them “models.” While a few axiomatic mathematical theories do exist in biology—the Price equation being the most famous—for the most part formal evolutionary models are attempts to understand the consequences of assumptions and to explore the sufficient or necessary conditions for a given outcome. With such models, we can study which kinds of strategies and population structures can possibly produce a phenomenon—these are possible sufficient conditions. We can also study which assumptions can be omitted, still producing a phenomenon—these are then not necessary conditions.

The demonstrable success of the modeling strategy across the sciences recommends it well. While most of the work in evolutionary anthropology and evolutionary psychology is concerned with more proximate phenomena than the population dynamics of joint gene-culture systems, understanding both the details of the psychological differences between humans and other apes and the different population dynamics of human and ape societies will be necessary, before we have a satisfactory set of answers for how humans evolved, both in relation to other primates and broader trends in animal societies.

2.1 What does it mean to say that culture is an inheritance system?

While evolutionary principles are equally applicable to almost any dynamical system, many researchers approach models of cultural transmission and evolution via an analogy with genetical evolution. Analogies are often useful, but can disguise important differences. This has rightly led some to be concerned about the strength of the gene-culture analogy (Sperber 2000). If cultural variants are not discrete, are prone to “mutation,” and are strongly affected by learning biases, then is it useful to speak of “transmission” of culture at all?

While I have no particular attachment to the term “transmission,” the answer is definitively “yes.” Even if all the above is true, culture can still be an evolving system that leads to cumulative adaptation. This does not mean that evolved psychology has no role to play in how culture evolves (on the contrary, psychology has a huge role to play in understanding culture), but

it does mean that dismissing cultural evolution on the basis of imperfection of the genetic analogy is unwarranted.

Many people—enthusiasts of the “meme” approach and critics alike—seem to have been persuaded by Richard Dawkins’ abstract statements on what is required for adaptive evolution to occur. In *The Extended Phenotype* (1982), he argued that any successfully replicating entity must exhibit (1) longevity, (2) fecundity, and (3) fidelity. The entity must last long enough (longevity) to make copies of itself (fecundity) that are reasonably similar to it (fidelity). Some have interpreted this to mean that anything with high mutation rates cannot be a successful replicator. Thus if cultural ideas change in the process of social learning, the conclusion is that they do not constitute an evolving system at all (see citations in Henrich and Boyd 2002). Similarly, if cultural variants are continuous and blended entities, then they never exactly replicate, and again cannot produce adaptive evolution.

These conclusions are unfounded. Read very generally, Dawkins’ conditions are necessary and sufficient—There must be some heritability for adaptive evolution to occur. However, there are many ways to produce heritable variation. So in the strict sense many people have read them, while Dawkins’ conditions are sufficient, they are definitely not necessary. Reverse-engineering DNA may tell us how inheritance *can* work, but it does not tell us how it *must* work. Henrich and Boyd (2002) examines the problems with this reverse-engineering in greater depth (see also Henrich et al. *ming*).

It is understandable that there is confusion about what is needed for adaptive evolution—even textbooks are confused. Before the union of genetics and Darwinism, most biologists, including Darwin, thought that inheritance was a blending process: offspring were a mix of parental phenotypes. Darwin was troubled by Fleeming Jenkin’s (1864) argument that natural selection could not produce adaptations, because inheritance would quickly deplete the variation natural selection relies upon. Fisher’s (1918) argument reconciling genetics with continuous phenotypic variation purportedly rescued Darwin. Many textbooks repeat this version of the history, reinforcing the notion that low-error discrete entities like genes is a necessary condition for adaptive evolution. But in reality both Jenkin’s argument and those who think Fisher saved Darwin are simply wrong: blending inheritance can preserve variation, and particulate inheritance is neither necessary nor sufficient to preserve variation. Maynard Smith (1998) has a chapter that examines this problem, still not entirely resolved within modern population genetics (see also Barton and Keightley 2002).

2.2 Is extra-genetic “inheritance” common?

I think this history lesson teaches us that our verbal models of even genetical evolution are sorely lacking. If so, then our informal appreciation of non-genetic influences on behavior are also suspect. In many baboons, females inherit dominance rank from their mothers and sisters (Silk and Boyd 1983). In these species, fitness is strongly affected by this extra-genetic inheritance: any female adopted at birth into a high-ranking matriline would be better off than if she were adopted into a low-ranking matriline. And this female will have her dominance rank before she fights a single member of her social group. Dominance is heritable, has important effects on fitness, and yet the mechanism of inheritance is at least partly non-genetic. The rules of how this inheritance works are complicated and very unlike genes. It probably depends upon the composition of one's own matriline, the composition of the entire social group, and local resource density and feeding competition. And yet no primatologist could completely understand baboon biology without taking this complicated extra-genetic pedigree into account. Its existence may lead females to strive for rank because of its downstream consequences, in addition to its immediate resource access effects (Boyd 1982, Leimar 1996).

Extra- or “epigenetic” (Maynard Smith 1990) systems like this are increasingly recognized: everywhere biologists look, they find hints of inheritance systems either built on top of genes or built from entirely different mechanisms. If the key question is what mechanisms account for heritable phenotypic differences among organisms, then the answer appears to be “many.” Jablonka and Lamb's *Evolution in Four Dimensions* (2005) mounts the empirically rich argument that heritable differences in many species are due to the action of several inheritance systems (genetic, epigenetic, behavioral and symbolic), sometimes interacting, sometimes acting in parallel.

