HISTORY AND EVOLUTION

edited by

Matthew H. Nitecki

and

Doris V. Nitecki

1992

State University of New York Press
How Microevolutionary Processes Give Rise to History

Robert Boyd and Peter J. Richerson

Over the last decade a number of authors, including ourselves, have attempted to understand human cultural variation using Darwinian methods. This work is unified by the idea that culture is a system of inheritance: individuals vary in their skills, habits, beliefs, values, and attitudes, and these variations are transmitted to others through time by teaching, imitation, and other forms of social learning. To understand cultural change we must account for the microevolutionary processes that increase the numbers of some cultural variants and reduce the numbers of others.

Social scientists have made a number of objections to this approach to understanding cultural change. Among these is the idea that culture can only be explained historically. Because the history of any given human society is a sequence of unique and contingent events, explanations of human social life, it is argued, are necessarily interpretive and particularistic. Present phenomena are best explained mainly in terms of past contingencies, not ahistorical adaptive processes that would erase the trace of history. Like other scientific (rather than historical) explanations of human cultures, the argument goes, Darwinian models cannot account for the lack of correlation of environmental and cultural variation, nor the long term trends in cultural change.

In this chapter, we defend the Darwinian theories of cultural change against this objection by suggesting that several cultural evolutionary processes can give rise to divergent evolutionary developments, secular trends, and other features that can generate unique historical sequences for particular societies. We also argue that Darwinian theory offers useful tools for those interested in understanding the evolution of particular societies. Essentially similar processes act in the case of organic evolution. Darwinian theory
is both scientific and historical. The history of any evolving lineage or culture is a sequence of unique, contingent events. Similar environments often give rise to different evolutionary trajectories, even among initially similar taxa or societies, and some show very long-run trends in features such as size. Nonetheless, these historical features of organic and cultural evolution can result from a few simple microevolutionary processes.

A proper understanding of the relationship between the historical and the scientific is important for progress in the social and biological sciences. There is (or ought to be) an intimate interplay between the study of the unique events of given historical sequences and the generalizations about process constructed by studying many cases in a comparative and synthetic framework. The study of unique cases furnishes the data from which generalizations are derived, while the generalizations allow us to understand better the processes that operated on particular historical trajectories. We cannot neglect the close, critical study of particular cases without putting the data base for generalization in jeopardy. Besides, we often have legitimate reasons to be curious about exactly how particular historical sequences, such as the evolution of *Homo sapiens*, occurred. On the other hand, it is from the study of many cases that we form a body of theory about evolutionary processes. No one historical trajectory contains enough information to obtain a very good grasp of the processes that affected its own evolution. Data are missing because the record is imperfect. The lineage may be extinct, and so direct observation is impossible. Even if the lineage is extant, experimentation may be impossible for practical or ethical reasons. Potential causal variables may be correlated in particular cases, and so understanding their behavior may be impossible. The comparative method can often clarify such cases. "Scientists" need "historians" and vice versa.

**Darwinian Models of Cultural Evolution**

Over the past two decades, a number of scholars have attempted to understand the processes of cultural evolution in Darwinian terms. Social scientists (Campbell 1965, 1975; Cloak 1975; Durham 1976; Ruyle 1973) have argued that the analogy between genetic and cultural transmission is the best basis
for a general theory of culture. Several biologists have considered how culturally transmitted behavior fits into the framework of neo-Darwinism (Pulliam and Dunford 1980; Lumsden and Wilson 1981; Boyd and Richerson 1985; Richerson and Boyd 1989a; Cavalli-Sforza and Feldman 1983; Rogers 1989). Other biologists and psychologists have used the formal similarities between genetic and cultural transmission to develop theories describing the dynamics of cultural transmission (Cavalli-Sforza and Feldman 1973, 1981; Cloninger et al. 1979; Eaves et al. 1978). All of these authors are interested in a synthetic theory of process applying to how culture works in all cultures, including in other species which might have systems with a useful similarity to human culture. Note that this last broadly comparative concern is likely to be useful in dissecting the reasons why the human lineage originally became more cultural than typical mammals.1

The idea that unifies the Darwinian approach is that culture constitutes a system of inheritance. People acquire skills, beliefs, attitudes, and values from others by imitation and enculturation (social learning), and these "cultural variants" together with their genotypes and environments, determine their behavior. Since determinants of behavior are communicated from one person to another, individuals sample from and contribute to a collective pool of ideas that changes over time. In other words, cultures have similar population-level properties as gene pools, as different as the two systems of inheritance are in the details of how they work. (In one respect, the Darwinian study of cultural evolution is more Darwinian than the modern theory of organic evolution. Darwin not only used a notion of "inherited habits" that is much like the modern concept of culture, but also thought that organic evolution generally included the property of the inheritance of acquired variation, which culture does and genes do not.)

Because cultural change is a population process, it can be studied using Darwinian methods. To understand why people behave as they do in a particular environment, we must know the nature of the skills, beliefs, attitudes, and values that they have acquired from others by cultural inheritance. To do this we must account for the processes that affect cultural variation as individuals acquire cultural traits, use the acquired information to guide behavior, and act as models for others. What processes increase or decrease the proportion of people in a society who hold particular ideas
about how to behave? We thus seek to understand the cultural analogs of the forces of natural selection, mutation, and drift that drive genetic evolution. We divide these forces into three classes: random forces, natural selection, and the decision-making forces.

