EVOLUTION AND COGNITION

Why Humans Cooperate

A CULTURAL AND EVOLUTIONARY EXPLANATION

Natalie Henrich
Joseph Henrich

Contents

Abbreviations

CC	Cooperation with Chaldeans
CNC	Cooperation with non-Chaldeans
DG	Dictator Game
EI	Ethnic Index
IMO	Income Maximizing Offer
MU	Monetary Unit
MW	Mann-Whitney test
PD	Prisoner’s Dilemma
PGG	Public Good’s Game
SMUG	Strategy Method Ultimatum Game
TFT	Tit-for-tat
UG	Ultimatum Game

WHY HUMANS COOPERATE

1
Evolution, Culture, Cooperation, and the Chaldeans

> A man who was not impelled by any deep, instinctive feeling, to sacrifice his life for the good of others, yet was roused to such actions by a sense of glory, would by his example excite the same wish for glory in other men, and would strengthen by exercise the noble feeling of admiration. He might thus do far more good to his tribe than by begetting offspring with a tendency to inherit his own high character.
>
> —Charles Darwin, "On the Development of the Intellectual and Moral Faculties," chap. 5 of *The Descent of Man*

This book is an inquiry into one of the great puzzles in the human sciences: the evolution of cooperation and altruism in the human species. Unlike many works on this question, we seek to simultaneously draw together formal theoretical work on the evolution of cooperation, rich ethnographic descriptions of human social life, and a wide range of experimental results from both the laboratory and the field. On the theoretical front, we provide an introduction to the puzzle of human cooperation and a unified theoretical framework that integrates culture, psychology, and evolution in a manner that makes these concepts accessible to nonspecialists. From this general framework a set of theoretical foci emerge to provide a backbone for the book and place this work in a context that will evoke a sense of familiarity for those who have studied human societies. These foci are (1) kinship, (2) reciprocity and reputation, (3) social norms, and (4) ethnicity. Though we have attempted to present this theoretical material at an introductory level, we believe that more advanced students and scholars will find our synthesis of culture, coevolution, and cooperation worthwhile and provocative.

Breathing life into the deductive logic and equations that buttress our theoretical presentation, we develop a body of ethnographic material on social life from 18 months of study by Natalie Smith Henrich among the Chaldeans of metro Detroit, as they lived, worked, and socialized at the end of the twentieth century. The community studied in this book is made up of first-, second-, and third-generation Catholic immigrants from Iraq who are predominantly middle- and upper-class inhabitants of Southfield, Michigan. As a focused, problem-driven ethnography, this book uses the lens of evolutionary theory to examine the anatomy of cooperation in Chaldean life. Though we make no attempt to comprehensively explore their lifeways, readers will learn a fair amount about contemporary Chaldean life in metro Detroit in working their way through the book. After laying the theoretical foundations of our inquiry in chapters 2 and 3, we open the ethnographic stream in chapter 4 by providing a brief introduction to Chaldean culture and history, with an eye on issues relevant to our work among the Michigan population.

In general, we feel that the use of both experimental and ethnographic methods is crucial for advancing theory. Experiments allow for precise measures, controlled comparisons, and specific manipulations, which often permit the direct testing of alternative theories. However, it is often unclear whether what is measured and manipulated in experiments actually affects decisions and behavior in real life. In contrast, ethnography, although messier than experiments and only occasionally amenable to controlled tests, allows us to systematically look for a consistency between real life and theory. Experiments on decision-making, for example, allow students using computers in a university laboratory to play the role of hypothetical farmers. Given full information on their screens about prices, yields, and labor costs related to their "crop choices," students make systematic trade-offs among the three variables in deciding "what to plant." Ethnography, however, has shown that many real farmers don't have any idea what the yields and prices are of the crops they don't plant—thus, they cannot be making the trade-offs seen in the laboratory. The experiments are missing the dynamics driving real behavior and decision making (Henrich 2002). To address the inadequacy of relying entirely on either experiments or ethnography, we have used experiments throughout this book to test one theoretical idea or another, and then employed ethnographic data to examine the consistency between laboratory-tested theories and real life.

Specifically, we have integrated experimental data with our ethnographic work in two ways. First, at the beginning of each of the relevant chapters, we summarize the most important experimental findings—predominately from behavioral economics—vis-à-vis the pertinent evolutionary theories. Second, we report results from our own use of experimental tools among the Chaldeans. As we show in chapters 5–9, and have already shown elsewhere (Henrich et al. 2004; Henrich and Smith 2004), the combination of experimental and ethnographic data can be more powerful than either is alone. Thus, in addition to the findings relevant to our focus herein, we hope to persuade experimentalists to read more ethnography and history, and more ethnographers to equip themselves with the rich set of experimental tools available in economics and psychology.

In our emphasis on culture and culture-gene coevolution, this book is unlike most other works in the field of evolution and human behavior. Broadly following the research programs laid out by Charles Darwin (1981) and James Mark Baldwin (1896a; 1896b; 1968), our theoretical approach brings “culture” (as socially learned behaviors, beliefs, values, etc.) into evolutionary theory (Richerson and Boyd 2005). In contrast to noncultural evolutionary approaches, such as those typically found under the rubric of “evolutionary psychology” (Tooby and Cosmides 1992), our framework adds two interrelated elements that we believe are crucial for understanding humans. First, we draw ideas from a large and growing body of evolutionary theory that views our “evolved minds” as a result (at least in part) of the coevolutionary interaction between genes and culture (Boyd and Richerson 1985; Henrich 2004; Henrich and McElreath 2003; Laland 2000; A. Rogers 1989). Given the massive reliance that humans place on social learning (i.e., culture: Tomasello 1999; Tomasello 2000) and the tendency for these socially transmitted behaviors to alter both our local physical and social environments (Durham 1991), it is quite difficult to imagine how cultural transmission could not have affected genetic evolution.[1] Second, the unique nature of human cultural transmission creates chains of learning that operate over successive generations. That is, our cultural capacity gives rise to cultural evolution and suggests that some important aspects of human behavior cannot be understood without considering the cultural evolutionary history of the particular social group in question.

In the coming chapters, we argue that both elements of this dual inheritance approach, culture-gene coevolution and culture history, are of great importance for understanding cooperation in our species and can be effectively incorporated under the umbrella of Darwinian theory (Boyd and Richerson 2002; Henrich et al. 2003; Richerson and Boyd 1998; Richerson, Boyd, and Henrich 2003). Thus, our theoretical perspective allows researchers to consider and develop explanations that involve both genetically evolved psychologies and cultural transmission (and culturally evolved psychologies: Levinson 2003; Nisbett 2003) without manifesting the very popular, false, and intellectually destructive dichotomy between evolutionary and cultural explanations. As we hope to show with regard to the Chaldeans, fully incorporating the Darwin-Baldwin vision of culture-gene coevolution expands our abilities to explain human behavior without losing the critical linkages between humans and the rest of the natural world that evolutionary theory provides.

As an urban ethnography, and especially one that is driven by evolutionary theory, this book is either unique or rare. Evolutionary anthropologists have often been accused of focusing their research on "small-scale," "tribal," or "primitive" human social groups at the exclusion of people in industrialized societies. Moreover, when anthropologists do study populations in an urban context, we tend to either study impoverished groups, marginalized groups, or university students. Here, we offer an urban ethnography of a fairly wealthy ethnic group in a major U.S. city. Thus, we demonstrate that evolutionary theory is as useful in explaining the behavior of middle- and upper-class Americans in an urban environment as it is to understanding people living in small-scale societies.

The Organization of the Book

This book is organized as follows: chapter 2 sketches the broad theoretical background by introducing culture-gene coevolutionary theory, explaining how culture and cultural evolution can be incorporated into evolutionary theory, and detailing a few key aspects of human cultural learning. Chapter 3 narrows our theoretical focus and charts the course for the rest of the book by applying both evolutionary and coevolutionary theory to the puzzle of human cooperation. Chapter 4 sets the scene for our ethnographic discussions by introducing the Chaldeans, their history, some ethnographic background on their life in metro Detroit, and a description of the ethnographic field methods used. Chapters 5 through 9 further develops each of the major classes of theories of cooperation introduced in chapter 2 by supplying further discussion, reviewing relevant experimental work, and examining each vis-à-vis Chaldean social life and cooperation. Specifically, chapter 5 deals with kinship, chapter 6 with reciprocity and reputation, chapters 7 and 8 with social norms and punishment, and chapter 9 with ethnicity. Chapter 10 concludes by considering the cooperative dilemmas associated with some of the environmental, economic, and public health problems that currently confront our society, both locally and globally. The appendices provide more in-depth background material that may interest some readers and are referenced at various points through the book.[2]

2
Dual Inheritance Theory
The Evolution of Cultural Capacities and Cultural Evolution

> The tendency to imitate may come into direct conflict with the prudential teachings of pleasure and pain, and yet may be acted upon. A child may do, and keep on doing, imitations which cause him pain.
>
> —James Mark Baldwin, chap. 10 of *Mental Development in the Child and the Race*

> With the exception of the instinct of self-preservation, the propensity for emulation is probably the strongest and most alert and persistent of the economic motives proper.
>
> —Thorstein Veblen, chap. 5 of *The Theory of the Leisure Class*