If one thinks about cell division for a moment, it is obvious that processes other than the replication of DNA are needed to explain how it works. Organelles need to be copied (Sheahan et al. 2004), and the genetic code itself needs to be copied (and this is not contained in the DNA, nor could it be). Beyond cell division, adult phenotypes depend upon imprinting and other forms of learning that may channel the environments offspring are exposed to (a kind of niche construction—Odling-Smee et al. 2003). And finally, most biologists believe that DNA was certainly not the first form of hereditary biological material (Szathmáry and Maynard Smith 1995). Thus some inheritance systems must be able to sometimes create complementary and

even usurping inheritance systems.

In light of these plausible “inheritance systems,” it appears that human culture may not be so special or surprising at all, in the sense of being a non-genetic system of inheritance. Organisms as diverse as *Arabidopsis* (a small plant related to mustard that is a favorite of geneticists), common fruit flies and single-celled microscopic animals such as paramecia exhibit heritable differences due at least in part to mechanisms other than the sequence of nucleotides in their DNA. The existence of social learning as a system of inheritance and adaptation that functions in complement to DNA may turn out to be unremarkable.

There will always be aspects of human behavior and evolution that can be usefully modeled as culture-free. These are, after all, just models: all of them are wrong, but some are more useful than others. To someone who makes formal models of evolutionary systems, the question that we must answer is what qualitatively different phenomena we miss when we represent human (or any other organism’s) evolution with just state variables for its alleles. If we sometimes require state variables for early childhood experience, imprinting, or behaviors acquired via social learning, to make useful models of our own evolution, then attempts to construct culture-free models are simply scientifically inadequate. As with each of the possible systems above (e.g. Jablonka and Lamb 1991, Pál and Miklós 1999, Maynard Smith 1990), the specific dynamics and consequences of cultural learning may be rather unique and very important for understanding both micro- and macro-evolution.

3 Simultaneous evolution of innovation and social learning

The literature on gene-culture evolution often presents individual and social learning as alternatives. At some point in its lifetime, an organism is forced to choose between relying upon individual experience or socially-acquired information. This dichotomy obviously does not imply that social learning is free of inference or unguided by individually-learned theory. But evolutionary models almost always engage the strategic and population levels, not the psychological. At the level of abstraction of our models, these psychological platitudes are granted, and different strategic uses of information are the focus.

So while evolutionary models do treat individual learning and social learning as strategic alternatives, I know of no theory that treats them as orthogonal influences on behavior. Even in the simplest sort of model, individual and social learning interact at the population level and across generations (Rogers 1988). These interactions are antagonistic, in some cases. Social learning parasitizes innovation, in the same way that students who cheat on tests parasitize (or attempt to parasitize) students who study. To understand the pattern of behavior that emerges from these models, we have to appreciate that individuals are relying differentially upon both kinds of information.

Consider the genetic analogy again. Evolutionary theorists routinely speak of mutation and selection as separate “forces.” Each is governed in part by unique parameters, within an evolutionary model. And yet no one charges theorists with assuming mutation and selection do not interact. Indeed, there is still active theory development concerning the interaction of mutation and selection in determining patterns of genetic variation (Barton and Keightley 2002).

Similarly, gene-culture models have considered how individual and social learning interact, over an organism’s lifetime, as well as over evolutionary time. Boyd and Richerson (1985) devoted considerable space to the adaptive interaction of individually-acquired and socially-acquired information, *guided variation*. There are other, more subtle kinds of interactions. An adaptive synergy between individual innovation and social learning can arise from conditional effort in innovation (Boyd and Richerson, Enquist et al). Suppose that an organism can assess the efficacy of the information it acquires via social learning. The assessment can function through an initial trial or from analysis. If the organism judges that the behavior it acquired socially is below some threshold of efficacy, the organism has the option of then spending additional time innovating.

A second kind of adaptive synergy is when information gathered through individual learning can be build onto information gathered from social learning. When this is possible, the two kinds of learning interact to produce accumulations of information across generations (Boyd and Richerson 1996). There has been much less theory developed in the cumulative culture case. Previous models demonstrate that there may be a fitness valley between simple, inaccurate social learning that cannot accumulate complex behavior and more accurate social learning that can. Boyd and Richerson (1996) developed both a discrete behavior and truly continuous behavior versions of

their model. In both cases, there were values of the parameters for which cumulative cultural learning could not invade when rare but was stable when common. However, neither of these models allowed individual learning to co-evolve with social learning. Innovation rate was a fixed parameter, and this lack of feedback hid some interesting dynamics.

The model in the remainder of this chapter treats the simultaneous dynamics of innovation—an organism’s investment in acquiring new adaptive information directly from the relevant environment—and cumulative social learning—and organism’s investment in acquiring complex information from other individuals.

3.1 What is learning for?

Levins (1968) produced what is probably the first broad formal analysis of the evolution of simple phenotypic plasticity. He asked us to imagine a large number of different environmental states which may vary through time and across space. Suppose there is a unique allele that is optimal in each state. Now imagine an alternative strategy that, instead of providing a fixed phenotype, assesses the environment and attempts to change phenotype to match it. Levins extracted several general conclusions from analyzing models of this type.

1. Both spatial and temporal variation reduce fitness, compared to a uniform, static environment. Call this fitness cost the “naive cost.” It is what an unlearning, unresponsive organism pays, in the presence of environmental change.
2. An optimal strategy that either bet-hedges against change or learns can be found for any pattern of spatial/temporal change, but this optimal strategy cannot remove the entire naive cost.
3. A learning strategy itself imposes a cost on the organism, either because of sampling effort or processing costs. Call this cost the “learning penalty.” Therefore in order for learning to evolve, it must reduce the naive cost by an amount greater than the learning penalty.