Random forces are the cultural analogs of mutation and drift in genetic transmission. Intuitively, it seems likely that random errors, individual idiosyncrasies, and chance transmission play a role in behavior and social learning. For example, linguists have documented a good deal of individual variation in speech, some of which is probably random individual variation (Labov 1972). Similarly, small human populations might well lose rare skills or knowledge by chance, for example, due to the premature deaths of the only individuals who acquired them (Diamond 1978).

Natural selection may operate directly on cultural variation. Selection is an extremely general evolutionary process (Campbell 1965). Darwin formulated a clear statement of natural selection without a correct understanding of genetic inheritance because it is a force that will operate on any system of inheritance with a few key properties. There must be heritable variation, the variants must affect phenotype, and the phenotypic differences must affect individuals' chances of transmitting the variants they carry. That variants are transmitted by imitation rather than sexual or asexual reproduction does not affect the basic argument, nor does the possibility that the source of variation is not random. Darwin imagined that random variation, acquired variation, and natural selection all acted together as forces in organic evolution. In the case of cultural evolution, this seems to be the case. It may well be, however, that behavioral variants favored by natural selection depend on the mode of transmission. The behaviors that maximize numbers of offspring may not be the same as those that maximize cultural influence on future generations (Boyd and Richerson 1985).

Decision-making forces result when individuals evaluate alternative behavioral variants and preferentially adopt some variants relative to others. If many of the individuals in a population make similar decisions about variants, especially if similar decisions are made for a number of generations, the pool of cultural variants can be transformed. Naive individuals may be
exposed to a variety of models and preferentially imitate some rather than others. We call this force biased transmission. Alternatively, individuals may modify existing behaviors or invent new ones by individual learning. If the modified behavior is then transmitted, the resulting force is much like the guided, nonrandom variation of "Lamarckian" evolution. Put differently, humans are embedded in a complex social network through which they actively participate in the creation and perpetuation of their culture.

The decision-making forces are derived forces (Campbell 1965). Decisions require rules for making them, and ultimately the rules must derive from the action of other forces. That is, if individual decisions are not to be random, there must be some sense of psychological reward or similar process that causes individual decisions to be predictable, in given environments at least. These decision-making rules may be acquired during an earlier episode of cultural transmission, or they may be genetically transmitted traits that control the neurological machinery for acquisition and retention of cultural traits. The latter possibility is the basis of the sociobiological hypotheses about cultural evolution (Alexander 1979; Lumsden and Wilson 1981). These authors, among others, argue that the course of cultural evolution is determined by natural selection operating indirectly on cultural variation through the decision-making forces.

Like natural selection, the decision-making forces may improve the fit of the population to the environment. The criteria of fit depend on the nature of the underlying decision rules. This is easiest to see when the goals of the decision rules are closely correlated with fitness. If human foraging practices are adopted or rejected according to their energy payoff per unit time (optimal foraging theory’s operational proxy for fitness), then the foraging practices used in the population will adapt to changing environments much as if natural selection were responsible. If the adoption of foraging practices is strongly affected by consideration of prestige, say that associated with male success hunting dangerous prey, then the resulting pattern of behavior may be different. However, there will still be a pattern of adaptation to different environments, but now in the sense of increasing prestige rather than calories.
What Makes Change Historical?

It has often been argued that historical and scientific explanations are different in kind. Ingold (1985) gives two important versions of this argument. Some authors (e.g., Collingwood 1946) argue that history is uniquely human because it entails conscious perception of the past. The second view (e.g., Trigger 1978) is quite different, and holds that history involves unique, contingent pathways from the past to the future that are strongly influenced by unpredictable, chance events. We focus on the latter view here. For example, capitalism arose in Europe rather than China, perhaps because Medieval and Early Modern statesmen failed to create a unified empire in the West (McNeill 1980), and marsupials dominate the Australian fauna perhaps because of Australia’s isolation from other continents in which placental mammals chanced to arise. In contrast, it is argued, scientific explanations involve universally applicable laws. In evolutionary biology and in anthropology, these often take the form of functional explanations, in which only knowledge of present circumstances and general physical laws (e.g., the principles of mechanics) are necessary to explain present behavior (Mitchell and Valone 1990). For example, long fallow horticulture is associated with tropical forest environments, perhaps because it is the most efficient subsistence technology in such environments (Conklin 1969).

It has been argued, perhaps nearly as often, that this dichotomy is false. Eldredge (1989:9) provides a particularly clear and forceful example of a common objection: all material entities have properties that can change through time. Even simple entities like molecules are characterized by position, momentum, charge, and so on. If we could follow a particular water molecule, we would see that these properties changed through time—even the water molecule has a history, according to Eldredge. Yet, everyone agrees that we can achieve a satisfactory scientific theory of water. Historical explanations, Eldredge argues, are just scientific explanations applied to systems that change through time. We are misled because chemists tend to study the average properties of very large numbers of water molecules.

This argument explains too much. Not all change with time is history in the sense intended by historically oriented biologists and social scientists. To see this consider an electrical circuit composed of a voltage source, a
capacitor and a fluorescent light. Under the right conditions, the voltage will oscillate through time, and these changes can be described by simple laws. Are these oscillations historical? In Eldredge's view they are; the circuit has a history, a quite boring one, but a history nonetheless. Yet such a system does not generate unique and contingent trajectories. After the system settles down one oscillation is just like the previous one, and the period and amplitude of the oscillations are not contingent on initial conditions. They are not historical in the sense that "one damn thing after another" (Elton 1967:40) leads to cumulative, but unpredictable change.