Since the rise of human sociobiology in the 1970s, culture and biology or cultural explanations and evolutionary explanations have often been opposed, and the seeming opposition between the categories has led to a great deal of unnecessary dispute and debate. This dichotomy and its associated debates are now outmoded and unproductive. A wide range of human behaviors that most people would think of as purely cultural (dress, greetings, food taboos, etc.) are actually 100 percent cultural and 100 percent genetic. Many behaviors are cultural in that they are socially learned by observation and interaction in a social group—social learning can then be understood as the foundational capacity that underpins what is typically glossed as "culture." All culturally acquired behaviors, beliefs, preferences, strategies, and practices (hereafter, we refer to all these collectively as "cultural traits") are also genetic in the sense that their acquisition requires brain machinery that allows for substantial amounts of complex, high-fidelity social learning. We know that there must be "human genes" that somehow allow for culturally acquired behavior, as chimpanzees reared (enculturated) alongside human children do not acquire anything approaching adult human behavioral patterns or social norms. Interestingly, although human-reared chimpanzees seem to acquire little from their human families via imitation, these families' human children have been observed to readily acquire a number of behaviors from their physically more advanced chimpanzee "siblings," including knuckle walking (even after achieving full bipedality), shoe-chewing, a habit of scraping their teeth against interior walls, excessive biting, and a range of stereotypical chimpanzee food grunts and hoots.[1] More generally, although limited social learning abilities are found elsewhere in nature, social learning in our species is high fidelity, frequent, internally motivated, often unconscious, and broadly applicable, with humans learning everything from motor patterns to goals and affective responses, in domains ranging from toolmaking and food preferences to altruism and spatial cognition. If other animals are "cultural," then we are a hypercultural species (Henrich 2003).

Capacities for social learning—and culture—quite naturally emerge from a Darwinian framework once one sees the dualistic opposition between culture and biology for what it is. In their now classic treatise, *Culture and the Evolutionary Process*, Boyd and Richerson (1985) formally apply the logic of natural selection to the evolution of our capacities for social learning (culture). This work, and many subsequent papers, showed that social learning could be effectively understood as a genetically evolved adaptation for acquiring (learning) adaptive traits in complex, spatially variable, and fluctuating environments. To see this, imagine you were born into a band of hunter-gatherers and needed to figure out what to eat. You could wander around the environment sampling various potential foods until you found some you liked (assume liking is a cue for "good to eat"). This, however, risks eating something poisonous and will certainly waste a lot of time (which could be devoted to other activities), as you are most likely to encounter things that are either nonnutritious or indigestible. A social learner can avoid all of this by simply focusing on the healthiest members of his group and eating whatever they eat. This cultural approach obtains a pretty good diet and avoids the costs of searching, sampling, and possibly eating something poisonous. Similarly, imagine you are again growing up among foragers and need to figure out how to hunt. If you could not learn via imitation, you would have to individually figure out a huge amount of information about animal behavior, tracking, hunting skills, and how to manufacture the necessary technology. Simply making a bow and arrow, which often requires a poison for the arrows to be effective, demands complex skills and knowledge. One would need to figure out, among other things, which trees can be used for bows; how to strengthen and straighten the wood; which insects, larvae, or tree saps can be used for making a poison; and which animal sinews are best for bow strings—all things one is rather unlikely to discover by himself. A good imitator, however, can simply observe and learn from the other members of his group, thereby taking advantage of the accumulated experience and wisdom of previous generations. If people mostly learn from the previous generation but occasionally make additions and improvements through their own experience, experiments, or luck, culture can become an adaptive system of learned traits that accumulate through time. The result of this kind of cultural transmission is a second system of inheritance that, for as long as human ancestors have had sophisticated social learning, has run in parallel —and often intertwined—with our genetic inheritance system.

With the emergence of this cumulative cultural learning it is useful to distinguish three levels of explanation: (1) Ultimate level: natural selection builds the psychological capacities for cultural learning; (2) Intermediate level: culture evolves, accumulates, and adapts nongenetically to produce local skills, preferences, beliefs, values, and cognitive abilities; (3) Proximate level: psychological mechanisms, preferences, values, beliefs, and motivations, which are the joint products of genetic and cultural evolutionary history, propel individual actions, decision-making, and behavior in particular situations.

Emerging from this line of reasoning, Dual Inheritance Theory aims to take account of the evolution of our capacities for culture (sophisticated social learning), cultural evolution itself, and the interaction—or coevolution—of culture and genes over tens of thousands of years. This approach can be summarized with three key ideas:

1. Culture, cultural transmission, and cultural evolution arise from genetically evolved psychological adaptations for acquiring ideas, beliefs, values, practices, mental models, and strategies from other individuals by observation and inference. In the next section, we summarize how evolutionary theory has been used to predict the psychological details of these cultural learning cognitive capacities.
2. These psychological mechanisms for social learning led to behaviors that were, on average, adaptive in the varying ancestral environments of the human lineage. Any particular individual's behavior or group's cultural practices may be adaptively neutral or maladaptive. By specifying some of the psychological details of these cultural learning abilities (see #1 above), cultural evolutionary models enable us to predict the patterns and conditions of maladaptation, and thus provide theories of adaptation, maladaptation, and the dynamic process of cultural change (Boyd and Richerson 1985: chap. 7; Henrich 2004b). This is an advantage over the models traditionally used in sociobiology and human behavioral ecology, in which no proximate mechanisms for achieving adaptation are specified—and thus the dynamics of change cannot be explored, and behaviors are either "adaptive" or inexplicable.
3. The emergence of cultural learning capacities in the human lineage created population processes that changed the selective environments in which genes develop. For example, suppose the practice of cooking meat spread by imitative learning in ancestral human populations. In an environment of 'cooked meat', natural selection may have favored genes that shortened our energetically costly intestines and altered our digestive chemistry. Such a reduction of digestive tissue may have freed energy for more "brain building." In this way, human biology adapted to culturally transmitted behavior. The interaction is called culture-gene coevolution. As discussed below, this interaction may be critical for understanding some aspects of human cooperation, particularly large-scale cooperation among nonrelatives (Baldwin 1896b; Boyd and Richerson 2002; Boyd and Richerson, 2005; Durham 1991; Henrich 2004a, 2004b; Richerson and Boyd 1998; Richerson and Boyd 2000).

Below, we summarize the theory and especially the evidence for certain key aspects of cultural learning, which will be important to our discussion of the evolution of cooperation. For a more complete understanding of this approach to culture and evolution, readers should begin with Henrich and McElreath (2003) or Richerson and Boyd (2005).

Evolved Psychological Mechanisms for Learning Culture

The approach to understanding culture using evolutionary theory begins by considering what kinds of cognitive learning abilities would have allowed individuals to efficiently and effectively extract adaptive ideas, beliefs, and practices from their social worlds in the changing environments of our hunter-gatherer ancestors (Richerson, Boyd, and Bettinger 2001). This approach diverges from mainstream evolutionary psychology in its emphasis on the *costly information hypothesis* and on the evolution of specialized social learning mechanisms. The costly information hypothesis focuses on the evolutionary trade-offs between acquiring accurate behavioral information at high cost and gleaning less accurate information at low cost. By formally exploring how the costly information hypothesis generates trade-offs in the evolution of our social learning capacities, we can formulate predictive theories about the details of human cultural psychology (Henrich and McElreath 2003). When acquiring information by individual learning is costly, natural selection will favor cultural learning mechanisms that allow individuals to extract adaptive information—strategies, practices, heuristics, and beliefs—from other members of their social group at a lower cost than through alternative individual mechanisms (e.g., trial-and-error learning). Human cognition probably contains numerous heuristics, directed attentional biases, and inferential tendencies that facilitate the acquisition of useful traits from other people.

Such cultural learning mechanisms can be categorized into (1) *content biases* and (2) *context biases*. Content biases, or what Boyd and Richerson (1985) have called *direct biases*, cause us to more readily acquire certain beliefs, ideas, or behaviors because some aspect of their content makes them more appealing (or more likely to be inferred from observation). For example, imagine three practices involving different additives to popcorn: the first involves putting salt on popcorn, the second favors adding sugar, and the third involves sprinkling sawdust on the kernels. Innate content biases that affect cultural transmission will guarantee that sawdust will likely not be a popular popcorn additive in any human society (sawdust is not very tasty). Both salt and sugar have innate content biases for sensible evolutionary reasons: foods with salty or sugary flavors were important sources of scarce nutrients and calories in ancestral human environments. Thus, natural selection favored a bias to acquire a taste for salty and sweet foods so that we would be motivated to obtain and eat them. Of course, if you grew up in a society that only salts its popcorn you may steadfastly adhere to your salting preference even once you find that sugar is the standard popcorn seasoning in other societies. Thus, human food preferences are simultaneously cultural (culturally learned) and innate (influenced by innate content biases).

Content biases may be either reliably developing products of our species-shared genetic heritage (i.e., innate) or culture specific. People may culturally learn beliefs, values, and/or mental models that then act as content biases for other aspects of culture. That is, having acquired a particular idea via cultural transmission, a learner may be more likely to acquire another idea because the two “fit together” in some cognitive or psychological sense.[2] For example, believing that a certain ritual in the spring will increase the crop harvest in the summer may favor the acquisition of a belief that a similar ritual will increase a woman’s odds of conception, a healthy pregnancy, and/or successfully delivering a robust infant.