These generalizations have remaining quite robust to the specific forms of models.

Spatial and temporal variation do generate different results, however. Germane to the arguments in this chapter, when the environment varies

through time, natural selection will favor a bet-hedging strategy that maximizes the geometric mean fitness of the organism, over the environmental states it experiences. Temporal variation does not maintain polymorphic fixed phenotypes, in the absence of learning. Spatial variation, on the other hand, may favor polymorphism in a species, depending upon the details.

The rest of this chapter builds a model that allows for both investments in individual exploration and, later, the ability to copy behavior and strategy from other individuals. In this section, I'll construct the individual learning and innovation core of the model. I'll analyze this core, before laying on social learning and analyzing the simultaneous dynamics of both cumulative social learning and innovation. By presenting the model in this way, I hope to lose fewer readers and better explain how cultural learning alters the outcomes.

Consider the evolution of simple learning, a form of phenotypic plasticity that uses information in the environment during an individual's development to alter phenotype, whether it is morphology or behavior. The evolution of the strategy of learning incorporates information into the genome that biases learning to be adaptive, rather than self-destructive.

Let's represent an individual's "genotype" with $d > 0$. The notion is that a number of regulatory and other genes combine to produce this continuous genotype that influences the amount of exploration and innovation. Individuals search and gather and process information about the environment so that each gains $q = d$ units of adaptive knowledge, in the current environment. These units translate into fitness, $\pi(q)$, with diminishing returns. I have explored a number of specific diminishing returns functions, but found all of them to produce the same qualitatively behavior. The easiest to analyze is:

$$\pi(q) = \frac{bq}{\beta + q} = \frac{bd}{\beta + d},$$

where $b > 0$ determines the strength of selection, and $\beta > 0$ is a parameter that determines the rate at which marginal fitness benefits decline. When β is large, $\pi(q_i)$ is approximately linear, and there are no diminishing returns to knowledge. When instead β is small, the fitness benefits of increasing knowledge diminish rapidly. Because of the specific form of the function π , β turns out to be exactly the value of q_i that produces half of the maximum value of π , $b/2$. That is, $\pi(\beta) = b/2$.

Suppose there are a very large number of environmental states. In each state the environment could take, different phenotypes are favored. Each

generation, there is a chance u that the environment changes to another random state. Since the variation here is stochastic, in the absence of phenotypic plasticity, a bet-hedging strategy will evolve that pays Levins’ naive cost of variability. Let fitness after paying this naive cost be w_0 .

Individuals can do better than this baseline, by attempting to learn the current state of the environment and use information from it to reduce the naive cost of variation. An individual who invests d in learning pays a cost cd , the learning penalty. Since investment in learning is continuous, this cost scales with it. As the fitness benefits of environmental knowledge, q , have diminishing returns, eventually the marginal benefits and costs of learning equal. At this point, selection will favor no further investments in plasticity.

These assumptions give us the following fitness, for an individual with genotype d :

$$w(d) = w_0 + \pi(d) - dc.$$

We find the evolutionarily stable investment in individual learning, \hat{d} , by solving $\partial w/\partial d = 0$ for d . This yields:

$$\hat{d} = \sqrt{\beta b/c} - \beta. \tag{1}$$

This is greater than zero, provided $b/c > \beta$, in which case selection favors learning. If this condition is not met, however, selection favors instead the bet-hedging fixed strategy that suffers the full naive cost of variation.

Are we ready yet to answer the question: what is learning for? According to this model, learning allows an organism to recoup fitness lost to temporal environmental variation. Note that I have assumed so far that this is an entirely asocial process. Fitness is not frequency dependent and there is no learning from conspecifics. In the next section, however, I add the possibility of social learning to the model. Then learning can be for building complex adaptations that fit the environment beyond the amount q .

3.2 Adding cumulative social learning

Many organisms are capable of phenotype plasticity. All primates—and indeed all mammals—are capable of individual learning of the kind modeled above. In novel circumstances, animals employ search strategies that may allow them to adaptively exploit new environments. One of the best understood of these is foraging in rats—rats explore trash eagerly, but sample

in small amounts and remember and avoid foods that make them ill (Galef 1996). This strategy allows rats to exploit varied urban environments, but it is possible because information about how to search for and use relevant information has been built into their genome by natural selection.

Animals also sometimes exhibit specialized adaptations for using conspecifics as cues of adaptive behavior. When Norway rats smell food on the muzzle of another rat, they are more likely to eat that same food (Galef 1996). Information in the rat’s genome makes this possible, by directing attention to odors on conspecifics and enhancing memory of socially-encountered foods.

In humans, the motivations and psychological adaptations that we might call “social learning” involve symbolic communication, abstraction, and substantial individual practice. Speech is a good model—while substantial social input is necessary for any human to learn the speech patterns of his or her community, a lot of individual practice with sounds is needed, because the inputs (sounds) are quite different than the information that an individual eventually needs to encode in order to produce them (motor memory). Every individual has a differently shaped vocal tract, and so in order to “imitate” another speaker, all of us had to experiment with sound production. Likewise, acquiring a complex skill like hunting or agriculture may require years of instruction and practice. Readers who have learned to play a musical instrument may find it to be a rich source of intuitions about the assumptions of this model. Playing the cello takes many years of individual practice, but this practice is much more effective when guided by a master cellist. A lone cellist may eventually attain the skill of a master, after many years of individual effort, but it is much easier to match or surpass the master, if the master provides instruction or simply allows observation. The purely “social” component of social transmission may be quite small, in terms of the time it occupies. But very little transmission, if any, is possible without the social component.