What then makes change historical? We think that two requirements capture much of what is meant by "history." These two requirements pose a more interesting and serious challenge for reconciling history with a scientific approach to explanation. Consider a system like a society or a population that changes through time both under the influence of internal dynamics and exogenous shocks. Then we suggest that the pattern of change is historical if:

A. Trajectories are not stationary on the time scales of interest. History is more than just change – it is change that does not repeat itself. On long enough time scales, the oscillations in the circuit become stationary, meaning that the chance of finding the system in any particular state becomes constant. Similarly, random day-to-day fluctuations in the weather do not constitute historical change if one is interested in organic evolution because, on long evolutionary time scales the there will be so many days of rain, so many days of sun, and so on. By choosing a suitably long period of time, we can construct a scientific theory of stationary processes using a statistical rather than strictly deterministic approach. In the case of nonstationary historical trajectories, a society or biotic lineage tends to gradually become more and more different as time goes by. There is no possibility of basing explanation on, say, a long-run mean about which the historical entity fluctuates in some at least statistically predictable way, because the mean calculated over longer and longer runs of data continues to change significantly. One of the most characteristic statistical signatures of nonstationary processes is that the variance they produce grows with time rather than converging on a finite value. Note that a process that is historical in one spatio-temporal frame may not be in another. If we are not too interested in a specific species or
societies in given time periods, we can often average over longer periods of
time or many historical units to extract ahistorical generalizations. Any given
water molecule has a history, but it is easy to average over many of them and
ignore this fact.

B. Similar initial conditions give rise to qualitatively different trajectories.
Historical change is strongly influenced by happenstance. This requires that
the dynamics of the system must be path dependent; isolated populations or
societies must tend to diverge even when they start from the same initial
condition and evolve in similar environments. Thus, for example, the spread
of a favored allele in a series of large populations is not historical. Once the
allele becomes sufficiently common it will increase at first exponentially, and
then slowly, asymptotically approaching fixation. Small changes in the initial
frequencies, population size, or even degree of dominance will not lead to
qualitative changes in this pattern. In separate but similar environments,
populations will converge on the favored allele. Examples of convergence in
similar environments are common—witness the general similarity in tropical
forest trees and many of the behaviors of the long fallow cultivators who live
among them the world over. On the other hand, there are also striking
failures of convergence—witness the many unique features of Australian
plants, animals, and human cultures. The peculiar hanging leaves of eucalypts,
the bipedal gait of kangaroos, and the gerontocratic structure of Aus-
tralian Aboriginal societies make them distinctively different from the in-
habitants of similar temperate and subtropical dry environments on other
continents.

It is important not to blur the distinction between simple trajectories
and true historical change; it is easy to see how evolutionary processes like
natural selection give rise to simple, regular change like the spread of a
favored allele or subsistence practice. However, it is not so easy to see how
such processes give rise to unique, contingent pathways. Scientists tout the
approach to steady states and convergence in similar situations as evidence
for the operation of natural "laws," so it seems natural to conclude that a
lack of stationarity and convergence are evidence of processes that cannot be
subsumed in the standard conceptions of science. Our argument is that
things are not at all that simple. There is every reason to expect that per-
fectly ordinary scientific processes, ordinary in the sense that they result from
natural causes and are easily understood by conventional methods, regularly generate history in the sense defined by these two criteria.

**How Do Adaptive Processes Give Rise to History?**

Let us begin with the two most straightforward answers to this question. First, it could be that most evolutionary change is random. Much change in organic evolution may be the result of drift and mutation, and much change in cultural evolution may result from analogous processes. The fact that drift is a very slow process would explain long-term evolutionary trends. Raup (1977) and others argue that random-walk models produce phylogenies that are remarkably similar to real ones. The fact that cultural and genetic evolutionary change is random would allow populations in similar environments to diverge from each other. It seems likely that some variation in both cases evolves mainly under the influence of nonadaptive forces — for example, much of the eukaryotic genome does not seem to be expressed and evolves under the influence of drift and mutation (Futuyma 1986:447). Similarly, the arbitrary character of symbolic variation suggests that nonadaptive processes are likely to be important in linguistic change and similar aspects of culture. In both cases, isolated populations diverge at an approximately constant rate on the average. However, to understand why a particular species is characterized by a particular DNA sequence, or why a particular people use a particular word for mother, one must investigate the sequence of historical events that led to the current state.

It also possible that historical change is generated by abiotic environmental factors (Valentine and Moores 1972). Long-term trends in evolution could result from the accurate adaptive tracking of a slowly changing environment. For example, during the last hundred million years there has been a long-term increase in the degree of armoring of many marine invertebrates living on rocky substrates and a parallel increase in the size and strength of feeding organs among their predators (Vermeij 1987; Jackson 1988). It is possible that these biotic trends have been caused by long-run environmental changes over the same period — for example, an increase in the oxygen content of the atmosphere (Holland 1984). Similarly, beginning
perhaps as much as 17,000 years ago, humans began a shift from migratory big game hunting to sedentary, broad spectrum, more labor-intensive foraging, finally developing agriculture about 7,000 years ago (Henry 1989). Many authors (e.g., Reed 1977) have argued that the transition from glacial to interglacial climate that occurred during the same period is somehow responsible for this change. Similarly, differences among populations in similar environments may result from the environments actually being different in some subtle but important way. For example, Westoby (1989) has argued that some of the unusual features of the Australian biota result from the continent-wide predominance of highly weathered, impoverished soils on this relatively undisturbed continental platform. Perhaps the failure of agriculture to develop in or diffuse to Aboriginal Australia, despite many favorable preconditions and the presence of cultivators just across the Torres Strait, also reflects poor soils.