Context biases, on the other hand, exploit cues from the individuals who are being learned from (we will term these individuals “models”), rather than from features of the thing being learned (the cultural trait), to guide social learning. There is a great deal of adaptive information embodied in both who holds ideas and how common the ideas or practices are. For example, because information is costly to acquire, individuals will do better if they preferentially pay attention to, and learn from, people who are highly successful, particularly skilled, and/or well respected. Social learners who selectively learn from those more likely to have adaptive skills (that lead to success) can outcompete those who do not. A large amount of mathematical modeling effort has been expended in exploring the conditions under which different context biases will evolve, how they should be constructed psychologically, and what population patterns will emerge from individuals using such learning mechanisms. Moreover, and perhaps more significant, a vast amount of field and laboratory data confirm that these learning biases are indeed an important part of our cognition (data are sketched below), that they are used by both children and adults, and that they influence economic decisions, opinions, judgments, values (e.g., altruism), eating behavior, rates and methods of suicide, and the diffusion of innovations. Our remaining discussion of psychological mechanisms focuses on the theory and evidence for two categories of context biases in cultural learning: (1) success and prestige biases, and (2) conformity biases.

Success and Prestige Biases

Once an individual is learning from others, she would be wise (in an adaptive evolutionary sense) to be selective about whom she chooses to learn from (Henrich and Gil-White 2001). A learner should use cues, attributes, or characteristics of the individuals in their social world to figure out who is most likely to have useful ideas, beliefs, values, preferences, or strategies that might be gleaned, at least partially, through observation. For example, an aspiring farmer might imitate the strategies and practices of the most skillful, successful, or prestigious farmers who live around him. Simply figuring out who obtains the biggest yields per hectare and copying that person’s practices is a lot easier than doing all the trial-and-error learning for the immense variety of decisions a farmer (or anyone else) has to make. On his own, an individual learner would have to experiment with many types of crops, seeds, fertilizers, planting schedules, and various plowing techniques. The variety of combinations creates an explosion of possibilities, making it virtually impossible for an individual to figure out the best farming strategy by relying entirely on experimentation. This is true of many real-world decisions that everyone faces. However, along with figuring out who is the most successful or most skilled, learners should also be concerned about how the things they might learn will fit with their own abilities, the expectations of their role or gender, and their personal context. Learners should assess certain kinds of "similarity" and weigh this alongside their assessments of "skill" and "success." Following this logic, we argue in the next chapter that learners should preferentially learn social norms from individuals who share their ethnic markers (e.g., their dialect, language, or dress; see McElreath, Boyd, and Richerson 2003).

Figuring out who possesses the more adaptive skills, strategies, preferences, and beliefs is no straightforward task. To achieve this, people rely on a range of cues related to skill (or competence), success, and prestige. For rhetorical purposes, this tripartite distinction is helpful because it captures the continuum of cues from direct observation by the learner (of skill or competence) to completely indirect assessments based on prestige. Noting someone's skill or competence, for our purposes, means that one has directly observed their technique or performance. An apprentice might watch two craftsmen working side by side, one hitting all of his marks and gliding right along to a perfect final product (say a handmade chair) as the other struggles, cuts himself twice, curses a bit, and produces something that only the bravest of his friends would sit on.

Cues of success are less direct and take advantage of easily observable correlates of competence (which are hopefully difficult to fake). Depending on the domain and society, such cues might be measured by house size, family size, number of wives and/or children, number of peer-reviewed publications, costliness of car, number of tapirs killed, number of heads taken in raids, size of biggest yam grown, and so on, each of which, in particular social contexts, is related to some domain of skill. Though these cues provide only an indirect measure of competence, they are sometimes more accurate than direct observations of competence. If performances are highly variable, the observations of a small sample of total performance may lead a learner to misperceive competence. However, cues of success often average over many performances, which can help reduce the error in the learner's assessment of whom to learn from.

The evolutionary theory underpinning this form of model-based cultural learning proposes that once the psychological machinery that makes use of competence- and success-based cues for targeted cultural learning has spread through the population, highly skilled and successful individuals will be in high demand, and social learners will need to compete for access to the most skilled models. This created a new selection pressure for learners to pay deference to those they assess as highly skilled (those judged most likely to possess adaptive information) in exchange for preferred access and assistance in learning. Deference benefits may take many forms, including coalitional support, general assistance (helping on laborious projects), public praise, caring for the offspring of the skilled, and gifts (Gurven 2001).

With the spread of deference for highly skilled individuals, natural selection can take advantage of the observable patterns of deference to further save on information-gathering costs. Naive entrants (say immigrants or children), lacking detailed information about the relative skill or success of potential cultural models, may take advantage of the existing pattern of deference by using the amounts and kinds of deference that different models receive as cues of underlying skill. Assessing differences in deference provides a best guess to the skill ranking until more information can be accumulated and integrated. Assessing and learning about whom to learn from using the distribution of deference is merely a way of aggregating the information (opinions) that others have already gleaned about who is a good person to learn from.

In addition to patterns of deference, people unconsciously indicate who they think is a good model through a series of ethological and behavioral phenomena that arise directly from efforts to imitate these individuals. These patterns relate to attention, eye gaze, verbal tones and rhythms, and behavioral postures. As other learners seem keenly attuned to these subtle patterns, it appears that natural selection has favored attention to both patterns of deference and those arising from targeted imitation, as a means of assessing whom to pay attention to for cultural learning. As we will discuss below, a mechanism such as “copy the majority” (conformist transmission) provides an effective way to aggregate the information gathered by observing and listening to others. In this case, conformist transmission can be used to figure out whom to pay attention to for cultural learning.

To understand the difference between cues of prestige, success, and skill, consider the following stylized example of an academic department. A new Ph.D. entering a department and aiming at tenure might assess her senior colleagues in order to figure out whom she should learn from (to achieve her goal of getting tenure). Initially, she can glean a measure of people’s prestige by listening to and observing how people act toward one another. If she’s really serious, she might pull up everyone’s curricula vitae and count their publications (and divide by their “years since Ph.D.”). This would give a measure or cue of success. Finally, if our fresh Ph.D. still has not given up all hope of finding a good model, she might read everyone’s papers (or at least those who rank high in “success” and “prestige”) and watch them teach. This would give our learner a measure of skill. Aggregating all these measures, she’d have a decent estimate of whom to start imitating.

Interestingly, the indirect nature of assessing another person’s utility as a cultural model (i.e., a person’s possession of adaptive information that is useful to the learner) creates an important phenomenon. In a complex world, such indirect measures do not tell the learner which of the model’s behavior, ideas, practices, and strategies causally contribute to his success or competence. For example, are people successful in farming because of what they plant, when they plant, how they plant, or how they make sacrifices to the spirits—or all four? Because of this ambiguity, humans may have evolved the propensity to copy successful individuals across a wide range of cultural traits, only some of which may actually relate to the individuals' success. When information is costly, it turns out that this strategy will be favored by natural selection even though it may allow neutral and even somewhat maladaptive traits to hitchhike along with adaptive cultural traits.

Evidence of Selective Model-Based Cultural Learning

Is there actually any evidence for these learning mechanisms? Yes, huge amounts, and it comes from across the social sciences. The evidence shows that children and adults will preferentially learn all kinds of things from individuals demonstrating particular cues of competence, success, and/or prestige—and there need not be any particular relationship between domains of prestige or competence and the things being learned. Unfortunately, the details don't go much beyond that. For example, we would like to know how different kinds of information are integrated. How important is observed competence compared to prestige? And how important is individually acquired information when it contradicts the behavior of highly successful people? Having looked at a wide range of social learning evidence, it is clear that the tendency to imitate prestigious and successful people is one of the most powerful aspects of cultural learning—a point highlighted by the effect of prestige- and model-based cues (e.g., gender and ethnicity) on the imitation of suicide (which we'll discuss below).

In summarizing some of the evidence for the broad power of success- and prestige-biased cultural learning, we emphasize six main points. First, these imitative patterns spontaneously appear in incentivized circumstances (in which choices influence monetary payoffs or other kinds of returns) and nonincentivized circumstances, and in both nonsocial and social situations, including situations that involve direct competition among the learners.[3] Second, the effects emerge broadly across contexts, including economic decisions, opinions, food preferences, beliefs, styles, dialects, and strategies in situations of conflict. Third, consistent with theory, the amount of cultural learning observed depends critically on the degree of uncertainty found in the environment. As uncertainty increases, so does cultural learning. Fourth, these learning patterns emerge even when the model's domain of competence, success, or prestige is apparently unrelated to the domain in question. Fifth, diverse findings from laboratory experiments in both economics and psychology, using very different experimental paradigms, consistently converge—giving us confidence in their robusticity across contexts. Sixth, the patterns of cultural learning observed in economics and psychology laboratories fit closely with field data—giving us confidence that the effects observed in the artificial context of experiments actually matter in the real world. Below, we first summarize some of the laboratory findings to illustrate points 1 through 5, and then describe a few key field studies that illustrate point 6.