Begin with the model of individual innovation presented above. Assume now that there is another set of loci that influence an individual’s ability and motivation to learn socially. The “genotype” at those loci is represented by s , and an individual with $s > 0$ can successfully copy a fraction s of the adaptive behavior displayed by an adult from the previous generation. In order to separate innovation and social learning, I restrict $s < 1$, such that social learning will never accidentally generate behavior that is more adaptive than what was observed. Investments that increase s may be attentional—improvements in studying and representing the behavior of

other individuals—or motivational—increases in the extent to which goals and ways of achieving goals are open to social input. In both cases, greater investments in time or ability to acquire complex behavior from others results in the eventual acquisition of a larger portion of previously innovated behavior. If s is large enough, innovations generated over several generations may accumulate, generating behavior more complex than any individual could innovate in a single lifetime. If s remains low, however, then no amount of innovation will result in these complex behaviors, because each generation has to re-invent too much.

Increasing the accuracy of social learning is, however, costly. It costs the learner energy for upkeep and use of the psychology that makes social learning possible, and it costs the learner time in observation, practice, and missed opportunities to enhance fitness in other activities. I represent the total cost of social learning of this kind by ks . The more an individual invests in accurate copying of information, the more the individual pays.

With these assumptions, we can define a new fitness expression, now for a rare (mutant) individual investing d in innovation and s in social learning, in a population in which the common type invests d^* and s^* in each, respectively.

$$w(d, s, d^*, s^*) = w_0 + \pi(q_t) - dc - ks.$$

where q_t is the individual's behavioral phenotype, after both social transmission and innovation. Because learned information can be maintained across generations now, q will depend upon the amount of accumulated adaptive behavior in the population. This in turn depends upon the common phenotype, d^*, s^* , and the rate at which the environment changes and renders previously innovated behavior non-adaptive. The correct expression for q_t is:

$$q_t = (1 - u_t)sq' + d,$$

where u_t is a random variable taking the value 1 or 0, depending upon whether the environment changed last generation (with probability u) or not (probability $1 - u$), respectively. The symbol q' defines a recursion for the dynamics of behavior that is transmitted across generations. The behavior available to learn socially depends upon the common genotype, not that of the individual whose fitness we are modeling. The dynamics of behavior from one generation to the next are defined by:

$$q' = (1 - u_t)s^*q + d^*,$$

where q above is the average behavioral phenotype in the previous generation. Because u_t and u_{t-1} are random variables, there is no equilibrium amount of adaptive behavior in the population. Instead, q is reset to zero after each change in the environment and then begin climbing until the next change. One could assume instead that a proportion of adaptive information is retained across changes in the environment, but all this does is reset q to some minimum, rather than zero. There still will never be a stable value of q across generations. To cope with this kind of stochastic system, we solve for the mean of the stationary distribution of q . While there is no equilibrium, in a linear system like this one, the distribution of q across generations will eventually settle down. This is the system's stationary distribution. We can compute the mean of the stationary distribution, by taking expectations across generations and solving for \hat{q} , the mean of the stationary distribution. Doing this yields:

$$\hat{q} = \frac{d^*}{1 - s^*(1 - u)}.$$

This expression tells us that the mean level of adaptive behavior increases with increases in d^* and s^* , but decreases as u increases. It is helpful to consider some limiting cases. Suppose for example that the population has yet to evolve any effective social learning, $s^* = 0$. Then the average level of adaptive behavior will be $\hat{q} = d^*$. No adaptive behavior accumulates beyond what individuals can learn for themselves. Now suppose instead that $s^* = 1$. Now $\hat{q} = d/u$ —if u is small enough, substantial adaptive behavior will accumulate, because social learning is very (unrealistically) accurate.

Of course d^* and s^* are evolving genotypes. In order to analyze the simultaneous dynamics of innovation and social learning, we need to substitute \hat{q} into the fitness expression:

$$\begin{aligned} \hat{w}(d, s, d^*, s^*) &= w_0 + \pi((1 - u)s\hat{q} + d) - dc - ks, \\ &= w_0 + \pi\left(\frac{(1 - u)sd^*}{1 - s^*(1 - u)} + d\right) - dc - ks. \end{aligned}$$

Note that the adaptive behavior available for the mutant individual to acquire depends upon the population genotypes d^* and s^* , while the accuracy of her own social learning and power of her own innovation depend upon the individual genotypes d and s . In this way, the invading genotype plays against the population in game theoretic fashion. Our goal is to find the values of d^* and s^* that cannot be invaded by any other values d and s , respectively.

3.3 Joint dynamics of innovation and cumulative social learning

Before deriving the un-invadable values of innovation and social learning, it is useful to summarize the combined, two-dimensional, dynamics of this model. This system can evolve to two qualitatively different outcomes. First, social learning may increase when rare and evolve until its theoretical maximum. Second, social learning may be unable to invade when rare. Which of these two outcomes is realized depends upon the amount of innovation favored, when social learning is rare. If innovation is cheap, for example, then enough of it might be favored when social learning is absent. Social learning will then increase from $s^* = 0$, because there is complex information in the population worth copying. Once social learning begins to increase, however, selection favors less innovation, because of the diminishing fitness returns on knowledge. Eventually innovation may fall to the same level it was at, before social learning invaded. However, social learning remains high in the population. Once social learning can get a start from initially high innovation levels, it can invade.