It is more difficult to understand how adaptive processes like natural selection can give rise to historical trajectories. There are two hurdles: First, adaptive processes in both organic and cultural evolution appear to work on rather short time scales compared to the time scales of change known from the fossil record, archaeology and history. Theory, observation, and experiment suggest that natural selection can lead to change that is much more rapid than any observed in the fossil record (Levinton 1988:342-47). For example, the African Great Lakes have been the locus of spectacular adaptive radiations of fishes amounting to hundreds of highly divergent forms from a few ancestors in the larger lakes (Lowe-McConnell 1975). The maximum time scales for these radiations, set by the ages of the lakes and not counting that they may have dried up during the Pleistocene, are only a few million years. The radiation in Lake Victoria (about 200 endemic species) seems to have required only a few hundred thousand years. Adaptive cultural change driven by decision-making forces can be very fast indeed as is evidenced by the spread of innovations (Rogers 1983). It is not immediately clear how very short time-scale processes such as these can give rise to longer term change of the kind observed in both fossil and archaeological record unless the pace of change is regulated by environmental change. In the absence of continuing, long-term, nonstationary environmental change, adaptive processes seem quite capable of reaching equilibria in relatively
short order. In other words, both cultural and organic evolution seem, at first glance, to be classic scientific processes that produce functional adjustments too rapidly to account for the slow historical trajectories we actually observe.

Second, it is not obvious why adaptive processes should be sensitive to initial conditions. Within anthropology the view that adaptive processes are ahistorical in this sense underpins many critiques of functionalism. Many anthropologists claim that it is self-evident that cultural evolution is historical, and that, therefore, adaptive explanations (being intrinsically equilibrist and ahistorical) must be wrong. For example, Hallpike (1986) presents a variety of data which show that peoples living in similar environments often have quite different social organization, and historically related cultures often retain similar social organizations despite occupying radically different environments. Because functionalist models predict a one-to-one relationship between environment and social organization, he argues, these data falsify the functionalist view. Indeed, functionalists like Cohen (1974:86) expect to see history manifest only in the case of functionally equivalent symbolic forms (see below). Biologists have generally been more aware that a population’s response to selection depends on phylogenetic and developmental constraints and, therefore, that evolutionary trajectories are, at least to a degree, path dependent. Nonetheless, lack of convergence is sometimes used to argue to the lack of importance of natural selection. Should selection not cause populations exposed to similar environments to converge on similar adaptations? Certainly, some striking convergences from unlikely ancestors do exist.

Here we argue that path dependence and long-term change are likely to be consequences of any adaptive process analogous to natural selection. Our claims are rather general, and are thus independent of the nature of the transmission process (genetic or cultural), and of the details of development. Let us begin with an especially simple example of genetic evolution. Consider a large population of organisms in which individuals’ phenotypes can be represented as a number of quantitative characters. Let us assume that there are no constraints on what can evolve due to properties of the genetic system itself. One model with this property assumes that the distribution of additive genetic values for each character is Gaussian, that there are no genetic
Figure 1. This figure shows two adaptive topographies. The axes are the mean genetic value in a population for two characters. The contour lines give contours of equal mean fitness. Populations beginning at different initial states all achieve the same equilibrium state. Fig. 1a, above, shows a simple unimodal adaptive topography. Fig. 1b, facing page, shows a complex, multimodal topography. Initially similar populations diverge owing only to the influence of selection.

correlations among characters, that no genotype-environment interactions exist, and that mutation maintains a constant amount of heritable variation for each character. Further, assume that the fitness of each individual depends only on its own phenotype, not on the frequency of other phenotypes or the population density, and there is no environmental change. With these assumptions it can be shown that the change in the vector of mean values for each character is along the gradient of the logarithm of average fitness (Lande 1979). In other words, the mean phenotype in the population changes in the direction that maximizes the increase in the average fitness of the population. This is the sort of situation in which selection, and similar processes in the cultural system, ought to produce optimal adaptations in the straightforward manner depicted in elementary textbooks.
In this simple model the evolutionary trajectory of the population will be completely governed by the shape of average fitness as a function of mean phenotype. If the adaptive topography has a unique maximum then every population will evolve to the same equilibrium mean phenotype, independent of its starting position, and once there be maintained by stabilizing selection. On the other hand if there is more than one local maximum, different equilibrium outcomes are possible depending on initial condition. The larger the
number of local maxima, the more path dependent the resulting trajectories will be (see fig. 1).

Unfortunately, we do not know what real adaptive topographies look like, and, as Lande (1986) has pointed out, there is little chance that we will be able to determine their shape empirically. In evolutionary texts, adaptive topographies are commonly depicted as a smooth three-dimensional surface with a small number of local maxima. However, if evolutionary "design problems" are similar to the engineering ones, this picture is misleading. Experience with engineering design problems suggests that real adaptive topographies are often extremely complex, with long ridges, multiple saddle points, and many local optima — more akin to the topographic map of a real mountain range than the smooth textbook surfaces.