Success and Prestige Biases in Nonsocial Situations Pingle (1995) confirms that people (well, university students) will imitate the strategies of successful individuals in nonsocial circumstances, especially when payoffs are on the line. Using a series of computerized decision situations, participants had to repeatedly select the amount of three different inputs (e.g., fertilizer, seed, and labor) into a production problem for either 21 or 31 rounds, depending on the treatment. Before each decision—i.e., before setting the final amounts, (χ_1, χ_2, and χ_3) of the three inputs for a given round—subjects could pay to find out what profit they'd get if they used different sets of inputs. In the baseline treatment, subjects could learn only from their own analyses and direct experience (what they earned each round from their chosen inputs). To calculate profit in each round, the subject's inputs were run through a preset production function. This function, which was unknown to players, had only one set of optimal inputs (χ^*_1, χ^*_2, χ^*_3)—these inputs would make the most money for the subjects. In four other treatments, opportunities for imitation were introduced in varying ways and with different costs. Participants in all treatments faced the same environment (the same production function) for rounds 1 to 11 (Block 1). At round 12, the environment shifted and again remained constant through round 21. For treatments 2–4 and the control, there was also a "competitive" environment that commenced in round 22 with an environmental shift and lasted through round 31 (Block 3). During this block, the optimal set of inputs shifted dynamically and depended on what other players had done. This means that participants faced a new environment beginning in rounds 1, 12, and 22.

The different treatments manipulated the information available for imitation: in treatment 1, during each round (starting in round 2), participants could—at a cost—look at the inputs and output of one other subject who had previously played that round. In treatment 2, participants could—at a cost—look at a list of inputs and outputs for that round for all the participants that had gone before them. In treatment 3, before the play for each block commenced, participants were given the best outputs and corresponding inputs of previous players for that block. In treatment 4, each subject watched two other subjects complete all 31 rounds before playing themselves. Each treatment used different participants, and participants were paid according to the profit they earned, which was determined by their choices of inputs.

A comparison of the findings from across the treatments highlights several important points about imitation:

1. Participants use imitation, often to a substantial degree, even when decisions are financially motivated and cost-benefit analysis is possible (but costly). The pattern of results across all four experiments —vis-à-vis the non-imitation control—shows the strength of our propensity for imitation: in round 2 of treatments 1 and 2, which can be compared directly to round 2 in the no-imitation control, people

imitated 87 percent and 57 percent of the time, respectively.

2. Imitation tendencies remain strong even in competitive environments. About 43 percent of subjects imitated in round 22 of treatment 2.
3. People tended to imitate (the inputs of) more successful players (those who got higher outputs). The patterns in the data are explicable only if people are looking at the difference in performance and using that as a cue about when to imitate.
4. Uncertainty causes a substantial increase on the reliance on imitation. In rounds 2, 12, 22, when a new environment is first encountered, rates of imitation are highest.
5. The availability of imitative opportunities, even costly ones, improves the average performance of the group. As a group, subjects in imitation treatments outperformed those of the control.
6. The "imitation environment" affects the average performance of the group. Average performance in treatment 3 and 4 exceeds that of treatment 1 and 2. Treatment 3, in which individuals have the best inputs from previous players, is the only informational environment that avoided a substantial degradation in group performance during the competitive block (Block 3).

Now, let's consider a different incentivized experiment, again from economics. Kroll and Levy (1992) ran two treatments of an investment game. In the control treatment, MBA students had to allocate money among three possible investments, each with different expected returns, return variances, and correlations between returns. With all this known, participants set up their portfolios across the three investments and realized some return each round for 16 rounds. Between rounds, they could adjust their allocations among investments but had no idea what was happening to other players. The "imitation-possible" (IP) treatment was identical except that the allocations and performance of all other subjects were posted (without names or labels) between rounds for all participants to see.[4]

Two findings from this study confirm those discussed above. First, using a round-by-round regression analysis, Kroll and Levy found (to their surprise) that participants strongly mimicked the top performers in the IP treatment. That is, participants used success-biased cultural learning to figure out a "wise" investment strategy. Moreover, a comparison of the control and IP treatments shows that adding the possibility of imitation brought the whole group, even after 16 rounds of direct experience, much closer to the optimal allocations predicted by Portfolio Theory. As in the previous experiment, a competitive environment in which individuals pursue their own interests but are also allowed to freely learn from one another creates a social effect: the group does better.

Interestingly, these imitation mechanisms also can produce maladaptive behavior (i.e., bad decisions). Dual Inheritance Theory shows that cultural learning mechanisms exist because they did better on average than alternative decision-making or learning mechanisms in our ancestral environments (the Paleolithic). Confirming this, Kroll and Levy also note that those who overimitated the top performers in the investment game sometimes ended up losing everything catastrophically because they imitated players who had risked substantial amounts and gotten lucky.

So far, the experiments we've summarized have shown that people will copy other people's economic strategies. But, will they copy their *beliefs* about the state of the world? To test this and to more precisely examine whether people will preferentially copy more successful people under uncertainty, Offerman and Sonnemans (1998) designed the following experiment. To represent an "unknown state of the world," they used an urn with a large number of red and white balls (the variable p equals the probability of picking a red ball from this urn). Players know the urn exists and contains red and white balls, but they don't know how many of each color. They are also told that the number of each type of ball does not change. Each round, four balls are randomly chosen from this urn, which yields five possible conditions (0 red balls, 1 red ball, etc.). Players (students again) earn money in two ways. First, they state their beliefs about the likelihood of each of the different conditions occurring (each must specify five probabilities corresponding to 0 reds, 1 red, … 4 reds). Payments are made according to the accuracy of each reported belief compared to the true probability. Second, participants must decide whether to "invest" or "not invest." Not investing leads to receiving a fixed sum for that round. Investing leads to losses if the condition of the round is 0, 1, or 2 reds, but yields substantial monetary gains (about three times the fixed amount for not investing) if three or four reds are picked in that round. This means that, in general, someone who prefers more money ought to generally invest more if they believe p is high.

Three treatments were run for 10 rounds each. In the control treatment, players received no information about what other participants did. They gave one set of beliefs and made an investment decision in each round. In the rounds of treatments 1 (T1) and 2 (T2), players stated their beliefs and investment decisions, then received information about the beliefs, investment decisions, and payoffs of two other players, after which they again stated their beliefs and investment decisions for that round.[5] Treatments 1 and 2 differed in only one respect: in T1, all players faced the same conditions (e.g., 3 red balls and 1 white were drawn this round), so in a sense, everyone received the same information. In T2, each player faced a different independent draw from the same urn each round, so that in some sense, the beliefs and behavior of other players should indirectly contain additional information about the state of the world (p). If people were perfect individual learners, the behavior of others in T1 would not be worth paying attention to, as it does not add any new information. Treatment 2 adds new independent observations except that the learner does not get to observe the condition directly but only via the others' decisions and beliefs.

By comparing players' reported beliefs before and after receiving the information about the beliefs of the others, Offerman and Sonnemans assessed the effects of social information. Interestingly, two different learning biases emerged. First, of the three individuals, those with lower payoffs tended to move toward the beliefs of those with higher payoffs. However, individuals with the highest payoffs (of the three) still moved their beliefs toward the other two. It is as if players adjusted their beliefs by doing a kind of payoff-weighted average of their own beliefs and others'.

Comparing the control with the two treatments reveals a few additional findings that reiterate what we've seen above. First, adding cultural learning opportunities improves the average performance of the group. Players in T1 and T2 made more money than those in the control, and those in T2 did somewhat better than those in T1. Second, the imitation of beliefs occurs in both treatment conditions, although there was more in treatment 2. This is important because a critic might expect individuals to use information as perfect individual learners would (what economists call "rational Bayesians": McKelvey and Page 1990). A perfect individual learner, however, would not imitate in treatment 1 (and it remains a bit murky how exactly a perfect individual learner would extract the necessary information from the beliefs of others in T2). However, the use of other people's beliefs makes good sense from a Dual Inheritance perspective if we assume that human brains are imperfect information processors and that natural selection had to build a brain that could contend with its own imperfection. Culturally learning from others, even if they perceived identical events, can reduce the amount of error that creeps in as information is perceived, processed, organized, and stored.

Social Situations: Success-Biased Imitation in Strategic Conflicts Maybe people copy other people in situations involving purely individual decisions, but do they really imitate others in situations in which the behavior of others also influences their own payoffs? Yes. Let's begin by reviewing some experiments that have explored imitative learning in competitive market situations called Cournot markets. In these experiments, individuals are put into groups of two to nine individuals (depending on the experiment and treatment). The game is played for some number of rounds (between 20 to 60). Each round, players make one decision regarding the quantity of some product they will produce. As a result, they receive some profit that depends on the quantity they produce and consumer demand, which depends on the aggregate quantity that all players produce (demand goes down as the aggregate production across all the players in a group goes up). All this is calculated using fixed mathematical functions that are unknown to players. At the end, players are paid in real money according to their total profits. Using this basic setup, experimenters have varied the information available and/or the ease of use of that information, or the time available for decision-making. By varying the treatments and using learning models of various kinds (e.g., success-bias imitation, reinforcement, belief-based best-response, etc.) to predict how players change their behavior over the rounds of the game, researchers can study how people adapt to these situations and how different informational conditions draw forth different adaptive learning tools.

To summarize, the results from several studies demonstrate that players adapt by using imitation strategies along with individual learning strategies. When they do imitate, players have a strong preference for other players who have received higher payoffs. In synthesizing the findings of their experimental study, Alpesteguia, Huck, and Oeschssler (2005) write: "If individuals can imitate actions, most of them do so. And some do so almost all the time." These analyses explicitly show that the probability of imitating increases proportionally with the observed difference in payoffs (this finding is remarkably consistent with the evolutionary prediction; see Schlag 1998 and 1999). Interestingly, Offerman, Potters, and Sonnemans (2002), using a similar setup, found that participants used two kinds of cultural learning mechanism: sometimes they imitated the most successful in their trio, and other times they imitated people who chose more group-beneficial options.

