From: "Evolution and the Capacity for Commitment,"
edited by Randolph M. Nesse, Russell Sage Press, NY, 2001

Chapter 1

Natural Selection and the Capacity for Subjective Commitment

RANDOLPH M. NESSE

ONCE YOU recognize them, commitments are everywhere. The commitment we all know is marriage. By giving up the option to leave for someone else, spouses gain security and an opportunity for a much deeper (and more efficient!) relationship than would otherwise be possible.Yet many commitments are not at all nice. When John F. Kennedy committed the United States to eliminating missiles from Cuba, millions of human lives were put at risk. When a thuggish-looking man offers unrequested fire insurance to a small retail shop, the implied threat of arson may well influence the owner's behavior. In a more banal vein, nearly everyone has waited for a repairman who never arrives. From the personal to the political to the mundane, commitments are everywhere. Threats and promises change people's beliefs and, therefore, their actions.

In a world without commitments social exchanges arise mainly from helping relatives and trading favors. In such a world individuals can reliably be expected to act straightforwardly in their own interests. The advent of commitment changes everything. As soon as one individual finds a way to convince another that he or she will act other than in simple self-interest, social life is transformed. Now, individuals must consider the possibility that others may fulfill promises and threats. Sometimes the situation itself makes fulfillment worthwhile, but apparently irrational commitments also may influence others profoundly. A whole new category of social influence emerges.

Reputations become important predictors of behavior, and people begin spending vast amounts of effort to convince others that they will fulfill their promises and threats.

One way to convince others of one's credibility in a commitment is to give up options in order to change the incentives in a situation. Others observe that compliance is now in the person's interests, and change their expectations and behavior accordingly. By signing a lease, for instance, a person gives up the option to leave on short notice, and induces the landlord to give up the option of raising the rent in the middle of winter. Not all commitments, however, are backed by such tangible incentives. As Frank and Hirshleifer argue (Hirshleifer 1987; Frank 1988), people also can convince others that they will keep their promises by emotional displays that testify to their irrationality. Such commitments give rise to profound paradoxes. In order to influence others an individual must convince people that he or she will act in ways that are not in his or her interests. Sometimes it is possible to reap the benefits of this influence without having to fulfill the commitment. Kennedy did not want a fight, he simply wanted the missiles out of Cuba; because the Russians believed he would follow through on his threat, he didn't have to. More often, however, convincing others requires fulfilling the commitment to some degree. If your fiancée gets sick on the day you have tickets for the big game, too bad, you have to stay home. Also, once commitment strategies are widespread, reputation becomes so valuable that maintaining it requires fulfilling some commitments irrespective of the effects. The gang leader who rules by ruthlessness cannot show mercy for fear of losing his power. Before you know it, commitments lead people to do all kinds of things they would rather not do, whether this means carrying out spiteful threats or helping others who will never be able to reciprocate. Social life becomes rich and complex. Foresight and empathy become essential tools for social survival. In a world where commitments influence behavior a workable theory of mind becomes more useful than a sharp stone axe.

The capacity for using commitment strategies effectively is so important that natural selection may have shaped specialized capacities to make this possible. These capacities may help to explain the human tendency to emotionally extreme behaviors that often seem senseless. Such explanations may help to bridge the gap between the social sciences and evolutionary approaches to behavior. Even more important, they may help us find ways to cope with the great moral and intellectual crisis precipitated by the discovery that natural selection works mainly at the level of the gene, and its corollary, that natural selection cannot shape a capacity for altruism that gives a selective advantage to genes different from one's own This is turning out to be a major

psychic trauma to humankind, one that may be even more disturbing than the previous two: while the Copernican revolution shook religious cosmology off its foundations, it did little to challenge people's belief that human life has special meaning as the culmination of a divine plan; the second trauma—the discovery of natural selection and our evolutionary origins—was more personal (Richards 1987; Cronin 1991). Finding out we are only one species among many, all shaped by the mindless force of natural selection, fundamentally threatens our sense of the significance of human life (Ruse 1989; Dennett 1995). Resistance to this fact remains passionate, and creates an emotional fissure that still ruptures the political and social landscape.

Even as we struggle to accept the facts about our evolutionary origins, however, we confront a third trauma, this one a more direct threat to our individual moral identities. For decades biologists complacently had thought that selection shapes traits that benefit groups and species (Wynne-Edwards 1962). This assumption made it easy to view self-sacrifice for the sake of the group as entirely natural and expected. With simple but ruthless logic, Williams showed in 1966 that selection at the group level is feeble compared to selection at the individual level (Williams 1966). Natural selection, it turns out, acts mainly to benefit genes and individuals, not groups or species (Maynard Smith 1964). Many implications follow from this, but the most profound is the transformation of altruism from a natural tendency into an evolutionary mystery (Dawkins 1976; Badcock 1986; Barash 1977; Krebs 1970). E. O. Wilson called altruism "the central theoretical problem of sociobiology" (Wilson 1975, 3). Previously, animals were thought to help each other because natural selection shaped behavioral tendencies to sacrifice for the good of the group. It is now clear, however, that genes that lead to sacrifice for the group tend to become less and less frequent, except in very special circumstances. Any natural tendency to help others now must be explained in terms of how it benefits the actor's genes. If no such explanation can be found, the tendency becomes an anomaly in need of special explanation. Notice that the object of explanation here is a tendency, not a behavior. Individuals may decide to behave in ways that are not in their reproductive interests, and social structures may foster many such behaviors. Yet all tendencies shaped by natural selection must provide an inclusive fitness benefit in the long run, otherwise they will be eliminated.

Most people think of altruism as a costly effort that helps others. Many people also identify with their genes. Learning that organisms are shaped to act in ways that benefit their genes can make people feel that helping their relatives is somehow helping themselves and, therefore, not generous in the same way that helping a nonrelative

would be. From this viewpoint, genuine altruism can appear impossible, or at least contrary to nature (Richards 1993). Furthermore, many people, on grasping that all organisms are necessarily designed to act in the interests of their selfish genes, instantly (and incorrectly) conclude that individuals must be selfish by nature. Some take this a step further, concluding that our moral passions are mere pretensions or, worse, self-deceptive strategies for manipulating others (Ghiselin 1969).

Of course, individuals do help each other. To explain this, Williams's insight was quickly complemented by the recognition of two specific ways in which helping others can benefit the helper's genes. The first is kin selection. As Hamilton pointed out, because related individuals share a proportion of genes that are identical by descent, natural selection can shape tendencies to act in ways harmful to the self if there is enough benefit to kin (Hamilton 1964). Thus, natural selection can increase the frequency of a genetic tendency that makes a mother blue jay risk her life to defend her eggs because of the benefits to identical copies of the same genes that exist in other individuals (who are not yet hatched). The second way in which altruistic acts can offer benefits is by reciprocal exchange. As Trivers made clear, trading favors can yield a net reproductive benefit to both parties (Trivers 1971). In the long run mutual helping gives a net payoff, so long as one avoids being exploited. On these two pillars, a new theory of sociality is being constructed (Trivers 1985). Every social tendency has been attributed to benefits from some combination of kin selection and reciprocity. Those that cannot be explained in these terms have become anomalies that require interpretation as abnormal behavior, or products of manipulation or socialization in our novel modern environment.

The scientific impact of these developments has been enormous. They have transformed social ethology from a descriptive science to a predictive one based firmly on evolutionary theory (Trivers 1985; Alcock 1997). Hundreds of studies now investigate the role of kin selection in behaviors ranging from mating strategies to food sharing and defense (Alcock 1997). Tom turkeys court females in cooperative groups; it turns out that the groups are almost always composed of brothers. The sentinel prairie dog that warns the group about approaching coyotes is especially likely to have many relatives in the group. Human infants are eighty times more likely to die from child abuse if they are in a family with a stepparent (Daly and Wilson 1987). Kin selection is one of the great discoveries in our time.

The principle of reciprocity is an equally powerful advance for explaining social behaviors among unrelated individuals (Trivers 1985; Cronin 1991). Examples range from mating alliances between male

chimpanzees to blood sharing in vampire bats (Dugatkin 1997). Individuals who trade favors judiciously do better than those who go it alone. Often, of course, individuals trade favors with relatives, thus getting benefits via both mechanisms. Together, kin selection and reciprocity are widely thought to fully explain social behavior. They are certainly hugely important, but are they sufficient? Or, as many have suggested (Boehm 1999; Hirshleifer 1999; Humphrey 1999), might there be other routes to social behavior that have been neglected? Certainly there are. The tendency to use reciprocity to stand for all cooperative relationships is itself a vast oversimplification. Among other possibilities, cooperation also can arise from mutualism, coercion, and social organizations that control incentives. Mutualism, in particular, has been neglected. A great proportion of social cooperation is mutualistic; individuals get benefits only if they contribute, so cheating is not possible. (For more on this discussion see Adams in chapter 5.) Much additional cooperation is coerced; an individual with control over resources can impose punishments that make cheating of no value. These are among several routes to cooperation that are often neglected. Here we are concerned in particular with one additional means of social influence—commitment. Our goal is to examine how commitment strategies work, how they are used in practice, and the core question of whether the ability to make and assess commitments gives selective advantages that have shaped our minds. The importance of this core question, and its place in the explanation of social behavior, will be clear only after outlining the full magnitude of the current crisis and the responses it has provoked.