A computer design problem discussed by Kirkpatrick et al. (1983) provides an excellent example. Computers are constructed from large numbers of interconnected circuits each with some logical function. Because the size of chips is limited, circuits must be divided among different chips. Because signals between chips travel more slowly and require more power than signals within chips, designers want to apportion circuits among chips so as to minimize the number of connections between them. For even moderate numbers of circuits, there is an astronomical number of solutions to this problem. Kirkpatrick et al. (1983) present an example in which the 5000 circuits which make up the IBM 370 microprocessor were to be divided between two chips. Here there are about $10^{1503}$ possible solutions! This design problem has two important qualitative properties:

A. It has a very large number of local optima. That is, there is a large number of arrangements of circuits with the property that any simple rearrangement increases the number of connections between chips. This means that any search process that simply goes uphill (like our model of genetic evolution) can end up at any one of a very large number of configurations. An unsophisticated optimizing scheme will improve the design only until it reaches one of the many local optima, which one depending upon starting conditions. For example, for the 370 design problem, several runs of a simple hill-climbing algorithm produced between 677 and 730 interconnections. The best design found (using a more sophisticated algorithm) required only 183 connections.
B. There is a smaller, although still substantial, number of arrangements with close to the optimal number of interconnections. That is, there are many qualitatively different designs that have close to the best payoff. In the numerical example discussed above there are on the order of $\sqrt{5000} \approx 70$ such arrangements.

Kirkpatrick et al. (1983) show that two other computer design problems, the arrangement of chips on circuit boards, and the routing of wiring among chips, have similar properties. These three computer design problems are not unlike evolutionary "design" problems in biology— the localization of functions in organs, the arrangement of organs in a body, and the routing of the nervous and circulatory networks— that are likely to generate complex adaptive topographies. Moreover, as anyone experienced with the numerical solution of real-world optimization problems will testify, these results are quite typical. To quote from the introduction of a recent textbook on optimization "...many common design problems, from reservoirs to refrigerators, have multiple local optima, as well as false optima, that make conventional [meaning iterative hill-climbing] optimization schemes risky" (Wilde 1978). Thus, if the analogy is correct, small differences in initial conditions may launch different populations on different evolutionary trajectories which end with qualitatively different equilibrium phenotypes.

It is important to see that this history-generating property does not depend on the existence of genetic or developmental constraints. At least as defined in Maynard Smith et al. (1985) there are no genetic or developmental constraints in the simple model of selection acting on a complex topography. Every combination of phenotypes can be achieved, and there is no bias in the production of genetic variation. Path dependence results from the facts that different characters interact in a complex way to generate fitness, and that the direction of natural selection depends on the shape of the local topography.

Of course, developmental constraints could also play a major role in confining lineages to historically determined bauplans, as many biologists have argued (e.g., Seilacher 1970). Further, complex topographies and developmental constraints may be related. Wagner (1988) hypothesizes, based on a model of multivariate phenotypic evolution, that fitness functions will generally be "malignant," and that developmental constraints act to make pheno-
types more responsive to selection. By malignant Wagner means that the fitness of any one trait is likely to depend on the values of many other traits. For example, larger size may be favored by selection for success in contests for mates, but only if many traits of the respiratory, skeletal, and circulatory systems change in concert to support larger size. If phenotype is unconstrained, response to selection will be slow because of the need to change so many independent characters at once, whereas developmental constraints confine the expression of variation to a few axes which can respond rapidly to selection. Thus, the bill is a simple, rather constrained part of the anatomy of birds, yet selection has remodeled bills along the relatively few dimensions available (length, width, depth, curvature) to support an amazing variety of specializations. Developmental constraints may be a solution to the complexity of adaptive topographies, albeit one that limits lineages to elaborating a small set of historically derived basic traits as they respond to new adaptive challenges.

Path dependence can arise from the action of functional processes in a cultural system of inheritance as well. For example, decision-making forces arise when people modify culturally acquired beliefs in the attempt to satisfy some goal. If people within a culture share the same goal, this process will produce an evolutionary trajectory very similar to one produced by natural selection – the rate of change of the distribution of beliefs in a population will depend on the amount of cultural variation and the shape of an analog of the adaptive topography in which fitness is replaced by utility (the extent to which alternative beliefs satisfy the goal) (Boyd and Richerson 1985:chap. 5). The details of the transmission and selective processes are not crucial, as long as the processes that lead to change can be represented as climbing a complex topography.

It is unclear whether adding genetic constraints will increase or decrease the potential for path dependence. One sort of genetic constraint can be added by allowing significant genetic correlations among characters (Lande 1986). This assumption means that some mutants are more probable than others. As long as there is some genetic variation in each dimension, the vector of phenotypic means will still go uphill, but not necessarily in the steepest direction. The population will come to equilibrium at one of the local peaks, although this might be quite distant from the equilibrium that
the population would have reached had there been no genetic correlations (Lande 1979, 1986). More generally, most genetic architectures do not result in Gaussian distributions of genetic values (Turelli and Barton, in press), and analyses of two locus models suggest that dynamics resulting from the combination of linkage and selection may create many locally stable equilibria even when the fitness function is unimodal (Karlin and Feldman 1970). This suggests that adding more genetic realism would increase the potential for path dependence. On the other hand, computer scientists (Holland 1975; Brady 1985) have found that optimization algorithms closely modeled on multilocus selection are less likely to get stuck on local optima than simple iterative hill-climbing algorithms. The issue of genetic constraints is still open.

The situation in cultural evolution is similar, even if not so well studied. On the one hand, many anthropologists stress the rich structure of culture. To the extent that such structure exists, path dependence is likely to be important. On the other hand, Bandura (1977), a pioneering student of the processes of social learning, argues that there is relatively little complex structuring of socially learned behavior. The many examples of cultural syncretism and diffusion of isolated elements of technology suggest his view ought to be taken seriously. Perhaps complex structure is most important in the symbolic aspects of culture, but symbolic variation may be only weakly constrained by functional considerations (Cohen 1974). According to Cohen, we have to use purely contingent historical explanations for things such as linguistic variation, while simple functional explanations suffice for economic, political, and social-organizational phenomena.