The preceding studies are important because decision making is incentivized, and the available information is rigorously controlled. Qualitatively, however, these findings from economics merely confirm older empirical insights from psychology. We will summarize some of the work from psychology to show that these findings are robust across disciplines, experimenters, experimental traditions, and scientific time. Together, results from the two traditions form a potent argument. Not only do these cultural learning mechanisms operate in incentivized decision making, as studied by the economists, but they also appear in nonincentivized situations in which behavior, opinions, and preferences shift both spontaneously and unconsciously.

In testing the early observations of Miller and Dollard (1941) about prestige and social learning, psychologists Rosenbaum and Tucker (1962) used an extremely artificial experimental setup in which pairs of subjects had to pick the winners of horse races. Of course, as with much early work, there were no horses or real money, but only isolation booths, buttons, and colored lights. Although crude, this work did show that "model competence" (or the frequency of correct answers made by a model), strongly affected the subject's propensity to imitate the model's choices (which horse to bet on), even when those answers were unconnected to the imitator's circumstances. Baron (1970), using a similar setup, provided a confirming result. In addition, numerous other studies have also shown how model success biases imitation (Chalmers, Horne, and Rosenbaum 1963; Greenfield and Kuznicki 1975; Kelman 1958; Mausner 1954; Mausner and Bloch 1957). Qualitatively, these studies report basically the same patterns seen above: people copy the choices of successful individuals.

However, people imitate more than choices, decisions, and strategies. Our opinions are deeply influenced by the opinions expressed by prestigious others. About half a century ago, Tannenbaum (1956) investigated the effect of different sources on opinion change and found evidence for prestige learning biases. The sources used were: (1) a prominent individual, (2) a prominent newspaper, and (3) a prominent social group. His 3x3 design tested for how the same source would bias subjects who (1) had prior respect for the source, (2) were previously neutral toward it, or (3) had prior disrespect for it, on an attitude/opinion item they (1) previously agreed with, (2) were previously neutral toward, or (3) previously disagreed with. He tested the effects of the source's positive and negative attitudes and found that subjects' attitudes were pulled closer to the source's even when subjects' prior opinions were contrary.[6] With similar findings, Haiman (1949) also showed that a model's prestige and perceived competence significantly affected the opinions of others, while perceived physical attractiveness, "likeability," and "fair-mindedness" did not.[7]

Experimental data also show that opinions are influenced even when the prestigious individual's domain of prestige is seemingly unrelated to the opinion domain in question. Ryckman et al. (1972) showed that prestigious individuals whose expertise is in the domain in question and those with prestige from other areas of expertise both influence opinion change. In one set of conditions, participants first did an opinion survey on student activism and watched a live speech by either an individual introduced as an expert in student activism or a college sophomore. In each treatment, the same person played the expert and the sophomore and gave the same speech in both roles. After this, participants repeated the opinion survey. In another treatment with the identical setup, participants listened to either an expert on the Ming Dynasty or a college sophomore. In both conditions, people's opinions moved toward those expressed by the prestigious person after hearing the expert but not after hearing the sophomore. In a similar study using the topic of national budget priorities, Ritchie and Phares (1969) obtained similar results with a leading economist versus a college sophomore in the high- versus low-status manipulation.

Same Patterns Are Observed Outside the Laboratory The same patterns observed in contrived laboratory experiments are observed in "real people" (non-student adults) outside the laboratory. One might wonder, then, why laboratory work is necessary at all. The answer is simple: the laboratory allows for tightly controlled methods of isolating cues and of measuring their relative strengths. The real-world data cannot do this as effectively, but it can confirm that what we see in the laboratory does influence real life.

In his massive review of the literature on the diffusion of innovations, Everett Rogers (1995, p. 18) summarizes some of the lessons from 50 years of research as follows:

> Diffusion investigations show that most individuals do not evaluate an innovation on the basis of scientific studies of its consequences, although such objective evaluations are not entirely irrelevant. … Instead, most people depend mainly upon a subjective evaluation of an innovation that is conveyed to them from other individuals like themselves who have previously adopted the innovation. This dependence on the experience of near peers suggests that the heart of the diffusion process consists of the modeling and imitation by potential adopters of their network partners who have adopted previously.

Rogers goes on to devote an entire chapter to explaining how the diffusion of new ideas, technologies, and practices is strongly influenced by “local opinion leaders.” Compiling findings from many diffusion studies, Rogers describes these individuals as: (1) locally high in social status (e.g., high status within the village or village cluster), (2) well respected (indicating prestige), (3) widely connected, and (4) effective as social models for others. Rogers’s insights are particularly important here because they confirm that success-and prestige-biased cultural learning are important for the spread of novel technologies and practices.

Naturalistic studies using a jaywalking manipulation have consistently found that people preferentially copy the behavior of high-status models. In these naturalistic experiments, high-status models wear business suits, and low-status models appear disheveled and impoverished. In a meta-analysis of seven studies on jaywalking, Mullen et. al. (1990) show that a high-status obedient model increases others’ compliance to “no jaywalking” rules, whereas a low-status obedient model has no significant effect.

Studies by sociolinguists show that these processes apply equally well to the transmission of dialect. Labov (1980, 1972) has shown that dialect change is driven by locally prestigious people. In Philadelphia, upper-class working women pioneered novel sound changes, which then spread through the local social strata. In Martha’s Vineyard, most folks are not aware of the dialect differences between themselves and mainlanders. Yet they seem to have granted considerable social status to local fishermen—who exemplify the local spirit of resistance and tradition—which has led to inadvertent copying of the salty dialects spoken by these locally prestigious people.

Using the idea of selective model-based cultural learning (context biases), we can even illuminate some of the robust patterns observed in studies of suicide. Data from industrialized societies show that committing suicide and its associated methods are imitated according to prestige and self-similarity (Stack 1987, 1990, 1991, 1996; Wasserman, Stack, and Reeves 1994). For prestige, many studies show that suicide rates spike after celebrity suicides (Kessler, Downey, and Stipp 1988; Stack 1987). This pattern has been observed in the United States (Stack 1990), Germany (Jonas 1992), and Japan (Stack 1996). The individuals who killed themselves after celebrities did tended to match their models on age, sex, and ethnicity, and they even copied the method. In fact, even fictitious suicides are imitated, although real suicides result in effects that are four times greater (Stack 2000). The time trends of these suicides do not show regression to the mean during the subsequent month, so these are not individuals who would have committed suicide in the near future anyway.

Because suicide is strongly influenced by imitation, suicide can spread in epidemic fashion, showing pattern similar to diseases, novel cultural practices, and innovations. In Micronesia (Rubinstein 1983), beginning in 1960 and lasting for at least 25 years, just such an epidemic spread through certain island populations. This case is particularly stark because the suicides were geographically patterned and distinctively stereotyped. The typical victim was a young male between 15 and 24 (modal age of 18) who still lived at home with his parents. After a disagreement with his parents or girlfriend, the victim was visited in a vision by past suicide victims who "called him to them" (we know this from parasuicides). Heeding the call, the victim performed a "lean hanging" from either a standing or sitting position, usually in an abandoned house, until he died of anoxia or was accidentally discovered and saved. In 75 percent of the cases, there was no prior hint of suicide or depression. Moreover, this pattern was restricted to those who could trace their ethnic descent to Truk or the Marshals, and emerged in localized sporadic outbreaks among socially interconnected adolescents (which parallels the U.S. pattern: Bearman 2004), and could sometimes be traced to the precipitating suicide of a prominent son of a wealthy family. This is all the more stark when contrasted with suicide patterns elsewhere in Oceania. For example, the strong gender bias and lean-hanging method contrast with the pattern in Western Samoa, where male-female rates are nearly equal and most suicides (and attempts) involve drinking paraquat, a deadly herbicide (Booth 1999).

Conformity Biases

What do you do when any observable differences in skill, success, and prestige among individuals do not covary with the observable differences in behavior, beliefs, practices, or values? For example, suppose everyone in your village uses blowguns for hunting, except one regular guy who uses a bow and arrow and obtains fairly average hunting returns. Do you adopt the bow or the blowgun? One solution for dealing with such information-poor dilemmas is to copy the behaviors, beliefs, and strategies of the majority (Boyd and Richerson 1985; Henrich and Boyd 1998). Termed conformist transmission, this mechanism allows individuals to aggregate information from the behavior of many individuals. Because these behaviors implicitly contain the effects of each individual's experience and learning efforts, conformist transmission can be the best route to adaptation in information-poor environments. To see this, suppose every individual is given a noisy signal (a piece of information) from the environment about what the best practice is in the current circumstances. This information, for any one individual, might give a 60 percent chance of noticing that blowguns bring back slightly larger returns than bows. Thus, using individual learning alone, individuals will adopt the more efficient hunting practice with probability 0.60. But if an individual samples the behavior of 10 other individuals, and simply adopts the majority behavior, his chances of adopting the superior blowgun technology increase to 75 percent.