The sketch described here uses intentionally bold brushstrokes to illustrate the crisis at the intersection of science and morality created by the demise of group selection and its replacement by kin selection and reciprocity. Subtleties and caveats abound, but here they are set aside intentionally to highlight the dark simplicity of the problem, summarized best in Dawkins's phrase "the selfish gene." He knows perfectly well that genes have no motives, but his metaphor of genes as self-serving agents has enabled thousands of readers to view life from this dramatically new perspective (Dawkins 1989). We are jerked to attention by picturing ourselves as lumbering robots acting at the behest of our genes. Given how natural selection works, we know that the genes that influence behavior in any species are those that have given rise to actions that tend to increase the numbers of copies of those genes in future generations. All evolved tendencies, including those motivating generosity and morality, somehow must have increased the frequency of the genes that give rise to them. The extrapolation that individuals must therefore be selfish, however, is incorrect. Indeed the very existence of sympathy and moral passions

suggests that they likely give a selective advantage; yet how do such traits advance the interests of an individual's genes? One possibility is by improving the individual's capacity to benefit from commitment.

Is this issue really so dramatic? Is all this talk about psychic traumas and moral crises just the latest attempt by academics to inflate the importance of their arcane arguments about human nature? I think not. What we believe about ourselves and human nature is important because it influences how we act (Beck 1976). Those actions shape our societies that in turn shape our beliefs, thus setting long-running cultural cycles in motion (Fukuyama 1995). Evidence that much is at stake can be seen in the intensity of reactions to these ideas (Caplan 1978; Ruse 1982; Rose and Rose 2000). Many people find it repugnant to think that humans are inherently selfish, and despicable to think of altruism as just another way to serve one's interests. They assume that people who advocate such ideas must themselves be selfish and lacking in respect for society's rules. Attacking such apparently immoral nonconformists is sanctioned and even required by many social groups. Thus, criticisms of evolutionary approaches to human behavior often segue, without notice, qualms, or apologies, into ad hominem attacks against individuals and moral condemnations of whole groups. If the tone of recent letters in the *New York Review of Books* provides insufficient evidence, consider the subtitle of a new book: *Arguments Against Evolutionary Psychology*. (Rose and Rose 2000). The acrimony in this debate has created a dust storm of rhetoric and anger that obscures many legitimate issues. The battle has engaged the general public, whose insights often are considerable, thanks to the availability of accessible yet meaningful treatments of the main issues (Wright 1994; Ridley 1997; Wilson 1993). For professional audiences, the number of books and articles exceeds anyone's ability to keep up (Campbell 1975; Schwartz 1986; Ruse 1986; Alexander 1987; Oyama 1989; Maxwell 1990; Frank 1992; Nitecki and Nitecki 1993; Wilson 1993; Bradie 1994; Midgley 1994; Petrinovich 1995; Hurd 1996; Katz 2000).

Reactions to the Trauma

Responses to this psychic trauma take a variety of forms. A brief treatment cannot catalog all of them, nor can it adequately explain and justify any of them in full detail; it can, however, map the landscape of these controversies. Moreover, the intensity and diversity of reactions are testimony to the central role of the problem of altruism. A review of twelve kinds of reactions to this crisis will set the stage for further consideration of the role of commitment as a possible, albeit partial, solution.

- Many people take one look at this controversy and dismiss the idea that all motivational systems must benefit the actor's genes. The theory seems to be contradicted by too many examples. Soldiers dive on grenades to save their buddies. Enraged spouses murder their partners, knowing that this will mean decades spent in prison. Suicide bombers destroy themselves to advance their causes. Heroes dive into icy rivers. Ascetics meditate for years. Some people decide not to have children. Some even decide not to have sex. Others quit jobs to care for their disabled children or demented spouses. Anonymous benefactors give money to charity. Saints sacrifice their lives for their beliefs. Do such examples suffice to demonstrate that an evolutionary view is simply wrong? It is possible that we may attend to such examples precisely because they are so anomalous, but they still offer a challenge. They suggest that models based only on kin selection and reciprocity are inadequate; some other principle is needed.
- The crudest form of opposition is also the most prevalent—to deny the very phenomenon of evolution, or at least its application to humans. Fundamentalist Christian theology fosters pride in such denial. According to a 1999 Gallup poll, 47 percent of Americans believe "God created human beings pretty much in their present form at one time within the last 10,000 years or so" (Moore 1999). Scientific evidence has little impact on such beliefs, except to spawn more subtle and apparently scientific versions of creationism (Behe 1996; Dembski 1998) that are so insidiously confusing that they befuddle not only school boards but even some otherwise clear-thinking scientists. Conversations with creationists often reveal high motives, however; many sincerely believe that natural selection is incompatible with the possibility of human goodness and meaning in life.
- Others challenge the fundamental genetic perspective of evolutionary thinking. Oyama, in particular, views organisms as systems that sustain themselves by passing along various characteristics by diverse means, genetic material being only one source of information among many (Oyama 2000). Her developmental systems theory interprets individuals as products of overlapping cycles of gene-environment interaction far more complex than those usually considered. She wants to challenge what she calls the "central dogma" that phenotypes are genetically programmed, and to replace the nature-nurture dichotomy with "a radical reformulation of both." Interestingly, these views become the basis for an attack on the notion that individuals are naturally selfish (Caporael et al. 1989; Oyama 1989, 48). Whether these ideas simply lend themselves to such arguments or helped to motivate their development is hard to determine.

- Scientists do not have the option of denying the fact of natural selection, and most seek to understand nature through descriptions that increase simplicity, not complexity. Those who object can, however, try to blunt the impact of this new knowledge in other ways. One is to emphasize the limits of natural selection—its stochastic nature, its inability to create perfection, and the difficulties of studying adaptations (Gould and Lewontin 1979; Kitcher 1985). Such criticisms are certainly justified when applied to silly just-so stories and the uncritical and naïve view of adaptation that gives rise to them. Some critiques, however, tend to damn the whole enterprise of using evolutionary theory to understand adaptations in general, behavior in particular, and human behavior especially. As Wright points out, they also may give fuel to creationists (Wright 2000, chapters 18 and 19). Daunting challenges remain, however, including finding better criteria for determining what is an adaptation and what is not, and better understanding how natural selection shapes maladaptations as well as adaptations. These tasks are difficult, but by no means impossible (Williams 1992; Rose and Lauder 1996). Criticism of specific proposals is valuable, but even more valuable are well worked-out tests of adaptationist hypotheses.
- This brings us to objections based on admittedly political aims that sometimes come with the trappings of science (Caplan 1978; Segerstråle 2000). For instance, in critiques of so-called genetic reductionism, otherwise careful scientists create confusion by confounding questions about evolution and adaptation with quite different questions about genes and mechanisms (Gould and Lewontin 1979; Lewontin, Rose, and Kamin 1984). Closely related is the tendency to fail to distinguish evolutionary studies about the functions of behavior from fundamentally different behavioral genetic studies that seek to explain individual differences as a result of genetic differences. Such conflation fosters overtly political attempts to associate modern evolutionary approaches to understand human behavior with eugenics and the horrors of the Third Reich (Lewontin, Rose, and Kamin 1984). As the political motivations of these gambits are better recognized, they are having less and less effect, at least in the natural sciences (Segerstråle 2000). Furthermore, any association of evolutionary biology with conservative causes is dissolving because most in the field today lean toward the political left. Some even argue that evolutionary principles give direct support to such liberal goals as combating inequality, sexism, and racism (Konner 1999; Singer 2000). The general debate over the utility of evolutionary approaches to behavior is coming to a close. Sociobiology has demonstrated its utility and has developed into a strong new

branch of science (Alcock 2001), even if sometimes called by other names, even if proponents sometimes advocate unsupportable theories, and despite the difficulty of finding the best ways to formulate and test some specific hypotheses.