Long-term nonstationary trends in evolution can result if there is some process that causes populations to shift from one peak to another, and if that process acts on a longer time scale than adaptive processes like natural selection. So far we have assumed that populations are large and the environment is unchanging. With these assumptions, populations will usually rapidly reach an adaptive peak and then stay there indefinitely. They will not exhibit the kind of long-run change that we have required for change to be historical. Wright (e.g., 1977) long argued that drift plays an important role in causing populations to shift from peak to peak, and then competition among populations favoring the population on the higher peak. Chance
variations in gene frequency in small populations could lead to the occasional crossing of adaptive valleys and the movement to higher peaks. Recently, several authors have considered mathematical models of this process (Barton and Charlesworth 1984; Newman et al. 1985; Lande 1986; Crow et al. 1990). These studies suggest that the probability that a shift to a new peak will occur during any time period is low; however, when a shift does occur it occurs very rapidly. If this view is correct, drift should generate a long-run pattern of change in which populations wander haltingly up the adaptive topography from lower local peaks to higher ones. It is also implausible that environments remain constant either in space or in time. As environments change, the shape of the adaptive topography shifts, causing peaks to merge, split, disappear, or temporary ridges to appear, connecting a lower peak to a higher one. Thus, populations will occasionally slide from one peak to another. As long as such events are not too common, environmental change will also lead to long-run change. Such change might appear gradual if there are many small valleys to cross, or punctuational if there are a few big ones.

Adding social or ecological realism to the basic adaptive hill-climbing model of evolution probably increases the potential for multiple stable equilibria. In the simple model, an individual's fitness depended only on its phenotype. When there are social or ecological interactions among individuals within a population, individual fitness will depend on the composition of the population as a whole. When this is the case, evolutionary dynamics can no longer be represented in terms of an invariant adaptive topography. However, they may still be characterized by multiple stable equilibria. Moreover, the fact that many quite simple models of frequency dependence have this property suggests that frequency dependence may usually increase the potential for path-dependent historical change.

Models of the evolution of norms provide an interesting example of how frequency dependence can multiply the number of stable equilibria. Hirshleifer and Rasmusen (1989) have analyzed a model in which a group of individuals interact over a period of time. During each interaction, individuals first have the opportunity to cooperate and thereby produce a benefit to the group as a whole but at some cost to themselves; they then have a chance to punish defectors at no cost to themselves. These authors show that strategies in which individuals cooperate, and punish noncooperators
and nonpunishers, are stable in the game theoretic sense. However, they also show that punishment strategies of this kind can stabilize any behavior—cooperation, noncooperation, wearing white socks, or anything else. We (Boyd and Richerson, in prep.) show that the same conclusions apply in an evolutionary model even when punishment is costly. This form of social norm can stabilize virtually any form of behavior as long as the fitness cost of the behavior is small compared to the costs of being punished.

More generally, coordination is an important aspect of several kinds of social interactions (Sugden 1986). In a pure game of coordination it does not matter what strategy is used, as long as it is the strategy that is locally common. Driving on the left versus right side of the road is an example. It does not matter which side we use, but it is critical that we agree on one side or the other. This property of arbitrary advantage to the common strategy is shared by many symbolic and communication systems, and allows multiple equilibria whenever there are multiple conceivable strategies. In many other common kinds of social interactions elements of coordination and conflict are mixed. In such games, all individuals are better off if they use the same strategy, even though the relative advantages of using the strategy differ greatly from individual to individual, and some individuals would be much better off if another strategy were common. As long as the coordination aspect of such interactions is strong enough, multiple stable equilibria will exist. Arthur (1990) shows how locational decisions of industrial enterprises could give rise to historical patterns due to coordination effects. It is often advantageous for firms to locate near other firms in the same industry because specialized labor and suppliers have been attracted by preexisting firms. The chance decisions of the first few firms in an emerging industry can establish one as opposed to another area as the Silicon Valley of that industry. More generally, historical patterns can arise in the many situations where there are increasing returns to scale in the production of a given product or technology. Merely because the QWERTY keyboard is common, it is sensible to adopt it despite its inefficiencies.

Interactions between populations and societies (or elements within societies such as classes) can give rise to multiple stable equilibria. Models of the coevolution of multiple populations have many of the same properties as frequency and density dependent selection within populations, although
the theory is less well developed (Slatkin and Maynard Smith 1979). The evolution of one population or society depends upon the properties of others that interact with it, and many different systems of adjusting the relationships between the populations may be possible. For example, Cody (1974:201) noted that competing birds replace each other along an altitudinal gradient in California, but latitudinally in Chile. Given the rather similar environments of these two places, it is plausible that both systems of competitive replacement are stable and which one occurs is due to accidents of history.

The stratification of human societies into privileged elites and disadvantaged commoners derives from the ability of elites to control high-quality resources and/or to exploit commoners using strategies that are similar to competitive and predatory strategies in nature. Insko et al. (1983) studied the evolution of social stratification in the social psychology laboratory. They showed that elites could arise in both an experimental condition that mimicked freely chosen trade relations and one that mimicked conquest. Elites were approximately as well off in both conditions, and insofar as they controlled things, would have no motivation to change social arrangements. It seems plausible that the diversity of political forms of complex societies could result from many arrangements of relations between constituent interest groups being locally stable. The distinctive differences between the Japanese, American, and Scandinavian strategies for operating technologically advanced societies could well derive from historic differences in social organization that have led to different, stable arrangements between interest groups, in spite of similar revolutionary changes in production techniques of the last century or two.