Any potential evolutionarily stable values of d^* and s^* are found where $\partial w / \partial d|_{d=s=d^*,s^*} = 0$ and $\partial w / \partial s|_{d=s=d^*,s^*} = 0$. Call the evolutionarily stable values \hat{d} and \hat{s} . Setting $d^* = \hat{d}$, $s^* = \hat{s}$ and solving these equations for \hat{d} and \hat{s} yields one possible equilibrium for d^* :

$$\hat{d} = (\sqrt{\beta b/c} - \beta)(1 - \hat{s}(1 - u)). \quad (2)$$

Again note that innovation does not always evolve. If $\beta > b/c$, the expression above is negative and no innovation is favored by natural selection. The possibility of social learning, however, affects the stable amount of innovation. As \hat{s} increases, \hat{d} decreases. If the environmental rate of change u is large, however, then the effect of social learning on reducing \hat{d} is reduced.

Instead of having an equilibrium value, s^* can either decrease or increase until it reaches zero or one (or another theoretical upper limit). That is, $\hat{s} = 1$ or $\hat{s} = 0$, depending upon the parameters. The condition for social learning to increase from zero, and so invade a population, is given by $\partial w / \partial s|_{d=\hat{d},s=s^*} > 0$. This reduces to:

$$u < 1 - \frac{k}{\sqrt{b\beta c} - \beta c}.$$

If the environment changes too quickly, social learning is never favored. But if the marginal cost, k , of social learning is low enough and fitness benefits of behavior do not diminish too rapidly, then social learning will invade and increase until its theoretical limit.

These expressions do not immediately reveal what is happening, however. It is easier to understand the behavior of this model, by visualizing the joint evolution of innovation and cumulative social learning. Figure 1 shows the phase diagram of this model, for two different sets of parameter values. In each plot, position along the horizontal axis represents the value of d^* , from zero to one. Position along the vertical axis represents s^* , also from zero to one. Arrows represent the direction and magnitude of change for the system, at each point. The red point in each plot is the eventually evolutionarily stable combination of innovation and social learning, in each case. On the left, the cost of innovation is set high, but not so high as to prevent individual learning from evolving at all. The high cost, however, does prevent d^* from ever evolving to high enough values to provide enough adaptive behavior to be worth investing in accurate social learning. Therefore, where ever the system begins, selection will eventually reduce social learning to its minimum. Cumulative culture does not evolve in this case, although rather fancy behavior is invented each generation, because of the non-zero equilibrium value of d^* .

On the right, the cost of innovation is slightly reduced. Suppose the system begins in the lower-left corner, at $d^*, s^* = 0$. Now innovation can increase to a higher level than on the left, before the arrows turn the other way and selection no longer favors any increases in innovation. Innovation can reach a high enough level, in fact, that the behavior that is invented each generation is now worth copying through investments in accurate social learning. Therefore the system evolves towards the interior, favoring increasing amounts of social learning as it heads for the top of the figure. As selection favors social learning, however, it also favors less innovation (Expression 2). Thus the eventual equilibrium has highly accurate social learning (\hat{s} is near one), but lower levels of innovation than the plot on the left. The behavior invented each generation is modest in comparison to the population in which social learning did not evolve. However the mean level of adaptive behavior is twice as large. On the left, $\hat{q} = \hat{d} = 0.29$. On the right, $\hat{q} = 20, \hat{d} = 20(0.029) = 0.58$.

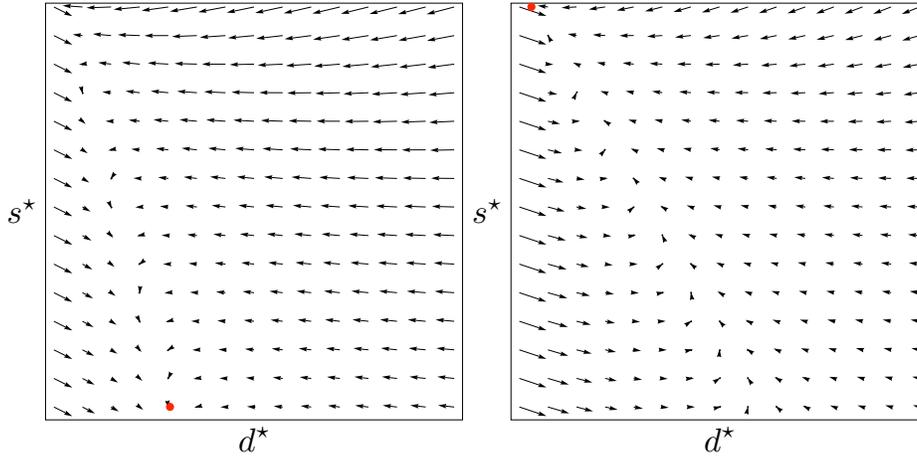


Figure 1: Evolutionary dynamics of innovation and cumulative social learning. Arrows show the direction and magnitude of evolutionary change at each point in the possible state space of the population, defined by the average investment in innovation, d^* , and the average investment in social learning, s^* . Red dots in each panel show the only stable equilibrium in each case. In both panels, $b = 5, u = 0.05, \beta = 1, k = 1$. Left panel: $c = 3$. Right panel: $c = 2$. On the left, higher costs of innovation prevent individually learned information from reaching high enough levels for natural selection to favor cumulative social learning. Without much behavior worth copying, the system remains at a high level of innovation, but no imitation evolves. On the right, a slightly reduced cost of innovation leads initially to a higher investment in individual learning, a higher level of individually acquired behavior, and eventually to the invasion of social learning. As social learning increases, however, natural selection favors reduced investments in innovation, because of the diminishing fitness returns to knowledge. This system comes to rest where innovation is lower than the panel on the left, but social learning is highly accurate.