- Another way to attempt to escape the controversy is to try to go back, to resurrect the principle of group selection. One tireless, outgoing scientist, David Sloan Wilson, has made this his mission. In a series of articles, books, interviews, and lectures he has succeeded in making it appear that a mistake was made, that group selection is important after all (D. Wilson 1975; Wilson and Sober 1994). In fact, his model of trait group selection is considerably different from the original naïve models of group selection. Trait group selection for genes is certainly theoretically possible, and Wilson describes a few examples from nature, although others contest their significance (Wilson and Sober 1994). What is clear, however, are his motives. He and his colleagues want to find a mechanism by which natural selection can shape motivations for moral behavior (Sober and Wilson 1998; Smuts 1999). I share that motive and agree that some behaviors and emotions cannot be explained by kin selection and reciprocity alone; however, group selection is not the only possible alternative. I also agree that emergent properties of social groups are powerful forces of selection that are likely to explain important aspects of human nature, including our moral capacities. Along with most other scientists (Dawkins 1982; Williams 1992; Smuts 1999; Maynard Smith 1998), however, I remain unconvinced that differential success of human social groups is a viable potential explanation for the genetic tendencies that shape brain mechanisms that give us moral capacities. Furthermore, if group selection indeed were powerful, it would account equally well not only for cooperation within the group, but also for competition between groups—including tendencies to dehumanize, exploit, and kill members of out-groups, hardly typical examples of moral behavior. While attention to different levels at which selection acts is essential (Maynard Smith and Szathmáry 1995; Williams 1992), trying to resurrect the phrase *group selection* creates unnecessary confusion that distracts attention from trying to understand exactly how the emergent properties of social groups may have shaped the moral passions.
- A related explanation is better supported. In two decades of work Richerson and Boyd have developed models that show how cultural group selection can account for many complex traits, including cooperation (Boyd and Richerson 1985; Richerson and Boyd 1999; see also chapter 9). In their model these traits, not genes, give benefits to groups and therefore spread. The benefits these traits

bestow on the group can themselves influence the course of natural selection. Boyd and Richerson do not challenge the predominance of selection's power at the level of the individual; they instead show how cultural selection of socially transmitted traits can help to explain the rise and fall of groups and the nature of culture, and how emergent properties of social groups can become forces of genetic selection for at least two instincts, ethnocentrism and moralistic aggression (Richerson and Boyd 1999).

- Others present culture as if it were an alternative to evolutionary explanations. When pressed, they usually describe examples of people fulfilling the expectations of cultural roles, even when that seems to do little for their reproductive success. Or, they describe the vast variety of human behaviors in different cultures in order to show that human nature is profoundly malleable, that our behavior is not susceptible to explanation as a product of a few drives and fixed action patterns (Lewontin 1982). There is no disputing that humans live in grandly diverse ways, and that our capacity for learning allows us to adapt to vastly different environments. It remains an elementary mistake, however, to posit culture as an alternative to evolutionary explanations (Tooby and Cosmides 1989; Barkow 1989a). In fact, our capacity for culture has almost certainly been shaped by natural selection, and it seems likely that emergent properties of cultures are powerful forces of natural selection (Durham 1982; Tooby and Cosmides 1989; Cosmides and Tooby 1989; Tooby and Cosmides 1992; Flinn and Alexander 1982). Understanding culture is important, even essential, if we are to understand how cooperation is possible, as is well known by those who have developed models of gene-culture coevolution (Cavalli-Sforza and Feldman 1981; Lumsden and Wilson 1981; Durham 1991; Wilson 1998).
- Then there is human will. People can decide to do things that harm their reproductive success. They can decide to inhibit deep impulses. They can decide not to eat, not to sleep, to sacrifice their firstborn, or to be generous even when they are being exploited. This profoundly human capacity—a property of mental mechanisms shaped by natural selection—is an observable reality. It is admirable when it motivates helping, terrifying when it motivates enduring attempts to harm an enemy at all costs. Whether such a capacity offers the best explanation for altruistic behavior is far less certain. After providing extensive reviews of the distressing discoveries described here, several books on evolution and human behavior conclude with the hope that we can use this sad and frightening new knowledge about ourselves to make willful decisions to hold

back our primal tendencies in ways that can improve our societies and ourselves (Wilson 1978; Dawkins 1989; Wright 1994; Ridley 1997). This may be correct, yet there is a canyon of ignorance between this hope and realizing the dream of a more peaceful world.

- The distress aroused by these discoveries does not always result in opposition. Some, especially those whom William James would characterize as "hardheaded," look at the facts with an unwavering gaze and proclaim that genuine morality cannot emerge from nature, and we should get used to it. Barash, in particular, seems to have been deeply affected by the recognition that what had seemed selfless turns out to help one's genes (Barash 1979; Barash 1982). Williams, who was perhaps the first to confront these issues (Williams and Williams 1957), has even gone so far as to call Mother Nature a "wicked old witch," because all organisms are inherently selfish in the sense that natural selection has shaped them inexorably to do whatever maximally benefits their genes (Williams 1993). His position follows that of Julian Huxley (Huxley and Huxley 1989 [1893]), which in turn follows Thomas Henry Huxley's in his 1894 Romanes lecture (Huxley 1897). From this point of view, moral action is inherently in opposition to natural tendencies. While such views seem bleak, they often arise, it seems to me, from the same deep dismay that arouses others to oppose an evolutionary view.
- Others react by embracing and extolling the brutality of nature. The initial political offspring of Darwinian theory was, after all, Social Darwinism, with its grand leap from the facts of nature to advocacy of ethical principles that now are widely recognized as promoting vast social harm. The distinction between what is and what should be was trampled. Social Darwinism remains a signal example of powerful people using their resources to spread an ideology that benefits their interests. More recently, some with a cynical bent seem almost gleeful in their attempts to explain every apparently altruistic act as somehow deviously selfish. As Ghiselin puts it, "Scratch an altruist and watch a hypocrite bleed" (Ghiselin 1969, 247). Kohn cites Santayana as going even further: "Generous impulses are self-deceptive hypocrisies. Dig a little beneath the surface and you will find a ferocious, persistent, profoundly selfish man" (Kohn 1990, 205). Such statements make it clear that these issues are deep and emotionally charged. It should be no surprise that people's reactions are intense and markedly varied.
- Finally, one can respond to the difficulties by carrying out a careful analysis of exactly what we mean by altruism and whether it indeed follows that selfish genes necessarily make selfish individuals. Radcliffe-Richards has done just this, examining the flawed reason-

ing that concludes that a behavior is in the individual's self-interest just because it is in the interests of that person's gene (Radcliffe-Richards 2000). She goes further, still in a strictly gene-selectionist paradigm, to consider the philosophical problem of whether the altruism that is shaped by natural selection can be "*real* altruism." She notes that while natural selection cannot shape a completely altruistic organism, "*genuine* and *limited* concern are not opposed. . . . Coming to show how altruism comes to exist no more shows that it is not real altruism than explaining how a cake was made shows it is not a real cake" (170, 175). This analysis focuses attention where it belongs, on the question of whether natural selection could have shaped capacities that induce people to fulfill their commitments even when they would rather not. Even if such actions advance a person's genetic interests in the long run, they are examples of genuine altruism.

This concludes our summary of twelve responses to the debate concerning altruism precipitated by recognition that selection acts at the level of the gene. Their extraordinary range and intensity are as important as their content. These reactions demonstrate the remarkable impact and continuing controversy arising from the discovery that innate tendencies to altruism can exist only if they have somehow increased the frequency of genes that shape them. While there is no hole in this logic, it leaves many observations unexplained. People often follow rules, do their duty, remain loyal, keep their promises and fulfill their threats, even when such actions are clearly not in their interests. How can this be possible? Wouldn't it be better to decide what is in one's best interests at each stage of the game? Why keep a promise if it requires a big sacrifice with no payoff? Why follow a rule if violating it will bring no punishment? Something seems to be missing.

Commitment

Signing an apartment lease, threatening nuclear retaliation, taking religious vows, burning bridges behind you, changing lanes without looking, caring for a spouse with Alzheimer's, threatening arson to enforce an extortion scheme, and waiting for a tardy deliveryman—these all have something in common. All are social strategies carried out not because they bring benefits via kin selection or reciprocity, but because they influence others via commitment. A credible commitment changes other people's expectations, and thus their behavior. For the dictator of a small country, launching nuclear missiles would be suicidal. Nonetheless, if he can convince others that he will really

do it, they will be extraordinarily cautious about violating his borders. A renter who signs a lease gives up the option to move out on a whim, but gains a secure place to live at a stable rent. An army that has burned its bridges motivates itself and intimidates its foe. Taking religious vows of celibacy certainly is usually not in one's reproductive or personal interests, but priests do it, and we therefore treat them differently. The taxi driver who changes lanes without looking will certainly cause an accident if others do not yield. So they do. It is one thing to promise to care for each other in sickness and in health, but when one spouse succumbs to Alzheimer's, what can be the sense of keeping a commitment to someone who does not even recognize you? Yet some do. As for the mob, it has to convince business owners that their businesses will burn if they do not pay protection money. To make the threat believable, the mob must, on occasion, commit arson.