Social or ecological interactions may also give rise to dynamic processes that are sensitive to initial conditions, and have no stable equilibria. Lande (1981) analyzed a model of one such process, sexual selection in which females have a heritable preference for mates that is based on a heritable, sex-limited male character. According to his model, when the male character and female preferences are sufficiently correlated genetically, female choice can create a self-reinforcing "runaway" process that causes the mean male character and the mean female preference to either increase or decrease indefinitely, even in the presence of stabilizing selection on the male character. Selection cannot favor female variants that choose fitter
males (in the usual sense of fitter) because most females are choosing mates with an exaggerated character. The "sensible" female's sons will be handicapped in the mating game. The direction that evolution takes depends on the details of the initial conditions in Lande's model. His quantitative character will be elaborated in one direction or the other depending on how evolution drifts away from an unstable line of equilibria. Although the interpretation of this model is controversial, it is easy to imagine that the exaggerated characters of polygynous animals like birds of paradise and peacocks
result from the runaway process. We (Boyd and Richerson 1985:chap. 8, 1987; Richerson and Boyd 1989b) have argued that quite similar processes may arise in cultural evolution when individuals are predisposed to imitate some individuals on the basis of culturally heritable characteristics. The use of some character associated with prestige (stylish dress for example) as an index of whom to imitate has the same potentially unstable runaway dynamic as Lande’s model of mate choice sexual selection, and even casual observation suggests that prestige systems do follow contingent historical trajectories. Fashions in clothing, for example, evolve in different directions in different societies, often without much regard for practicality.

Perhaps the most clearly historical patterns of change result when social or ecological interaction lead to "chaotic" dynamics. For example, Day and Walter (1989) have analyzed an extremely interesting model of social evolution in which population growth leads to reduced productivity, social stratification, and eventually to a shift from one subsistence technology to a more productive one. The resulting trajectories of population size are shown in figure 2. Population grows, is limited by resource constraints, and eventually technical substitution occurs, allowing population to grow once more. The only difference between figures 2a and 2b is a very small difference in initial population size. Nonetheless, this seemingly insignificant difference leads to qualitatively different trajectories – one society shows three separate evolutionary stages, and the second only two.

Conclusion

Scientific and historical explanations are not alternatives. Contingent, diverging pathways of evolution and long-term secular trends can result from processes that differ only slightly from those that produce rapid, ahistorical convergence to universal equilibria. Late nineteenth and early twentieth century scientists gave up restricting the term "scientific" for deterministic, mechanistic explanation and began to admit "merely" statistical laws into the fundamental corpus of physics (very reluctantly in some cases – recall Einstein's famous complaint about God not playing dice with the universe to express his distaste for quantum mechanics). Similarly, historical explanations cannot
be distinguished from other kinds of scientific explanations except that some models (and, presumably, the phenomena they represent) generate trajectories that meet our definition of being historical. These history-generating processes do not depend on exotic forces or immaterial causes that ought to excite a scientist’s skepticism; perfectly mundane things will do. There are challenging complexities in historical processes. For example, even well-understood processes will not allow precise predictions of future behavior when change is historical. However, all the tools of conventional scientific methods can be brought to bear on them. For example, it should be possible to use measurement or experiment to determine if a process is in a region of parameter values where chaotic behavior is expected. At the same time, the historian’s traditional concern for critically dissecting the contingencies that contribute to each unique historical path is well taken. Process-oriented "scientific" analyses help us understand how history works, and "historical" data are essential to test scientific hypotheses about how populations and societies change.

In the biological and social domains, "science" without "history" leaves many interesting phenomena unexplained, while "history" without "science" cannot produce an explanatory account of the past, only a listing of disconnected facts. The generalizing impulses of science require historical methods, because the phenomena to be understood are genuinely historical and because historical data are essential for developing generalizations about evolutionary processes. In return, generalizations derived from history and by the study of contemporary systems would seem to be essential for an understanding of particular cases. The amount of data available from the past is usually very limited, and the number of possible reconstructions of the past is correspondingly large. Some sort of theory has to be applied to make some sense of the of the isolated facts. Historians (e.g., Braudel 1972) and paleontologists (e.g., Valentine 1973) often cast their nets rather widely in search of help in interpreting the documents and fossils. McNeill (1986) advocates a "scientific," generalization-seeking approach to history much in this spirit. Consider the question of which of the potential history-producing processes we have discussed are most important in explaining the changes in human societies over the last few tens of thousands of years. Generalizing disciplines such as climatology and cultural ecology are certainly relevant to
the task in general, and to the understanding of how particular societies changed in particular environments (Henry 1989). At the same time, because these historical societies faced Pleistocene climates and the transition to the Holocene, and because they developed a series of technical, social, and ideological innovations that are the foundation of modern human societies by processes that are not open to direct observation, the historical and archaeological record provide crucial data not available from ahistorical study. To the extent that the processes we have described are important, "science" and "history" cannot be disentangled as separate intellectual enterprises.