The same logic can be applied to the effect of imperfect information about the relative success of others, who may be useful as cultural models. Some individuals may obtain accurate information that allows them to effectively pick out and copy the most successful individuals, whereas others may receive noisy (inaccurate) information about relative success, which prevents them from effectively distinguishing differences. This second group can still take advantage of the more accurate information received by the first group by adopting the traits adopted by the majority. To see this more clearly, imagine a group of 200 individuals, wherein 100 are experienced hunters and 100 are novices who need to figure out which technology to invest in learning. Of the 100 experienced individuals, suppose that 40 use bows and that 60 use blowguns for hunting. In their current environment (perhaps it recently changed), however, bows obtain more efficient returns, although the difference is small and hunting returns in general are highly variable. Nevertheless, using the returns of the experienced hunters, 40 of the 100 novices selected a bow hunter to learn from, 50 were left confused, and 10 picked a blowgun hunter to learn from (they got bad information because of the noise in returns). Thus 80 hunters now use bows (40+40), and 70 use blowguns. In their confusion, the 50 confused novices decide to use conformist transmission. This will result in more than 53.3 percent of the confused individuals' adopting bows. For example, of the confused 50, 40 might adopt bows, whereas 10 still decide to go with blowguns. After all of the transmission this generation, 120 hunters will use the more adaptive bows, and 80 will use blowguns. If the older (experienced) generation dies, 80 percent of the new generation will use bows.

This kind of verbal reasoning has been rigorously tested in both analytical models (Boyd and Richerson 1985: chap. 7) and extended to more complex environments using evolutionary simulations (Henrich and Boyd 1998; Kameda and Nakanishi 2002).[8] In their computer simulation, Henrich and Boyd investigated the interaction and coevolution of vertical transmission (parent-offspring transmission), individual learning, and conformist transmission in spatially and temporally varying environments. The results confirm that conformist transmission is likely to evolve under a very wide range of conditions. In fact, these results show that the range of conditions that favor conformist transmission are broader than those for vertical transmission alone—suggesting that if true imitation (via vertical transmission) evolves at all, we should also expect to observe a substantial conformist component. This work leads to several specific predictions about human psychology. First, learners will prefer conformist transmission over vertical transmission, assuming it is possible to access a range of cultural models at low cost (which is often but not always the case). Though a direct test of this prediction is lacking, we note that a substantial amount of research in behavioral genetics indicates that parents actually transmit very little culturally to their offspring—once genetic transmission is accounted for, vertical cultural transmission often accounts for less than 5 percent of the variation among individuals (J. Harris 1995, 1998; Plomin, Defries, and McLearn 2001). Second, as the accuracy of information acquired through individual learning decreases, reliance on conformist transmission (over individual learning) will increase. Third, as the proportion of models—in the learner's sample of models—displaying a trait increases, the strength of the conformist effect should increase nonlinearly as well. We address the second and third predictions below.

Consistent with the broad thrust of the above theoretical work, a substantial amount of empirical research shows that people do use conformist transmission in a wide range of circumstances, particularly when problems are complex or difficult to figure out on one's own. This work reveals that humans have two different forms of conformity that operate in different contexts. The first, often called informational conformity, matches the theoretical expectations from models of conformist transmission, and is used to figure out difficult problems. It results in people actually altering their private opinions and beliefs about something. The second, often called normative conformity, is conformity for the purposes of going along with the group, to avoid appearing deviant. Under this type of conformity, people alter their superficial behavior but often don't change their underlying opinions or beliefs. We argue that the ultimate origins of this second type of conformity can be explained by the evolutionary process that we describe under the rubric of "social norms" in the next chapter and in chapter 7. Below, we sketch some of the evidence for informational conformity (conformist transmission), which is consistent with the predictions derived above, from Dual Inheritance Theory.

Conformist Transmission in Nonsocial Situations

Baron et al. (1996) used a lineup experiment to systematically vary both the difficulty and importance of a task. Working alone, control subjects observed slide projections of a "perpetrator" for some fixed period and then attempted to pick him out of a lineup of four "suspects." Test subjects faced the same task, except each was in a room with two other "subjects" who were actually confederates of the experimenter. To obscure the actual structure of the experiment from the subjects, seven "critical trials" were interspersed with a series of other mundane trials. During these critical trials, both confederates would give the same incorrect answer before the actual subject gave his answer. Conformity was then assessed by comparing the answers of the test subjects on these critical trials to those of the control subjects. Difficulty was varied by changing the amount of time a subject could initially view the perpetrator. In the easy version, subjects viewed the perpetrator for five seconds, and in the more difficult problem they viewed the perpetrator for one-half second. Control subjects selected the correct person from the lineup 97 percent of the time on the easy version and 76 percent of the time on the harder version. Task importance was also varied. In the less important task, subjects performed for the "good of science." In the more important task, subjects performed for the "good of science" and received a lottery ticket in a $20 lottery for each correct answer. Figure 2.1 presents the results.

In Figure 2.1, the y-axis plots the mean number of conforming trials (out of seven possible trials), and the x-axis shows the "less important" versus "more important" (i.e., unpaid versus paid) versions. The dotted line connects the mean number of conforming trials for the treatments when the task was moderately difficult (control group got 76 percent correct), while the solid line shows them when the task was easy (control group got 97 percent correct). The key thing to note is that when importance increased, the social influence decreased as long as the problem remained quite easy. However, when the problem was moderately difficult, social influence increased.

To clarify, in the unpaid situation, subjects were motivated principally by a desire to conform to the two other (confederate) subjects, with only a slight (nonsignificant) effect of using the social information more in the ambiguous circumstance. However, when money was added as a motivation, it overcame much of the desire to conform to the other two (whom the subject did not know) and thereby decreased conformity for the easy problem (where subjects knew the right answer 97 percent of the time based on their own perceptions)—they wanted the money, so they were motivated to give what they were pretty sure was the correct answer. But when subjects were not so sure, but still wanted to get the answer correct, they shifted their reliance toward the choice made by the other two. When an individual's ambiguity about a behavior (or answer) increased and they were motivated to get correct answers, their reliance on social learning increased. This bears out a prediction that comes directly from evolutionary theory.

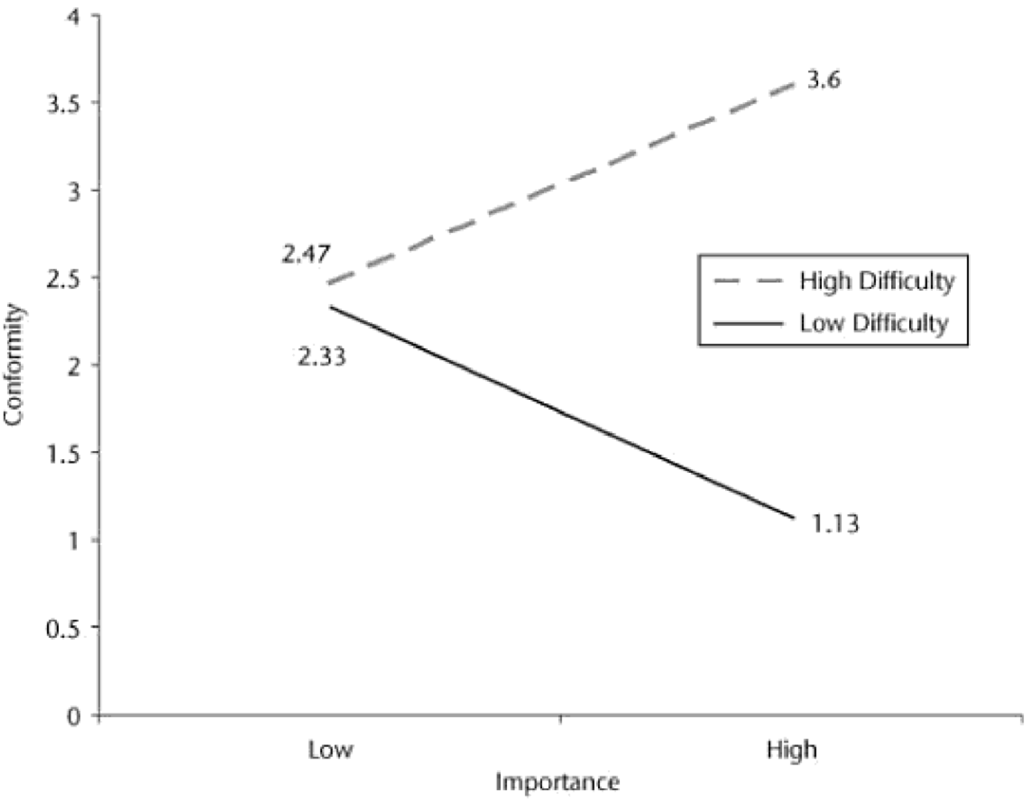

Figure 2.1 Mean number of conforming trials by Condition in the lineup task. Data taken from Baron et al. 1996. ▬▬▬ Low Difficulty, ▬ ▬ High Difficulty.

Using an experimental design that closely parallels the evolutionary simulation constructed by Henrich and Boyd, McElreath, Lubell, et al. (2005) tested the predictions of conformist transmission. Undergraduate subjects repeatedly faced an economic choice between two options, A or B, for 20 rounds. This was posed as a “farming decision” in which A and B were different crops with different yields and variances. Players did not know the mean yields or yield variances for the two crops, but they were told that the local environment might fluctuate such that the mean yields of the crops change. After each round, each player learned the yield realized in that year for his field, and could choose to look at the decisions (crop A or B, but not the yields) of other farmers in the past year. At the end of the 20 rounds, players were paid according to their total yield over the 20 seasons, and made between $4 and $8. Consistent with the theoretical predictions, McElreath, Lubell, et al.’s analysis found that (1) people increase their use of social learning when crop variance is high and decrease it when environmental conditions vary over time; and (2) a simple conformist learning rule (copy the majority) seems to capture an important part of decision making, although there is quite a bit of individual heterogeneity.