Note that the initial focus has shifted. The discussion began with the challenge of explaining altruism but moves now to commitment, which covers not only altruistic promises to help, but also terrible threats of harm. That threats and promises influence by different means is an argument for treating them separately, yet commitment binds them together as two ways of influencing people by changing their expectations.

In a series of books and articles starting in the 1960s Thomas Schelling laid out the framework for understanding commitments (Schelling 1960, 1978a, 1978b, forthcoming). He was preoccupied with cold war strategy, but he extended his analysis to personal relationships and problems of social life in general. This work has influenced many, not only economists but also philosophers who see that it can help to solve some long-standing puzzles involving morality (Parfit 1984; McClennen 1990).

Commitment is fundamentally different from kin selection and reciprocity. Kin selection changes behavior toward relatives owing to the benefits of helping those who share genes identical by descent. More a phenomenon than a social strategy, kin selection explains why organisms help kin even when an action is costly and gives no direct payoff. Reciprocity is based on the expectation of a net benefit from the actions of exchange partners who also act in their own calculated best interests. Commitment, however, is different—it changes behavior by giving up options and thereby changing people's beliefs.

A commitment is an act or signal that gives up options in order to influence someone's behavior by changing incentives or expectations. Some commitments change the objective situation so that fulfilling them becomes in a person's interests. Other commitments are pledges

to act in ways that will be contrary to obvious self-interest; they are not enforced by external incentives, but by some combination of reputation and emotion.

Examples are everywhere. If you believe your spouse will care for you in sickness, you will be much more likely to make and keep the same promise. If you believe your spouse will kill you if you try to divorce, you will be much more reluctant to leave. If either unfortunate circumstance arises, it will be disadvantageous to make good on these commitments. Caring for a sick spouse can sap life's energies; murder is likely to result in a long prison term (to say nothing of guilt). Yet people do make such commitments, and others believe them and are influenced because people often fulfill commitments even when doing so is costly. In the long run, however, a capacity for making commitments brings net benefits. These benefits may explain how natural selection shaped tendencies to follow rules and fulfill promises and threats even when that is not in a person's interests. In short, commitment may offer some of what has been missing from an evolutionary understanding of social behavior—a potential means by which natural selection could shape mental faculties for genuinely moral (and immoral) action.

This use of the word *commitment* follows its technical meaning in economics and game theory; other uses need to be distinguished. Sometimes people say they are committed to having dinner tonight or to going to a movie. These are only announcements of plans, not commitments in the sense used here. As Hirshleifer puts it, "A promise or a threat must be to do something that the individual would not otherwise be motivated to do. That is what distinguishes these pledges from mere forecasts" (Hirshleifer 1987, 309). Similarly, people sometimes say they have a commitment at two o'clock. Such obligations that a person cannot easily abrogate are related to the technical sense of commitment in that options are foreclosed, and many such commitments are intended to influence the behavior of others. If you make a commitment to meet others for lunch, they will likely be there.

In psychology commitment often refers to the strength of a person's determination to pursue a goal (Brickman and Coates 1987; Klinger 2000). The goal may be a specific task or a desire to change relationships, religion, or employment. This concept of psychological commitment is also similar to the meaning considered here; both indicate that a person will follow through to reach a difficult goal. The game theory sense of strategic commitment, however, emphasizes the influence on others, while commitment as studied in psychology more often refers to personal commitments individuals make to control their own future behavior—a commitment to lose weight, for in-

stance, or to write for three hours each morning. One strategy for controlling one's behavior is to change the incentive structure by giving up options so that short-term pleasures are set aside to achieve long-term gains. Some people keep their cigarettes in their car, where they cannot easily reach them; others mail their own computer cables to themselves to enjoy a weekend free from Internet addiction (Burnham and Phelan 2000). Gibbard sees such examples as evidence for a system "peculiar to human beings" of normative motivation that can conflict with what he calls the "animal control system" (Gibbard 1990, 57). Such internal mental conflicts, usually between short- versus long-term benefits, have long been recognized by psychoanalysts, although they tend to emphasize the mental conflict and neglect the real strategic dilemmas that give rise to them (Nesse 1990b). Brickman and colleagues have conducted a provocative early exploration of the interfaces between these several aspects of commitment (Brickman and Coates 1987), with particular attention to the utility of irrationality: "Behavior that is not rational, behavior that is chosen on logically inadequate grounds or carried out with utter disregard for costs, can still be functional and effective for the actor. . . . The apparent blindness that follows commitment may be irrational and still a precondition for effective action" (35).

Within the game theory concept of commitment addressed here, several distinctions can help to separate different subtypes. Some commitments are promises to help, others are threats to harm; the difference is whether others perceive a new potential benefit or a new danger. Threats and promises are treated here as two faces of commitment, although an argument could be made that they should be considered as independent phenomena. Commitments can be conditional or unconditional. Threats are usually conditional attempts to prevent another person's potential action; promises are more likely to be unconditional, though often they are conditional on the other fulfilling a matching promise. Commitment can be intentional or unintentional; some commitments are entered into specifically to influence people, others are not. Schelling's chapter describes these varieties in vivid detail.

Perhaps the most profound distinction between different kinds of commitments arises from how they are enforced—that is, why others believe the commitment will be fulfilled. Sometimes the commitment itself makes some options impossible (for example, burning bridges or being bound to the mast). In other situations the options remain available but are no longer advantageous after a third party is given control over tangible incentives (a lease, for instance). A pledge of reputation, such as a public oath, is another way to convince people

that a commitment is valid. Finally, commitments can be enforced by emotions such as pride and guilt. Observing such emotions may increase confidence in a commitment.

The first two kinds of commitments—those enforced by the situation itself or by third parties—can be thought of as "secured" because once such a commitment is made, its fulfillment becomes in the actor's interests. Many commitments are, however, unsecured by such tangible enforcers—their fulfillment depends on emotions and concerns about reputation; we will call these subjective commitments. While reputational commitments put real social resources at risk, they also are strongly enforced by emotions such as pride and guilt, so it seems appropriate to include them under subjective commitments. The fulfillment of externally enforced commitments also involves emotions, yet the emotions involved are not these special moral passions but rather the usual emotions that regulate goal pursuit (Nesse 1999). Many actual situations incorporate several of these mechanisms simultaneously. Marriage, for instance, is enforced by a mix that includes legal contracts, reputational pledges, and emotional bonds. Thomas Schelling has distinguished the subtleties of various kinds of commitments. He points out, for example, that belief in a deity who will reward goodness and punish evil transforms many situations from subjective to secured, at least in the believer's mind (Schelling, personal communication, 2000).

Subjective commitments are of special interest because they influence others only by convincing them that an individual will, in some specified future situation, act in ways not in his or her interests. The possibility that natural selection has shaped emotions to help to guarantee subjective commitments was brought to wide attention by Robert Frank's book *Passions Within Reason* (Frank 1988) and by Jack Hirshleifer's chapter "Emotions as Guarantors of Threats and Promises" (Hirshleifer 1987). Each recognized that the utility of commitment strategies could have provided a selection force that shaped emotions such as anger, trust, pride, and guilt, and that the very existence of these emotions might be evidence for this hypothesis. Hirshleifer and Frank summarize their positions in very similar terms:

> It is possible to analyze, in terms of effects upon rationally calculated self-interest, the consequences of non-self-interested motivations and of limitations upon the ability to calculate. The economist must go beyond the assumption of "economic man" precisely because of the economic advantage of not behaving like economic man—an advantage that presumably explains why the world is not populated solely by economic men. (Hirshleifer 1987, 322)

> The commitment model is a tentative first step in the construction of a theory of unopportunistic behavior. It challenges the self-interest model's portrayal of human nature in its own terms by accepting the fundamental premise that material incentives ultimately govern behavior. Its point of departure is the observation that persons *directly* motivated to pursue self-interest are often for that reason doomed to fail. They fail because they cannot solve commitment problems. These problems can often be solved by persons known to have abandoned the quest for maximum material advantage. The emotions that lead people to behave in seemingly irrational ways can thus indirectly lead to greater material well-being (Frank 1988, 258).

This insight is profound. It means that the rational pursuit of self-interest is sometimes an inferior strategy. People influence others by convincing them that they will do things that would not otherwise be in their interests. This is the essence of intimate social life. The social fabric has a warp of promises and a weft of threats. We spend our lives making and assessing commitments. We are constantly preoccupied with assessing other people's reputations and protecting our own. We are enraged when others do not fulfill deep commitments. Most of us are a bundle of anxieties when tempted to renege on a pledge. Elster (who remains unconvinced) points out that Frank and Hirshliefer have turned the usual idea of emotions on its head (Elster 2000). Instead of interfering with rational strategy they claim that emotional behaviors can be strategically superior to those based on rational calculation. If it turns out to be correct that natural selection has shaped emotional mechanisms for mediating strategic commitments, this will be of enormous importance. This book is devoted to further developing this idea and trying to determine where it does and does not apply.