Darwinian models of organic and cultural evolution illustrate how little distinction can be made between the two approaches. Such models can produce historical patterns of change by a rather large number of different mechanisms. We have argued that historical change is distinguished by two attributes: the tendency of initially similar systems to diverge, and the occurrence of long-term change. Evolutionary models, including those which assume that selection or analogous cultural processes increase adaptiveness in each generation, readily generate multiple stable equilibria. Populations with similar initial conditions may evolve toward separate equilibria. Random genetic drift and analogous cultural processes coupled with environmental change may cause populations to shift from one equilibrium to another. It is plausible that peak shifting by populations (or the shifting of peaks due to environmental change) occurs at a slow enough rate to explain long-term secular trends.

Many anthropologists take as their task the explanation of differences among human societies, and suggest that most such differences are historical in character. If explanation of such variation is mainly historical, then anthropologists might reasonably ask, what is the point of Darwinian models of cultural change when historical or "contextual" explanations will be much more productive. The reasons are as follows:

First, the premise is often incorrect. Genuine convergences are common, and explaining them requires some theory based on common processes of cultural change. Perhaps the most spectacular cultural example is the convergence of social organization in stratified, state-level societies in the Old and New Worlds. For example, Cortez in 1519 found that Aztec society
was quite similar to his own in important ways: it contained familiar roles, hereditary nobility, priests, warriors, craftsmen, peasants and so on. The bureaucracy was organized hierarchically. This convergence is remarkable because the Spanish and Aztec states evolved independently from a hunter-gatherer ancestry. The cultural lineages that resulted in these two states were without known cultural contact for several thousand years before state formation began in either (Wenke 1980).

Second, Darwinian models can make useful predictions. They can tell us why some forms of behavior or social organization are never observed and others are common. For example, kinship is an extremely common principle of social organization. Contrariwise, there would seem to be lots of advantages to a free market in babies — for the individual it would allow easy adjustment of family size, age composition, sex ratio and so on, and for society a division of labor in child rearing would allow better use of human resources. The sociobiological theory of kin selection explains why there are no societies with free trade in infants, and why kinship is generally an important feature of social organization. If most of the historic context is taken as given, Darwinian arguments can be very powerful heuristics. This is especially clear for genetic evolution. For example, given haplodiploidy, a theory based directly on the expected equilibrium outcome of natural selection can make surprising and extremely fruitful predictions about patterns of behavior in social insects. Who, for example, would have thought to connect sex ratio among reproductives and "slave making" in ant species? In recent years, similar ideas have been usefully applied to understanding human behavior. For example, Hill, Kaplan and their colleagues (reviewed in Hill and Kaplan 1988) have used theory from behavioral ecology to relate patterns of foraging, mate preference, and child care among Ache hunter-gatherers, and Boguehoff-Mulder (1988) has explained variation in bride price among Kipsigis pastoralists in terms of parameters that predict future female fitness.

Finally, it is useful in and of itself to know that even the most strongly functional Darwinian models can give rise to historical change. The same processes that give rise to convergence in one case can generate differences in another given only small changes in the structure of the process or in initial conditions. Brandon (1990) argues that "why possibly" explanations are useful in evolutionary biology. By this he means explanations that tell us
how some character could have evolved are useful even if we cannot determine whether the explanation is true. The theoretical models in population genetics provide a good example: Hamilton's (1964) kin selection models show how natural selection could give rise to self-sacrificial behavior. However, we usually do not know whether any particular case of altruism arose as a result of kin selection. The lack of any "why possibly" explanation would cast doubt on other aspects of our knowledge of how selection shapes behavior.

Understanding how adaptive processes could give rise to historical change is useful for analogous reasons. There is considerable evidence that people's choices about what to believe and what to value are affected by the consequences in material well being, social status and so on (e.g., Boyd and Richerson 1985). This view has a venerable history in anthropology (e.g., Barth 1981; Harris 1979), plays a foundation role in economics, and is taken for granted in many historians' explanations for particular sequences of events. If cultural change is affected by consequence-driven individual choice or natural selection, then it follows that there will be a process that will act to modify the distribution of cultural variation in a population in much the same way that natural selection changes genetic variation (Boyd and Richerson 1985:chaps. 4 and 5). The fact that functional processes like natural selection readily lead to history allows one to hold this view without having necessarily to search for external environmental differences to explain the differences among apparently similarly situated human societies.

Acknowledgments

We thank James Griesemer, Matthew H. Nitecki, Eric A. Smith, and two anonymous reviewers for most helpful comments on previous drafts of this paper.

Notes

1. This project is quite different from the better-known, classical studies of cultural evolution developed by Leslie White (1959) and other scholars in anthropology. This work focused de-
scriptively on the large scale patterns of cultural evolution rather than on the details of the processes by which cultural evolution occurs (Campbell 1965, 1975). The research tradition White represents derives from the progressivist ideas of Herbert Spencer, rather than from Darwin.

2. The additive genetic value of a particular individual for a particular character is the average value of that character for offspring produced when that individual mates at random with a large number of other individuals in the population. For example, the additive genetic value of a bull for fat content is the average fat content of all its offspring where mates were chosen at random. The distribution of genetic values is Gaussian when the probability that an individual has a given genetic value is given by the normal (or Gaussian) probability distribution. Genetic correlations exist when the distributions of genetic values for different characters are not probabilistically independent. For example, if bulls whose genetic value for size also tend to have a higher genetic value for fat content, then body size and fat content are genetically correlated. Genotype environment correlations arise when individuals with the same genotype develop different phenotypes in different environments.

References


Henry, D. O. 1989. *From Foraging to Agriculture: The Levant at the End of the Ice Age*. Phila-