3.4 How much cumulative culture?

There is an irony lurking within the solution above, however. While the evolution of social learning appears to have resulted in higher levels of adaptive behavior, “culture,” \hat{q} , in reality social learning has only provided a cheaper way to attain the same amount of adaptive behavior the population would have enjoyed, if it had relied entirely upon high levels of innovation. This is obvious, once we inspect the expression for stationary mean \hat{q} , at \hat{d}, \hat{s} . Let \hat{q} be the value of \hat{q} , evaluated at $d^* = \hat{d}, s^* = \hat{s}$. Then the expression for this expected average level of adaptive behavior is:

$$\hat{q} = \frac{\sqrt{b\beta c} - \beta c}{c} = \sqrt{\beta b/c} - \beta.$$

Note that this expression does not contain \hat{s} . Therefore, it does not depend upon social learning at all. Furthermore, it is the same amount of adaptive behavior we would expect from the a-cultural model presented earlier (Expression 1)! Cumulative social learning has evolved, but it has failed in this model to produce long-term information gains beyond what would already have been possible using (highly advanced) innovation.

What is happening in the evolutionary economics is that lower costs of innovation allow behavior to reach a threshold that then allows social learning to invade. Once social learning invades, selection favors less innovation. Because behavior has diminishing returns, individuals do better by investing in an optimal mix of innovation and social learning. This optimal mix trades off the costs of innovation against the potentially unreliable benefits of socially learned behavior. Because individual benefit is driving the evolution of both innovation and cumulative culture in this model, selection does not necessarily maximize the group benefits of cumulative culture.

One way out of this unsatisfactory result is to note that we have only modeled a single domain of behavior. Social learning ability will be applied potentially to other domains with much lower relevant rates of environmental change. Consider that bows and arrows continue to function, even when climate changes substantially. Therefore different technologies and strategies experience different rates of change (u in the model). If fitness gains from more-slowly changing domains are important enough, then social learning will be pulled up to a higher level of accuracy, even in fast-changing domains, than would be optimal, if we consider those domains alone.

But this is a hand-waving argument. Are there other theoretical solutions

that do not invoke large numbers of parameters that are poorly understood and potentially unmeasurable?

3.5 When social learning enhances innovation

How can we get selection to increase adaptive behavior beyond this selfish optimum? One way is by allowing social learning to improve the efficiency of innovation. This hypothesis is reasonable, if you believe that the psychological abilities that make cultural transmission possible also enhance an individual's ability to represent, remember, and explore new solutions. For example, language is a symbolic capacity that allows us to represent abstract systems, much like the model in this chapter. While language makes it possible to acquire complex behavior from other people, it also makes it possible to organize and transform information within one's own head. The dialog scientists carry on with themselves—sometimes out loud—suggests that at least some abilities that are possibly selected for enhancing social learning can simultaneously enhance imagination and innovation. If one of the things social learning does to human cognition is provide a quite open motivational and association system, so that we can remember arbitrary scripts and develop novel goals through communication, then any energies given to innovation may be able to tap these same abilities.

When we allow new synergy between social learning and innovation (so that social learning actually makes innovation cheaper), we have a new fitness expression:

$$w(d, s, d^*, s^*) = w_0 + \pi((1 - u)s\hat{q} + d) - (1 - s/z)cd - ks.$$

The parameter $z > 1$ determines the amount of synergy. When $z = \infty$, the above reduces to the previous fitness expression and the result is unchanged. When z is small, however, there may be substantial cost reductions to innovation as social learning abilities increase. When $z = s$, innovation is effectively free (note that this is impossible, by the constraint that $z > 1$). The new steady state accumulated culture becomes:

$$\hat{q} = \frac{bz}{\sqrt{bcz(z - \hat{s})}} - \beta. \quad (3)$$

Figure 2 plots this expression over all possible values of \hat{s} , for two values of z . As $z \rightarrow \hat{s}$ from above, this quantity increases rapidly. In biological terms, as

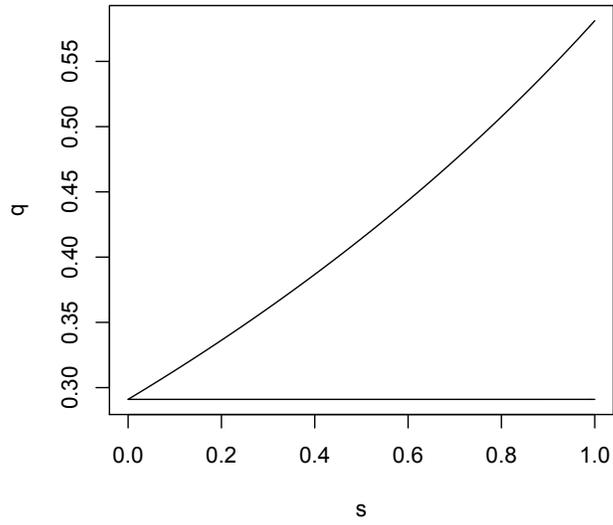


Figure 2: The expected amount of adaptive behavior, Expression 3, as a function of the amount of cumulative social learning in the population, \hat{s} . Horizontal line: $b = 5, c = 3, \beta = 1, z = \infty$. Sloped line: $z = 3$. When social learning reduces the costs of innovation, the evolution of social learning leads to increases in adaptive behavior, beyond what innovation alone could provide. Otherwise, selection adjusts the amount of innovation so that the amount of adaptive behavior remains the same, whether social learning invades or not.

social learning increasingly makes innovation cheaper and more efficient, the steady state amount of adaptive behavior increases. In contrast, as $z \rightarrow \infty$, the above approaches the previous expression for \hat{q} , $\sqrt{\beta b/c} - \beta$, resulting in no change in the amount of adaptive behavior as social learning increases.