To summarize, there are four main reasons to believe that a commitment will be fulfilled:

1. a commitment can be *self-enforcing* if it is secured by incentives intrinsic to the situation;
2. a commitment can be secured by *external incentives* controlled by third parties;
3. a commitment can be backed by a pledge of *reputation*; and
4. a commitment can be enforced by *internal emotional motives*.

In a sense these are four different kinds of commitments: *self-enforcing, contractual, reputational,* and *emotional*. As already noted, all four factors may influence a single commitment, so it is probably better to

think of them as factors rather than distinct kinds. Note that the first two types of commitment are secured, in much the same way that a loan is secured by collateral. Making the commitment changes the situation so that fulfillment becomes in the individual's interests. Such commitments give up options in ways that change the objective nature of the situation. The latter two types of commitment are quite different: they do not change the objective contingencies, only what a person says he or she will do; they are subjective commitments in that they involve a continuing option for reneging. This may be why there is a special need for emotions to maintain motivation to fulfill such commitments, and to try to convince others that the commitment will be fulfilled.

The Significance of Commitment

The importance of commitment is no doubt becoming clear, but a few points deserve emphasis. Benefits that come from an ability to make and assess commitments could shape mental mechanisms that induce individuals to follow rules, even when that is not in their interests. This requires the peculiarly human ability to inhibit short-term self-interested action, offers a potential explanation for the origins of motives for moral behavior, and suggests that the apparently irrational excesses of both love and hatred may emerge from the same source. Commitment offers a framework that engages much of the richness and complexity of human relationships, and it gives genuine moral capacities a place in human nature. This gives it the potential for decreasing the stigma associated with pursuing evolutionary studies of human behavior,thus allowing those studies to proceed with new understanding and support.

Commitment even offers a potential link with much recent social science based on postmodern theories. Criticism of the excesses of such thinking has obscured the value of recognizing how we create our own social environments. Our commitments change other people's expectations about what we will do. Although there is usually nothing coldly calculating about them, our commitments can be seen as strategies for creating beliefs in others. Taken together, these beliefs are the fabric of culture. If our capacity for inducing and assessing such subjective beliefs has been shaped to some degree by natural selection, this may offer a much-needed bridge between fields that otherwise seem determined to ignore or misunderstand each other. While much progress has been made in understanding culture in an evolutionary context (Barkow 1989a, 1989b; Boyd and Richerson 1985; Atran 1990; Sperber 1996; Diamond 1997), much remains to be done to find the mechanisms that make culture possible (Flinn and Alex-

ander 1982; Dunbar, Knight, and Power 1999) and the extent to which these mechanisms have been shaped by natural selection.

Similarly, commitment can offer insights into the origins and functions of religious beliefs. One main function of religious groups is to systematically cultivate and reinforce beliefs that make commitment more accessible and less risky. Many religions insist vehemently that their beliefs are not intended to bring gain, but to foster goodness, at least toward members of the church. Such explicitly subjective ideologies that disavow any motive for individual gain may provide stronger foundations for deep commitments than any appeal to rationality or self-interest (Irons 1996). When a religious group is difficult to enter, and involves most of a person's relationship partners, all of whom monitor one another's behavior, the cost of defecting can become so high that what began as subjective commitments become strongly secured. Religious communities, even more than other groups, provide members with protections that allow commitments to flourish (Dennett 1995; Wilson and Sober 1994). This may help to explain why so many religious groups oppose an evolutionary approach to human behavior, and perhaps why some Islamic cultures take such desperate measures to inhibit the importation of Western influences.

The major Western religions, according to Rue, offer commitment on a grand scale (Rue 1989). Worshipers enter into a covenant with God, agreeing to submit and obey unconditionally in return for a better life and, often, eternal life. The deepest blessings of such systems are, Rue says, in the mental changes that come from living in a world imbued with meaning and order, and the social benefits of living among other people who follow the rules. He contrasts this sharply with the alienated and meaningless worlds inhabited by so many nonreligious people in modern industrial economies (Klinger 1998; Emmons 1999). As noted earlier, if entering a religious group has high opportunity costs and requires major continuing investments of time, effort, and money, then the cost of leaving the group will be so substantial that the threat of expulsion is a potent enforcer of group norms. This is a classic example of commitment—giving up options in order to influence others in a way that changes their behavior and yields net benefits. Almost any emotional or mystical justification for the rules will do, but as soon as calculated self-interest comes to the fore, subjective commitment is fatally undermined.

Finally, as becomes especially clear from practicing psychiatry, our individual social worlds are much of our own making. Individuals who believe that everyone is out for himself or herself are incapable of making subjective commitments; their social worlds are populated only by exchange partners whose motives are always suspect. By con-

trast, individuals who can believe in other's subjective commitments reap the benefits of relationships that go beyond mere reciprocity. They also run the risk of major betrayal. Frank has conducted a series of studies to show that learning economic theory decreases contributions to public goods, although how widely the findings can be replicated and generalized remains unclear (Frank, Gilovich, and Regan 1993). One presumes that the same would be true for teaching about social evolution, but the study has not been done insofar as I know. The novelist A. S. Byatt recently looked at the conflict between social science and evolutionary science and concluded that it arouses such intense passions because we care so much about our fundamental human nature (Byatt 2000). We intuitively know that what we believe about human nature profoundly shapes how we conduct our lives and the structure of our societies. The belief that all behavior is fundamentally selfish itself changes behavior. As Oyama puts it, "Assuming cooperation to be, at best, a competitive strategy can make it conceptually unstable" (Oyama 2000, 207). It seems possible that the spread of simplistic notions about the evolutionary origins of social relationships could make individual relationships more conflicted and society even more brutal. One antidote may be an evolutionary approach to behavior that incorporates a capacity for commitment.

Background

The concept of commitment, as developed by Frank and Hirshleifer following the lead of Schelling, emerges from a long line of related thinking. Actually the precursors are better described as several lines, since many of them seem to be unaware of each other. This section outlines some of these ideas, starting with those from economics and game theory, and moving to those from philosophy and ethics.

The initial application of game theory to animal behavior by John Maynard Smith (1982) developed in parallel with the application of game theory in economics (Von Neumann and Morgenstern 1944). The possibilities for synergy seem substantial, especially since economists have long restricted their studies to a species well-suited to such explanations, *Homo economicus*. As Kohn puts it, "Egoism is not *an* assumption but *the* assumption underlying neoclassical economics" (Kohn 1990, 185). Such individual utility maximizers are little different from the animals studied by evolutionary ethologists. There is no room for genuine altruism, spiteful threats, or other such irrationality here. People are expected to do whatever best advances their own interests. This grand assumption generally works very well indeed. Economists claim to be able to predict not only what specimens of *Homo economicus* will buy, but also which ones will join a labor

Table 1.1 Social Emotions and the Prisoner's Dilemma

	Other Cooperates	Other Defects
You cooperate	Trust or friendship	Suspicion or anger
You defect	Anxiety or guilt	Rejection

Source: Nesse 1990a.

union, when they will choose to retire, and even whom they will marry (Becker 1991).

Studies of reciprocity relationships have prospered in this framework. The usual model is the prisoner's dilemma, a game in which two players each choose, on every move of the game, to either cooperate or defect. Mutual cooperation brings the greatest net benefit (say, three points each, for a total of six), but an individual who chooses defection gets a greater reward (say, five points) if the other chooses cooperation (and gets zero). This creates a continuing tension between a wish to get steady maximal mutual rewards, the temptation to defect, and the suspicion that the other might defect. Hundreds of experiments have confirmed that this is quite a good model for much cooperation. The strategy of doing whatever the other person did on the past move (tit for tat) has proved a robust strategy (Axelrod and Hamilton 1981; Axelrod 1984) that allows individuals to avoid exploitation while taking advantage of cooperation when that is available. Much work now compares tit for tat to other strategies, some of which are superior in some circumstances. Playing this game well is essential for members of a species where fitness depends on reciprocal exchange. As seen in table 1.1, the four situations defined by the four boxes in the prisoner's dilemma have arisen so often in the course of evolution that they seem to have shaped emotions specific to each situation (Trivers 1971, 1981; Gibbard 1990; Nesse 1990, 1999).