A naturalistic experiment using nonincentivized actions further confirms these conformist effects by showing the nonlinear influence of the frequency of a behavior (Coultas 2004). Here, subjects entered a computer lab one by one, not realizing they were in an experiment, and observed a “rare behavior” that involved placing the keyboard cover on top of the monitor. In pretesting, the experimenters confirmed that no one, without modeling, ever put the cover on top of the monitor—so without modeling, the expected frequency of placing the cover on the monitor is zero. The experimenters were able to manipulate the number of individuals placing the cover on their monitor by silently giving explicit instructions to a few through their computers. Others, not receiving these instructions, were observed to see if they placed the cover on top of the monitor. Figure 2.2 summarizes the results by showing how the frequency of models performing the cover placement affected a subject’s likelihood of making the same placement. The horizontal axis gives the percentage of individuals already in the room who had their keyboard covers on top of their monitor as the subject entered. The vertical axis gives the probability that the subject would then place his keyboard cover onto his monitor. As predicted, the likelihood of performing this behavior, which is not otherwise performed, increases nonlinearly as the percentage of models performing the behavior rises above 50 percent. One problem with this experiment is that it does not carefully distinguish informational from normative conformity.

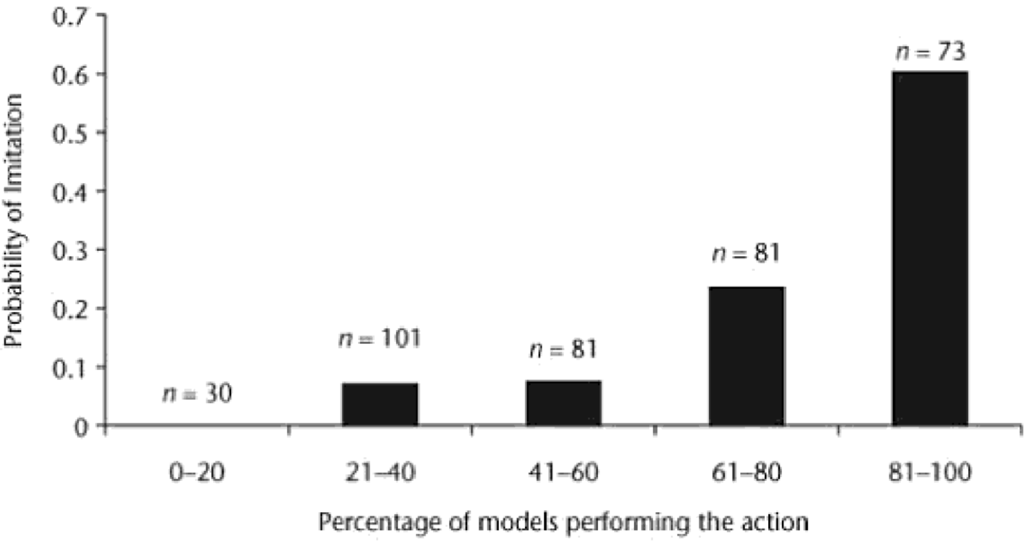

Figure 2.2 The figure shows how the percentage of models performing the "covers on the monitor" behavior influences the likelihood of others performing the same behavior. The n values above each bar gives the number of individuals observed for that bar (e.g., 73 subjects entered a room in which 81 percent to 100 percent of the people in the room had their keyboard covers on their monitors; about 60 percent of these subjects then put their covers on the monitor). The data are taken from Coultas 2004.

Conformist Transmission in Social Situations—Particularly in Cooperative Dilemmas

In a study of the effect of social influence on common-pool resource games, Smith and Bell (1994; also see Wit 1999) argue that players sometimes copy other players when they're uncertain about what to do. Two forms of a multi-round common-goods game were used. In both games, one subject and two confederates can withdraw "points" from a common pool that initially contains 15 points. The number of remaining points in the pool doubles every other round, but it cannot exceed 15 points. The game lasts until the common pool goes to zero or until 15 rounds are played. In version 1 of the game, players receive lottery tickets (a chance of winning real money) according to their personal point totals at the end of the game; in version 2, players receive lottery tickets according to their group's point total at the end of the game. In both versions, subjects show a significant reliance on mimicking confederates' behavior. When confederates underutilize the resources, subjects tend to underutilize them as well, and when confederates overutilize, subjects also tend to overutilize resources. Because subjects behave similarly when their self-interest equates with the group's interest and when it's opposed to the group, the authors argue that they mimic as a means of using social information under uncertainty and not as a means of competing.

Broadening this finding, Carpenter also finds evidence of conformist transmission in the behavior of players in a cooperative social dilemma (2004). Similarly, Denant-Boemont, Masclet, and Noussair (2005) observe a conformist effect on the punishment of nonpunishers in cooperative dilemmas where subjects can pay to punish (take money away from) those who don't cooperate, and the subjects even punish those who don't punish noncooperators.

These results, which show that people copy others, especially when the right thing to do is unclear when many others are doing the same thing, are consistent with the predictions of Dual Inheritance Theory and are important because they address economic decisions in social circumstances.

Cultural Learning of Altruism and Punishment

Faced with the preceding evidence for cultural learning, a skeptic of the importance of culture for understanding social behavior might argue that although the predicted effects are observed in the social learning strategies found in laboratory experiments and in the costless expression of opinions or dialects, they probably do not influence the acquisition of durable values or preferences that favor taking individually costly actions, such as those related to altruism or cooperation. We take this challenge seriously in this section by offering converging lines of evidence that demonstrate that people do use cultural learning to acquire intrinsic motivations (preferences or values) associated with at least one kind of altruism/cooperation.

Psychologists, mostly during the 1960s and 1970s, have intensively studied the imitation of a particular type of altruism (i.e., helping strangers in a nonrepeated context) in children. In the paradigmatic experimental setup, from which endless variations emerged, a child is brought alone to the experimental area (often a trailer on school grounds) to get acquainted with the experimenter. Then, the child is introduced to a bowling game and shown a range of attractive prizes that he or she can obtain with the tokens (or pennies or gift certificates) won during the bowling game. Subjects are also shown a charity jar for poor children where they can put some of their winnings if they want. This jar is next to the bowling game and often has a "March of Dimes" poster over it, or some facsimile. A model, who could be a young adult or another peer, demonstrates the game by playing 10 or 20 rounds. On winning rounds, which are preset, the model donates (or not, depending on the treatment) to the charity jar. After finishing the demonstration, the model departs (or not, depending on the treatment), and the child is left to play the bowling game alone, often monitored through a one-way mirror. Putting money in the jar results in the child getting fewer toys or other desirable items, so this is clearly costly altruism (and postgame interviews with subjects confirm that they understood the costs).

The results from numerous researchers involving hundreds of children (ages 6 to 11) and a wide range of setup variations demonstrate three robust findings for our purposes here. First, children spontaneously imitate either the generosity or selfishness of the model. That is, if the children see a model donate to the charity jar, they will donate more, whereas if they see a model fail to donate, they will donate less (than they would without any model). The more the model donates, the more the children donate, and the more opportunities a child has to observe the model, the greater the degree of imitation (Bryan 1971; Bryan and Walbek 1970b; Grusec 1971; Presbie and Coiteux 1971). Similar patterns occur whether models are peers or adults (Hartup and Coates 1967). Second, beyond merely imitating the practice and the amount, children also will readily imitate other aspects of the model's behavior. For example, if multiple charity jars are used (with different recipients), children are influenced by the model's distributional preferences (Harris 1970, 1971). Similarly, Bryan (1971) showed that children not only imitate the amount of generosity, but also the order of putting money in the charity jar versus their personal jar.[9] Third, the effects of exposure to a model endure over months in retests (without a model present) and extend to somewhat similar contexts (Elliot and Vasta 1970; Midlarsky and Bryan 1972; Rice and Grusec 1975; Rushton 1975).

Showing the power of cultural learning, the modeling of altruism in experimental situations such as this has consistently, and often to a substantial degree, emerged as the most effective means to instill altruistic giving in children. Several studies compare the effect and interaction of models that, on the one hand, practice generosity or selfishness and, on the other, preach ("exhort") either generosity or selfishness (e.g., "one ought to donate …"). Preaching alone has been shown to have little effect (or sometimes a negative effect; see Grusec, Saas-Kortsaak, and Simutis 1978). However, even when it has been observed to increase donating, it is always overshadowed by modeling effects. That is, children tend to ignore exhortation toward costly actions when those exhortations are inconsistent with the model's deeds (Bryan, Redfield, and Mader 1971; Bryan and Walbek 1970a, 1970b; Rice and Grusec 1975; Rushton 1975).[10] In postexperimental recall tests of what models said (preached) and did (donated), subjects had the most difficulty in accurately remembering the words and deeds of "inconsistent" models (Bryan and Walbek 1970a).