4 Where did culture come from?

Natural selection builds adaptive information into the genome. Learning gathers information about the environment, to be used by information in the genome. Cumulative social learning takes information from behavior—whatever its source—and allows it to be stored and accumulated in human brains. The obvious adaptive utility of the products of this process—technologies and strategies too complex for any individual to invent in his or her own lifetime—make puzzling the gap between humans and other apes in this regard. If “culture” is such a great adaptive trick for genes to acquire, then why aren’t other apes similarly cultural? This question is additionally puzzling, given the evidence of at least proto-cultural social learning abilities in chimpanzees (see Whiten, this volume).

The theory I have reviewed and developed in this chapter addresses the question of the origins of human cultural abilities. The first goal of the theory is to understand how natural selection on genes can fail to favor cumulative social learning and under what conditions it will lead to cultural evolution and accumulation. The second goal is to understand how the population-level adaptive benefits of this accumulation can appear, without these being the selective reasons for investments in learning.

4.1 Evolving cultural evolution

The first goal is addressed by the combined dynamics of innovation and social learning. When social learning allows accumulation and costs more and more as the complexity of what is copied or the accuracy with which it is copied increases, then a fitness valley can appear between an a-cultural population and a cultural population (Boyd and Richerson 1996). If individual learning is effective enough, however, the model in this chapter suggests that it can provide a way around this valley. If selection favors improvements in innovation, independent of cumulative social learning, eventually there is complex behavior that—while not accumulated across generations—is nev-

ertheless worth copying, because the costs of social learning are lower than those of innovation itself. Proximately, the lower costs of social learning may arise because innovation is an inherently harder activity. Many good ideas are hard to stumble upon, and much individually learned behavior takes a lifetime to assemble, despite not being a product of social learning. Once complex behavior is available, selection might favor acquiring it before any individual effort is made in innovation. Once this happens, selection trades off innovation against social learning, reducing the once-large innovation rate that was needed to cross the cultural valley. This may or may not be the right idea, to explain the gap between humans and other apes. But it points in the same direction as others have suggested (see for example Whiten, this volume). Innovation and social learning are potentially co-adapted in humans, and explaining one may ultimately require an equally deep understanding of the evolution of the other.

In the end, the theory here required investments in social learning to make innovation more effective or cheaper. A number of other alternations could be made to the model, to modify the relationship between innovation and cumulative social learning. An obvious one is to allow the amount of adaptive behavior that is socially learned, rather than the efficiency of social learning itself, to enhance innovation. The idea is that previously evolved information may make future innovation easier, because it defines the relevant parts of the problem and provides tools to finding solutions. Much of how modern science works involves the development and dissemination of tools, not products. In this way, science is as much about building intellectual and technological solutions for discovery as it is about discovery itself.

Similarly, many of the social institutions and cooperative arrangements in modern societies enhance innovation. Governments actively structure patent law, so that more innovation is encouraged than would be individually optimal for firms. While patent law doesn't necessarily become more effective as knowledge accumulates, further enhancing innovation, there are other institutions which might. Division of labor and the exchange institutions that make it possible also enhance innovation, in two ways. First, division of labor carries with it the benefits of specialization. Economies of scale make innovation easier in each domain of behavior, and new information can be traded among specialists more easily than it can be independently discovered by all of them. Second, as culture accumulates, eventually the sum of what the population knows exceeds what any individual can learn, even with advanced social learning. The readers of this chapter are probably among the most ed-

ucated people on the planet, and yet each is unlikely to be expert in more than one or two areas of science. Your author spent a decade learning to understand the intersection of anthropology and evolutionary ecology, and yet he still has little deep understanding of some branches of both anthropology and ecology. Like most scientists, he relies upon experts in other areas—combined with active skepticism and habits of thought—to keep track of relevant advances in neighboring fields. This division of labor allows knowledge in any particular domain—hunting large animals versus gathering palm fiber or processing medicinal plants versus childcare—to grow beyond the limits of individuals to learn and practice all domains.

4.2 Evolvability as a side effect

The second goal of this chapter has been to highlight the kind of theory that is required to understand the accumulation of socially-transmitted adaptive behavior, without the eventual highly-adaptive accumulations being the initial reason for the evolution of the psychology needed to make cumulative culture possible. Human societies are no doubt more adaptable than those of other apes—we have conquered (and according to many, ruined) nearly every environment on the planet, while other apes shrink in tropical refuges. Part of the explanation for our world dominance is the ability to generate complex, locally adapted behavior over generations (Richerson and Boyd 2005). Foragers in different parts of the world need quite different knowledge and strategy. The combination of innovation and social transmission makes local specialization and regional adaptability both possible.

And yet, selection does not favor costly social learning abilities, unless there is an immediate benefit to the organism. Our ancestors did not lug around brains capable of cumulative culture, because it would turn out to allow our species to dominate the planet. Instead, we have to seek short-term, individual fitness benefits in order to explain why an organism would cross the cultural gap. An acorn detects moisture when it decides whether or not to germinate, because acorns that were initially slightly sensitive to a moisture gradient produced more descendants. These descendants then had mutations that favored more sensitivity, until some rough optimum was reached. Selection favored every step, even though the eventual level of adaptation was higher than the initial. Similarly, the theory in this chapter hypothesizes that cumulative social learning began as a way to avoid the costs of innovation. Especially as learned behavior becomes more complex,

social learning allows an individual to rapidly acquire sensible locally-adapted behavior, saving time and energy for other activities. As each individual continues to add some continued improvement to what is learned socially, the average adaptiveness of behavior may increase over generations. However, selection favored each step along the way because of the benefits and costs at each step, not because of the population-level benefits that would eventually arise.