People do not, however, always do what is predicted by rational choice theory (Gintis 2000a, ch. 11). They do not always pursue straight self-interest, and their decisions are not always Bayesian rational. In the prisoner's dilemma, for instance, the game theory optimum for any limited sequence of plays against a game theory–rational opponent is to defect right from the start. This is because defection is the optimal last move of any finite game, and cooperation unravels backward from this endpoint. Yet that is not what people do. Most cooperate early on in the game, and stop cooperating only if the other defects or a specific endpoint is near (Axelrod 1984). Likewise, in public goods games, people cooperate vastly more than expected. According to game theory, an individual's best strategy is to accept the advantages while letting others contribute money to public radio, blood to the

blood bank, and land to preservation trusts. Actual human behavior is different. Gintis provides a comprehensive review, emphasizing, *"In neither the everyday nor the narrower economic sense of the term does rationality imply self-interest"* (Gintis 2000a, 243). Even in the laboratory people don't act according to the dictates of rational choice theory. When each subject contributes anonymously to a pot that is then sweetened by a percentage and then divided equally among all players in the game, the average contribution is about half of an individual's reserves in the early moves of the game. The individual variation is huge, however, from large contributions to none. When it becomes obvious that some individuals are not cooperating, contributions taper off and usually end. When individuals can find out who is not contributing, some will spend their resources to punish such individuals, even though they get no individual benefit from such actions (Fehr and Gächter 2000). The suggestion has even been made that complex human organizations became possible only when the ability to throw projectiles precipitously reduced the "cost of punishment" for coalitions trying to enforce their rules (Bingham 1999).

In the ultimatum game one person proposes how to divide a resource (say, ten dollars), and a second person decides either to accept that split or turn all the money back to the experimenter. The Subgame-Perfect Equilibrium is for the proposer to offer the smallest possible amount and for the responder to accept anything offered. A penny is better than nothing, after all. The corresponding game theory–rational strategy for the individual who proposes the split is to offer one penny to the other player. Yet this is not what most university student subjects do. Instead they tend to offer something close to an even split. And the recipients of a low offer often choose to forgo all benefits, thus getting spiteful revenge (Roth et al. 1991; Fehr and Falk 1999). Interestingly, there are wide cultural variations in patterns of response, with most individuals in some groups accepting low offers (Henrich 2000). Recent work has shown that men with low testosterone are more willing to accept an unfair split, while those with high testosterone are likely to turn down anything but a nearly fair division (Burnham, forthcoming.). Such tendencies seem closely related to the norms in honor-based societies, where even the slightest insult may lead to threats of violence (Nisbett and Cohen 1996; see also Cohen and Vandello, chapter 8 herein). Such threats signal an intention to defend resources from others who try to take them, even if that means accepting a huge cost. The alternative—fighting only when it is rational to do so—yields ridicule and impoverishment. Males who are not "demonic" may do poorly (Wrangham and Peterson 1996).

Most attempts to explain the anomalies in behavioral economics

are basically within the framework of rational choice theory (RCT) and its two assumptions: that people ultimately pursue self-interest exclusively, and that self-interest is maximized by acting on the basis of rational calculation. Binmore, for instance, emphasizes how difficult it is to escape the implications of RCT. He observes that people only rarely base their behavior on predictions that others will act in ways that are not in their self-interest (Binmore 1994), and he will have none of Gauthier's concept of constrained maximization (Binmore 1993). Likewise, Skyrms at first seems to have little sympathy for attempts to transcend RCT, saying, "A strategy that is not modular rational in these terms is just one that in certain circumstances would require such a rational agent to choose what she would not choose. . . . If expected utility theory is kept in mind, the idea of modifying the normative theory by somehow building in commitment appears quixotic" (Skyrms 1996, 41). He avoids contradiction by incorporating emotional outcomes as a part of utility, citing Frank and Hirshleifer. He proceeds to demonstrate how "Fairman" strategies can persist in equilibrium with other strategies, even when they might not survive otherwise, and concludes, "Evolution may—if the conditions are right—favor commitment over modular rationality strategy. Mixed populations that include individuals using strategies that are not modular rational in Darwinian fitness can evolve according to the replicator dynamics" (42).

Such models assume rational self-interested agents and pose the problem in terms of finding situations in which tendencies to cooperate can survive. Much depends on the details of the model. Axelrod has shown how the initial proportion of cooperators and defectors are powerful determinants of whether a group goes to increasing cooperation or devolves into constant defection (Axelrod 1986, 1997). Yet even here there is a stochastic element to the outcome, which is stable once it tips in one direction or the other. In actual human groups, as managers know all too well, organizational cultures have enduring power far beyond that of any individual.

Another line of work has emphasized the role of reputation. Alexander was one of the first to elaborate this in an evolutionary framework (Alexander 1982, 1985, 1987). He noted how hard it is to explain generalized cooperation and concluded that it may arise mostly from indirect reciprocity. This works because people observe whether or not others are cooperators and seek to create relationships with those who appear generous. In his framework, the optimal strategy is to appear very generous while actually contributing as little as possible. This leads to substantial emphasis on deception and self-deception, including the idea that our experiences of ourselves as moral agents is a self-deception that allows us to better deceive others (Alexander

1974; Trivers 1976; Alexander 1987; de Waal 1984; Mitchell and Thompson 1986; Lockard and Paulhus 1988; Nesse and Lloyd 1992).

Akerlof challenges the "core assumption of economic behavior . . . that all persons are totally selfish. . . . [T]his assumption is made for reasons of convenience, not because economists empirically assume that all persons act only out of selfishness" (Akerlof 1983, 54). He presents a model showing "that honest and cooperative behavior pays off; the honest person is not just a systematic 'sucker'" (55). As an example he cites the rewards of being a "bonded" courier. While the need for such bonding demonstrates the need for external enforcers to motivate honesty, Akerlof claims the system can lead to genuine honesty because "the *appearance* of honesty and class loyalty are beneficial; the easiest way to achieve these appearances is to *be* honest and loyal, even though honesty and loyalty themselves involve sacrifices" (61). Thus his model is parallel to Alexander's but emphasizes the benefits of integrity despite its costs, instead of the benefits of deception despite its risks. Note the central role that reputation plays in this argument.

Kitcher takes a very similar position, arguing that morality can persist in the face of cheaters because individuals can choose those with whom they play prisoner's dilemma games. "Discriminating altruists" eschew contact with defectors, to the substantial cost of those known to be defectors. His mathematical model demonstrates the viability of this strategy, and thus, presumably, how it could have been favored by natural selection (Kitcher 1993). These selective associations are the same engine that allows moral tendencies to prosper in Wilson's trait groups.

More recently, Nowak and Sigmund have developed a computer simulation that demonstrates how tendencies to provide benefits via indirect reciprocity can evolve even when the players meet each other only once (Nowak and Sigmund 2000). This is possible when information about an individual's cooperative "image" is public and when a reputation as a cooperator brings increased benefits via indirect reciprocity. They note, parenthetically and profoundly, "The evolution of human language as a means of such information transfer has certainly helped in the emergence of cooperation based on indirect reciprocity" (576). This work has been replicated and extended with experiments on humans (Wedekind and Milinski 2000; Nowak and Sigmund 2000). Reputation proves crucial to making indirect reciprocity viable. A capacity for commitment allows individuals to act in ways that reap the benefits of image scoring, and to avoid the costs of being recognized as an exploiter. In this and similar models individual agents generally are assumed to pursue one strategy. The strategy

may take account of the reputations of other individuals, but it has trouble coping with social groups in which individuals act differently toward each other based on a long history of prior interactions.

Clearly, reputation is a major mediator for the benefits of indirect reciprocity. So long as each player's reputation is revealed sufficiently, cooperation can grow. The optimal cost to spend on monitoring others depends on the ratio of cooperators to defectors in the population. In intermediate proportions monitoring allows cooperators to exclude defectors and to benefit from exchanges with one another. If either type is too rare, the costs will not be worth it. On a grand scale this has been used as the foundation for a theory of sociopathy as a frequency-dependent strategy that succeeds when it is rare in comparison to the proportion of cooperators who, when prevalent, let down their guard (Mealey 1995).

Elster emphasizes that rationality often requires taking steps to make precommitments in anticipation of temptations that will yield a short-term gain and a long-term cost (Elster 1979). By "depositing one's will" one gains global maximization. While he emphasizes the limitations of the human ability to resist temptation, and the workaround of making temptations inaccessible or ultimately more painful than the reward, he also describes strategies of building up one's character that make this much more feasible and yield moral behavior (Elster 2000). He sees perfect rationality as supporting something like Gauthier's constrained maximization, but he doubts that many humans have sufficient willpower to use such strategies.

Gintis defines *strong reciprocity* as a tendency to follow norms that promote cooperation with cooperators, and punishment of cheaters even when those acts are costly. He contrasts this with *weak reciprocity*, which is reciprocal altruism (Gintis 2000b). Strong reciprocity produces fewer replicas in the next generation, but the strategy can still invade and persist because those who violate the norms are recognized and excluded from the benefits of being in groups with strong reciprocators. This is not exactly group selection, because it does not depend on the relative success of different groups. It is, however, social selection, because emergent properties of groups become forces of natural selection that change the prevalence of genotypes, and thus individual behavior. He summarizes many laboratory experiments demonstrating otherwise unaccountable cooperation and presents a formal model demonstrating the viability of strong reciprocity. The key is that "defectors are excluded from participation in the community for a number of periods just sufficient to make defecting a suboptimal strategy, at zero cost to the community" (Gintis 2000a, 273). Many examples in the ethnographic literature also support the thesis

that tendencies to strong reciprocity evolved via this route. His position is summarized by his preference for viewing humans as *Homo reciprocans* rather than *Homo economicus*.