Figure 2.3 summarizes some experimental results that usefully illustrate several of the above points. Here, children were tested in the basic bowling experiment using six different treatments and were retested six months later. Along with a no-model control condition, children observed a model who (1) preached generosity, (2) preached selfishness, or (3) merely showed positive affect ("This is fun"). This preaching condition was paired with donating either (1) generously or (2) very little. The vertical axis gives the fraction of tokens the subjects donated to charity (to "poor children"), with the horizontal axis separating the treatment groups. This figure illustrates that (1) modeling donations has the biggest effect on children (note the vertically striped bars across preaching conditions); (2) modeling is augmented by showing either positive affect or preaching generosity (compare the vertically striped bars for "positive affect" and "preaching generosity" with "preaching selfishness"); (3) this effect endures on retests two month later; and (4) showing positive affect increases the effect of imitation more than either type of preaching (compare the difference between the vertically striped and black bars across conditions, and in the retests).

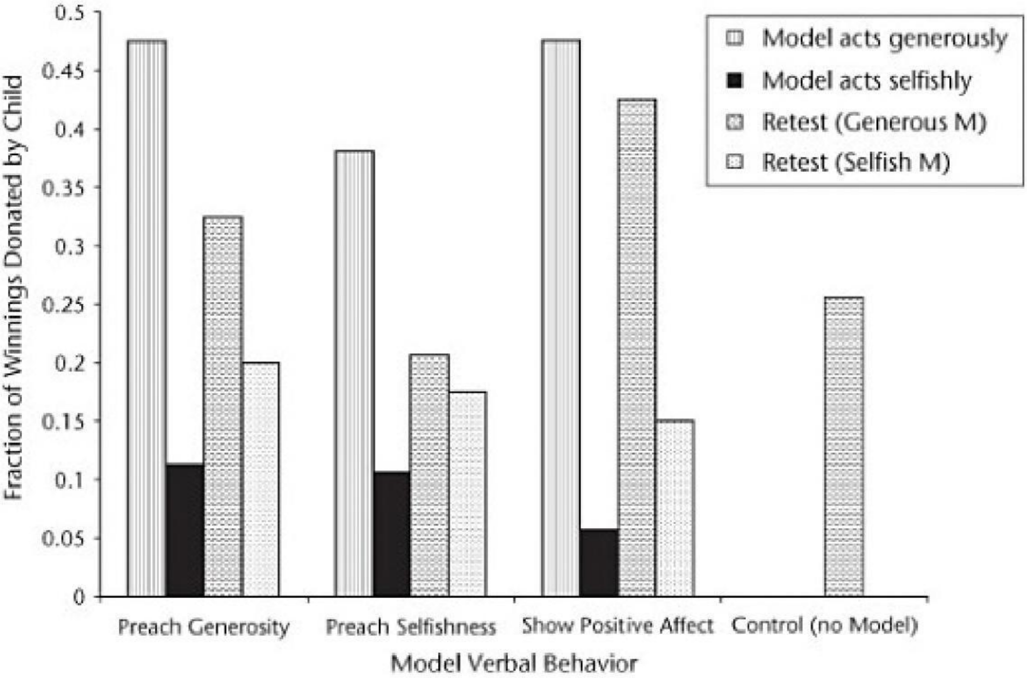

Figure 2.3 The figure shows the effects of various modeling treatments involving combinations of demonstrated behavior and verbal preaching, measured both immediately after modeling and on a retest two months later. The vertical axis shows the fraction of the subject's winnings donated to help "poor children." Data gleaned from Rushton 1975.

Finally, using the same type of bowling game setup, psychologists have shown that children will imitate standards for self-reward or self-punishment and readily impose those standards on others. In this work, children observe a model rewarding himself by taking tokens when he gets above a certain score, or punishing himself by returning tokens to a bin when he gets below a certain score (Bandura and Kuper 1964).[11] Having learned via imitation or instruction, these children are then allowed to teach another novice, on whom they spontaneously impose the same standards that they recently acquired. These effects endure in retests four weeks later (Mischel and Liebert 1966).

Of course, none of this is limited to children. Both experimental and naturalistic studies show that adults are influenced by imitation. We saw above that undergraduates with real money on the line readily imitate others in deciding how much to contribute in cooperative dilemmas, both when self-interest conflicts with the group's interest and when self-interested choices correspond to group-interested choices. Additionally, in quite natural settings, providing models for imitation has been shown to increase (1) volunteering to help in experiments (Rosenbaum and Blake 1955; Schachter and Hall 1952), (2) helping a stranded motorist, (3) donating to a Salvation Army kettle (Bryan and Test 1967), and (4) giving blood (Rushton and Campbell 1977). Modeling treatments often increase helping rates by 100 percent or more vis-à-vis control treatments.

These empirical findings on imitation and altruism set the stage for the acquisition of social norms via cultural learning, which we'll discuss in the next chapter. Summarizing the above, cultural learning appears to be the most powerful means through which children readily acquire certain kinds of altruistic behaviors (which can be generalized to similar situations) and standards of conduct or performance (including self-rewards or punishments), which they spontaneously apply to peers.

Culture-Gene Coevolution

As just highlighted, the power, fidelity, and importance of cultural learning in our species leads to a second system of inheritance that, in turn, can influence genetic evolution. In the next chapter, one of the theories we will discuss adds clarity to how cultural evolution, driven by competition among cultural groups, may have altered the selective environment faced by genes. We argue that this coevolutionary interaction, perhaps operating over tens of thousands or even hundreds of thousands of years, has deeply affected human sociality, altruism, and cooperation across the entire species. In chapters 6–9, we provide support for this coevolutionary approach by comparing its empirical entailments with those of alternative theories using a combination of ethnographic and laboratory findings. However, because some readers may not be familiar with the degree to which culture can influence genetic evolution in general, we will first highlight research that shows how cultural evolution has influenced recent genetic evolution (occurring in the last 10,000 years) and patterns of genetic diversity in our species. It is important to realize that in contrast to our later substantive arguments that focus on universal aspects of our coevolved social psychology, the findings summarized below are interesting because they sharply reveal "culture-gene coevolution in action" by taking advantage of the synchronic correlations between certain genes and certain cultural traits.

The classic example of culture-gene coevolution emerges from the historically strong correspondence between the prevalence of "lactose absorption" by adults and certain culturally transmitted practices. As background, we offer a few facts: most mammals and all primates have young that produce an enzyme called lactase, which allows them to process the lactose sugars found in milk (for example, from humans, cows, camels, and goats), and thereby access the available nutrition. The production of the lactase enzyme appears to be a developmental phase necessary for using mother's milk during infancy, and typically disappears well before adulthood. This dominant mammalian and primate pattern allows us to associate the ancestral human state with the inability to absorb milk sugar in adults. Interestingly, in some populations of our species—principally those with descent from northern Europe and nomadic pastoralist groups in Africa—most people possess the ability to process milk sugars through adulthood. The coevolutionary hypothesis that best fits the data goes like this: As populations began adopting the culturally transmitted practice of domesticating large mammals (cows) for meat and perhaps for milk to supplement children's diets, a novel selection pressure was created favoring genes that extend the lactase production into adulthood, thereby allowing adults to reap the benefits of the nutrition available in milk. In some places (northern Europe and pastoral Africa), natural selection eventually responded with a "genetic solution" (Bersaglieri et al. 2004; Mulcare et al. 2004). However, in places such as the Middle East and China, the cultural transmission of the practice of keeping large domesticated animals was followed relatively rapidly by the evolution of the technological know-how for (and practice of) turning milk into cheese and yogurt. In these forms, lactase is not required, and anyone can obtain the nutritional benefits of milk. In these populations, adaptive cultural evolution beat natural selection acting on genes to the punch (and thereby sapped the strength of selection on genes) by taking advantage of the opportunity created by the presence of domesticated mammals. As expected, we find the relevant alleles, lactase enzymes, and milk-drinking practices among populations who, historically, had domesticated animals but had not developed the cheese- and yogurt-making know-how or practices[12] (gleaned from: Bayless and Rosenwei 1966; Beja-Pereira et al. 2003; Bersaglieri, et al. 2004; Cavalli-Sforza 1973; Durham 1991; Holden and Mace 1997; J. Johnson et al. 1977; Simoons 1970).

Our second example comes from the Sino-Tibetan–speaking hill tribes of Thailand. Oota et al. (2001) explored the influence of two different postmarriage residential cultural practices on the genetic diversity of the Y-chromosome (which is transmitted from father to son) and on mitochondrial DNA (mtDNA is transmitted from mothers to their offspring of either sex). The research took DNA from people in three villages with matrilocal residence practices, where married couples take up residence in the wives' village, and compared this to three economically and ecologically similar villages with patrilocal residence practices, where couples live in the husband's village. The researchers predicted *apriori* that if culture influences genetic patterns, then the genetic diversity of mtDNA would be relatively greater in the patrilocal villages while the diversity on the Y-chromosome DNA would be greater in the matrilocal villages. Their findings strongly confirm the coevolutionary hypothesis.

These examples illustrate that, even on shorter evolutionary time scales (within the last 6,000 years for lactase), cultural evolution can powerfully influence genetic evolution. Unfortunately, evolutionary approaches to humans have been slow to recognize this and incorporate it into theory.

Summary and Onward to the Puzzle of Cooperation

Our goals in this chapter were threefold. First, we wanted to show that culture and cultural explanations have been incorporated into a fully