The specific model developed in this chapter suggests that one path to evolving cultural evolution lies in first getting selection to favor increases in innovation, as summarized just above. However any successful theory of the evolution of evolvability must contend with this same challenge. Students of the evolution of development (“evo-devo”) are fond of noting how animal body plans can make life very evolvable, over macro-evolutionary time. Developmental genes are organized in such a way as to make compartmentalized changes possible—the genome can make one set of limbs longer or even replace them with the genetic information for another specialized set (see Kirschner and Gerhart 1998). But while this source of evolutionary novelty may turn out to explain the very long term success of some groups of organisms (like bilaterally symmetric animals), it cannot be the reason the body plan arose in the first place. Higher-level selection, at the population or species level, can indeed explain the maintenance of such adaptations. A popular theory of the maintenance of sexual reproduction suggests that sex indeed makes populations more evolvable (Maynard Smith 1978).

References

- Barton, N. and Keightley, P. (2002). Understanding quantitative genetic variation. *Nature Reviews Genetics*, 3:11–21.
- Boyd, R. (1982). Density-dependent mortality and the evolution of social interactions. *Animal Behaviour*, 30:972–982.
- Boyd, R. and Richerson, P. (1996). Why culture is common, but cultural evolution is rare. *Proceedings of the British Academy*, 88:77–93.
- Boyd, R. and Richerson, P. J. (1985). *Culture and the Evolutionary Process*. Univ Chicago Press, Chicago.

- Cavalli-Sforza, L. L. and Feldman, M. W. (1981). *Cultural transmission and evolution: a quantitative approach*. Princeton University Press, Princeton.
- Dawkins, R. (1982). *The Extended Phenotype: The Long Reach of the Gene*. Oxford University Press, Oxford.
- Durham, W. H. (1991). *Coevolution: Genes, Culture, and Human Diversity*. Stanford University Press, Stanford.
- Fisher, R. A. (1918). The correlation between relatives on the supposition of Mendelian inheritance. *Transactions of the Royal Society of Edinburgh*, 52:399–433.
- Frank, S. A. (2009). Natural selection maximizes fisher information. *Journal of Evolutionary Biology*, 22:231–244.
- Galef, B.G., J. (1996). Social enhancement of food preferences in norway rats: A brief review. In *Social Learning and Imitation: The Roots of Culture*.
- Henrich, J. and Boyd, R. (2002). On modeling cognition and culture: Why replicators are not necessary for cultural evolution. *Journal of Cognition and Culture*, 2:87–112.
- Henrich, J., Boyd, R., and Richerson, P. J. (forthcoming). Five common mistakes in cultural evolution. In Sperber, D., editor, *Epidemiology of Ideas*. Open Court Publishing.
- Jablonka, E. and Lamb, M. (2005). *Evolution in Four Dimensions*. MIT Press, Massachusetts.
- Jablonka, E. and Lamb, R. M. (1991). Sex chromosomes and speciation. *Proceedings of the Royal Society of London, B*, 243:203–208.
- Jenkin, F. (1864). The origin of species. *North British Review*, 46:277–318.
- Kirschner, M. and Gerhart, J. (1998). Evolvability. *Proc. Natl Acad. Sci.*, 95:8420–8427.
- Leimar, O. (1996). Life history analysis of the Trivers-Willard sex-ratio problem. *Behavioral Ecology*, 7:316–325.

- Levins, R. (1968). *Evolution in Changing Environments*. Princeton University Press, Princeton.
- Maynard Smith, J. (1978). *The evolution of sex*. Cambridge University Press, Cambridge, UK.
- Maynard Smith, J. (1990). Models of a dual inheritance system. *Journal of Theoretical Biology*, 143:41–53.
- Maynard Smith, J. (1998). *Evolutionary Genetics*. Oxford University Press, Oxford.
- Odling-Smee, F. J., Laland, K. N., and Feldman, M. W. (2003). *Niche Construction: The Neglected Process in Evolution*. Number 37 in Monographs in Population Biology. Princeton University Press, Princeton.
- Pál, C. and Miklós, I. (1999). Epigenetic inheritance, genetic assimilation and speciation. *Journal of Theoretical Biology*, 200:19–37.
- Richerson, P. J. and Boyd, R. (2005). *Not by Genes Alone: How culture transformed human evolution*. University of Chicago Press, Chicago.
- Rogers, A. R. (1988). Does biology constrain culture? *American Anthropologist*, 90:819–831.
- Sheahan, M. B., Rose, R. J., and McCurdy, D. W. (2004). Organelle inheritance in plant cell division: the actin cytoskeleton is required for unbiased inheritance of chloroplasts, mitochondria and endoplasmic reticulum in dividing protoplasts. *The Plant Journal*, 37:379–390.
- Silk, J. B. and Boyd, R. (1983). Female cooperation, competition, and mate choice in matrilineal macaque groups. In Wasser, S. K., editor, *Social Behavior of Female Vertebrates*, pages 315–347. Academic Press, New York.
- Sperber, D. (2000). An objection to the memetic approach to culture. In Aunger, R., editor, *Darwinizing Culture: The Status of Memetics as a Science*, pages 163–173. Oxford University Press, Oxford.
- Szathmáry, E. and Maynard Smith, J. (1995). The major evolutionary transitions. *Nature*, 374:227–232.