Strong reciprocity is very nearly the same as subjective commitment. The main difference is that strong reciprocity models place less emphasis on individual communications of threats and promises and the associated complexities of relationships and signaling, and instead focus on groups that monitor deviations from norms and exclude nonconformists. Strong reciprocity may be possible because individuals have capacities for subjective commitments to enforce group norms, even when that is costly. One could equally well explain the evolution of the capacity for commitment as a result of the costs of being excluded from a group of cooperators, especially if you consider groups with only two members. As Gintis notes, his position is presaged by Campbell, who said in 1983, "mutual monitoring forc[es] altruism on group members who cannot survive without cooperative group membership" (Campbell 1983, 37). On the macro scale, the proportion of strong reciprocators may vary from society to society, perhaps with major economic consequences, such as lack of economic growth (Knack and Keever 1997). For instance, in low-trust environments levels of investment are low, possibly explaining the poverty trap in which some groups and countries find themselves mired (Zak and Knack 2001).

Ideas closely related to commitment are found throughout the development of evolutionary and game theory perspectives on relationships. Even in *The Origin of Species* Darwin was aware of a "special difficulty" with respect to sterile castes of insects, "which at first appeared to me insuperable, and actually fatal to the whole theory" (Darwin 1859). Hamilton, after answering Darwin's quandary, explored the psychological mechanisms that mediate cooperation (Hamilton 1975). Even earlier, Alexander had worked out the dynamics of evolved sociality and some implications for morality (Alexander 1974). Trivers, in the original paper on the evolution of reciprocity, said, "People who are altruistically motivated will make more reliable partners than those motivated by self-interest" (Trivers 1971, 157). He also recognized that "The tendency to like others, not necessarily closely related, to form friendships and to act altruistically towards friends and towards those one likes will be selected for. . . . In other words, selection will favor liking those who are themselves altruistic" (Trivers 1971, 48). He saw the possibilities for gross and subtle cheating in friendships and said that this would shape moralistic aggression, although it has turned out to be difficult to see exactly how benefits come to those who bear the costs of moralistic aggression. He further went on to show how emotions such as gratitude and guilt

could be shaped by natural selection, and crucial roles of trust and suspicion in regulating relationships (Trivers 1981).

McClennen (1990) notes that in the publication that established game theory, Von Neumann and Morgenstern anticipate commitment strategies in their distinction between "dead" variables and those that "reflect another person's will or intention." He goes on to say,

> What Von Neumann and Morgenstern proceed to do, however, and what others have done since, is build a theory denying, in the end, that the behavior of each other player could be parameterized. . . . In effect what they do is reduce the problem for each participant to that of a standard maximization problem. . . . The assumption that rational players are bound by the principle of maximization of expected utility proves to be quite paradoxical . . . it ensures that rational players . . . will have to settle for outcomes that are mutually disadvantageous . . . why, if the possibility of gain drives them to enforcement devices, does it not also drive them *beyond*, to an even more efficient form of coordination? (McClennen 1990, 258–60)

The usual solution—to create precommitments and constraints—is suboptimal and fails to explain the existence of moral constraints that cannot be set aside when they become disadvantageous. Justifications for following such rules have been divided, he says, between those, such as Hobbes and Hume, who emphasize the tangible benefits of acting according to moral principles, as compared to others "like Kant [who] postulate that they must be grounded in some radically different fashion" (264). Neither, he says, is satisfactory because both are stuck in a limited view of rationality. Building on Gauthier (Gauthier 1986), his solution, *resolute choice*, "might, one would hope, provide a way to bridge the gap between moral choice and rational interested choice" (McClennen 1990, 260). As important, resolute choice also postulates a force of selection that could shape the moral passions that make commitment possible. "Whereas the sophisticated self acts under a constraint on choice imposed by the logic of belief, the resolute self acts under a constraint of a different sort altogether, a constraint of *commitment*" (160).

Gauthier provides a rationalist foundation for Rawls's contractarian ethics (Rawls 1971; Gauthier 1986). Gauthier's notion of rationality transcends the usual standard of modular rationality, which expects only "straightforward maximization" and instead says there are advantages to having a disposition to practice "constrained maximization," by which he means fulfilling promises even when that will not be advantageous (Gauthier 1986). These advantages are the key to his thesis that it can be rational to follow rules even when that means sometimes acting in ways that are not in a person's self-interest (Sug-

den 1993). Gibbard takes this further in his suggestion that natural selection has shaped emotional dispositions supporting a "normative control system" to facilitate constrained maximization and thereby give a selective advantage (Gibbard 1990). In Frank's terms, the emotions are "self-control devices" (Frank 1988, 81) that allow people to protect themselves and their reputations from their impulses to act in their short-term interests.

All of this harks back to Hume and Adam Smith with their emphasis on the importance of sympathy. Hume made it the foundation of his moral theory. We cannot, he said, derive oughts from our observations of the world or from any theory. They come instead from our intuitions, from our emotions. Mill, for all his utilitarianism, also emphasized the emotional origins of social cooperation, especially "the fear of exposure" (Mill 1848, 135). This theme continues to the present with work by psychologists showing that our moral choices emerge more from emotion than reason (Kagen 1984) and philosophical treatments of morality based on the natural origins of "apt feelings" for making normative judgments (Gibbard 1990).

Adam Smith, subtle psychologist that he was, emphasized the role of natural emotions in human affairs, especially empathy.

> However selfish soever man may be supposed, there are evidently some principles in his nature, which interest him in the fortunes of others, and render their happiness necessary to him, though he derives nothing from it except the pleasures of seeing it. Of this kind is pity or compassion, the emotion which we feel for the misery of others. . . . That we often derive sorrow from the sorrows of others, is a matter of fact too obvious to require any instances to prove it. . . . Nature, when she formed man for society, endowed him with an original desire to please, and an original aversion to offend his brethren. She taught him to feel pleasure in their favourable and pain in their unfavourable regard. (Smith 1984 [1759], 91, 116)

His outrage at a cynical view of commitments makes it clear that at least one aspect of the so-called sociobiology controversy has been going on for centuries:

> There is, however, another system which seems to take away altogether the distinction between vice and virtue, and of which the tendency is, upon that account, wholly pernicious: I mean the system of Dr. Mandeville . . . Dr. Mandeville considers whatever is done from a sense of propriety, from a regard to what is commendable and praiseworthy, as being done from a love of praise and commendation, or as he calls it from vanity . . . I shall endeavour to show that the desire of doing what is honourable and noble, of rendering ourselves the proper objects of

> esteem and approbation, cannot with any propriety be called vanity. (Smith 1984 [1759], 308–9)

Hobbes, in 1651, said it best of all in his Third Law of Nature: "*That men performe their Covenants made*. . . . In this law of Nature, consisteth the Fountain and Originall of JUSTICE . . . the definition of INJUSTICE is no other than the not Performance of Covenant." He notes that "there must be some coërcive Power to compel men equally to the performance of their Covenants, by their terror of some punishment greater than the benefit they expect by the breach of their Covenant" (Hobbes, 1996 [1651] no. 2693, 100). But then he goes on, in the parable of the Foole, to ask why people should keep their promises if the violation will not be punished.

> The Foole hath sayd in his heart, there is no such thing as Justice; and sometimes also with his tongue; seriously alleaging that every mans conservation and contenement, being committed to his own care, there could be no reason why every man might not do what he thought conduced thereunto: and therefore also make or not make, keep or not keep Covenants, was not against reason, when it conduced to ones benefit. . . . From such reasoning as this, Successful wickedenesse hath obtained the name of Vertue. . . . Where there is no Civill Power erected over the parties promising . . . there is the question whether it be against reason, that is against the benefit of the other, to performe, or not. And I say it is not against reason. . . . He therefore that breaketh his Covenant, and consequently declareth that he thinks he may with reason do so, cannot be received into any Society that unite themselves for Peace and Defense, but by the errour of them that receive him . . . if he be left, or cast out of Society, he perisheth." (Hobbes, 1996 [1651], 100–102)

This Foole will be recognized, of course, as *Homo economicus*. Centuries ago Hobbes understood why commitment persists. Those who announce that they do not need to keep their promises are cast out from society, where they perish. What could be a more fitting description of how natural selection shapes motives for keeping commitments? Moreover, what explanation could better explain why people are so hostile toward economists and evolutionists who seem to portray human behavior as self-interested and cynically deceptive?

Natural Selection

The core question remains: What does natural selection have to do with the capacity for commitment? In one sense, it may not matter that much. There are wide cultural variations in patterns of commit-

ment, and much of this variation seems to arise from different beliefs about human nature. It should be possible to understand how commitment functions in relationships and societies without knowing whether natural selection shaped tendencies specifically to make commitments possible. Yet, as noted by Byatt, we care deeply about our biological nature. Perhaps this should not influence us, but it does. Perhaps we should realize that what is natural is no guide to what is good, but when we discover (or start to believe in the absence of any evidence) that a tendency is a part of our nature, in fact we behave differently. As Hume pointed out, it is logically incorrect to try to get oughts from what is in nature; emotionally, though, the leap from is to ought seems easy and natural. This tendency fuels the fire about debates on human nature (Ruse 1982; Rose and Rose 2000; Segerstråle 2000).

This enormous range of cultural variation indicates that tendencies to use commitment strategies are by no means as hardwired as the preference for sweets. This variation does not, however, mean that natural selection is unimportant, any more than the wide variation in the foods people eat means that food preferences are not shaped by natural selection. Is it possible that the capacity for commitment in fact has not been shaped by natural selection? Yes, but this seems unlikely. If selection has not shaped special faculties for making and assessing commitments, some other explanation must be found for our moral passions, our preoccupations with reputation and honor, and our irrational wishes to fulfill commitments to help loved ones and harm enemies. Furthermore, despite wide variation in specific patterns, the emotional capacity for commitment seems to be a human universal, one that is complex, costly, and peculiar, and therefore difficult to explain unless it somehow increases reproductive success.

Psychological studies of mechanisms that could mediate commitment are just beginning; increasingly these studies have used an evolutionary context. Cosmides and Tooby have pursued an extensive program of research designed to show that our minds have a built-in system for conducting social relationships, especially a "cheater detection module" (Cosmides and Tooby 1992). They find support in their extensive data showing that people make fewer mistakes on the Wason selection task when the content is related to social obligations as compared to nonsocial information (Tooby and Cosmides 1989). This cheater detection module presumably has evolved not just to calculate the odds that another player will cooperate on the next move, but whether the partner is pretending to be a friend when he is actually just a "banker" who is willing to invest only when a payoff can be guaranteed (Tooby and Cosmides 1984, 1996). This model is closely related to the notion of commitment developed here.

Theories of social relationships on a large scale likewise nearly always identify some relationships as intimate or based on "communal sharing" in which the boundaries of the individuals merge (Fiske 1991). People take great pains to differentiate these relationships from those based on mere exchange. Psychological studies have found that friendship satisfaction is highly predicted by self-disclosure and trust, and is *decreased* by an exchange orientation (Jones 1991). Most people experience friendships as communal relationships (Mills and Clark 1982) as distinct from exchange relationships, with a "deep structure" that is distinct from the "superficial structure" of exchange (Hartup and Stevens 1999).

Perhaps the strongest evidence that friendships are based on commitment and not reciprocity is the revulsion people feel on discovering that an apparent friend is calculating the benefits of acting in one way or another. People intuitively recognize that such calculators are not friends at all, but exchangers of favors at best, and devious exploiters at worst. Abundant evidence confirms this observation. Mills has shown that when friends engage in prompt reciprocation, this does not strengthen but rather weakens the relationship (Mills and Clark 1982). Similarly, favors between friends do not create obligations for reciprocation because friends are expected to help each other for emotional, not instrumental reasons (Mills and Clark 1994). Other researchers have found that people comply more with a request from a friend than from a stranger, but doing a favor prior to the request increases cooperation more in a stranger than a friend (Boster et al. 1995). Owing to the important differences between friendship and acquaintanceship, people seem to take care with new potential friends to communicate whether they do or do not intend for the relationship to become a friendship. Communications that blur this distinction make people distinctly uncomfortable (Lydon, Jamieson, and Holmes 1997).

The importance of commitment is also a likely explanation for why people care so much about not only what others do, but why they do it. The meaning of an act differs dramatically depending on whether it is motivated by desire to influence or by emotional concern for a person's well-being. This is consistent with the emphasis philosophers have long placed on motives in determining the moral status of an act (Darwall, Gibbard, and Railton 1997). The significance of the difference between calculated versus real altruism has been long recognized (Krebs 1970). The capacity for making and keeping commitments allows people to engage in relationships that are impossible for those who just exchange favors. The more negative side of this same capacity is seen in studies of honor cultures (Nisbett and Cohen 1996). Individuals who give a calculated response to aggression can

be exploited easily. Those who are taken over by rage are far more formidable opponents. This is expanded at length in the chapter by Cohen and Vandello herein.

Finally, contemporary studies of moral psychology are compatible with evolutionary perspectives on commitment. As Krebs notes:

> From an evolutionary perspective Kohlberg's stages represent cognitive structures that were selected in ancestral environments because they upheld adaptive systems of social interaction. For example, Stage 1 structures upheld social systems based on obedience to authority; Stage 2 judgments upheld cooperative systems based on instrumental, individual exchanges (e.g., tit-for-tat reciprocity) and Stage 3 judgments upheld cooperative systems based on harmonious in-group relations. The cognitive structures that uphold Stage 4 judgments may have evolved relatively recently to uphold the maintenance of social structures, such as legal systems, within more complex societies. . . . Put another way, the original function of moral judgment was to constrain others from advancing their interests at the expense of those with whom they formed cooperative relations. (Krebs, Denton, and Wark 1997, 131)

Demonstrating how such psychological traits give a fitness advantage is only one half of an evolutionary explanation; the other is discovering their precursors and how they were shaped to their present forms. This brings us to the precursors of capacities for commitment. While commitment itself is a strategy, not a trait, it offers benefits that may well have shaped various psychological mechanisms. Certainly capacities for commitment may have precursors in tendencies shaped by kin selection and reciprocity. As many have pointed out, we evolved in small kin groups where helping others often helps the actor's genes (Alexander 1985; Irons 1998; Richerson and Boyd 1999). In such a circumstance it may be more efficient to just help rather than to calculate. Likewise, while the reciprocity literature emphasizes sequential favor trading, many social relationships seem to be based on metacommitments that people make to keep cooperating despite occasional difficulties. Our relationships, even those based on commitment, depend on some balance between what is offered and what is received. Yet success in social life depends on the ability to maintain relationships during dry periods without being exploited. Likewise, conflicts are inevitable, and it appears that humans, like primates, have a built-in motivation to reconcile differences (de Waal 1989). Such motivations may be crucial precursors to capacities for commitment.

A closely related second possibility is that the capacity for commitment may have emerged from the benefits of political alliances. Chimpanzees spend their lives making alliances. Subdominant males

routinely make coalitions to compete against the alpha male who, in turn, must maintain alliances with others to keep his position (de Waal 1982, Goodall 1986; Runciman, Maynard Smith, and Dunbar 1996). Human hierarchies are different from those of other primates, with strong tendencies toward egalitarianism (Boehm 1999) and are based as much on prestige as they are on dominance (Henrich and Gil-White 2000). Nonetheless, political alliances are central to human competition, and modeling them as reciprocal exchanges misses much of how they work (Masters 1983; Trivers 1981). Instead of monitoring each move to anticipate patterns of cooperation and defection, political partners support each other in general, over multiple situations during long periods of time. This is exactly the kind of situation in which selection might well shape tendencies for emotional loyalty and commitment. Such relationships are, of course, continually threatened by third parties. Social life is composed of triangles (Kerr and Bowen 1988). An individual cannot always fulfill the wishes of both other parties, especially when each seeks an ally in a conflict with a fourth party. Furthermore, humans routinely use the strategy of *splitting*—flattering one member of a partnership while denigrating the other in order to disrupt the alliance and get a new ally. Ending all relationships at the first such difficulty is not an option; that would end participation in social life. Finding ways to overlook these deviations in order to continue can be of great value. Tact—the ability to suppress just the right information while talking with someone—may be deception of a sort, but it is the glue that makes social life possible (Nesse 1990b).

One further consideration may turn out to be enormously important. When potential mates are assessing each other, one of their top priorities is to find someone who is kind and honest. (This is the central but neglected finding of the cross-cultural study of mate choice by Buss [1989].) Otherwise, when bad times come, all the good genes and resources in the world will not be of much use. Miller (2000) has suggested, in the course of his work on selection for traits that display quality, that sexual selection may have augmented the capacity for altruism in a runaway escalation. If men and women each have a genetic tendency to seek mates who are kind and honest, this will select strongly for acting kind and honest, and the best way to accomplish this is to actually be kind and honest. In short, in addition to all the other selective advantages of a capacity for commitment, it may have been subject to sexual selection.

Social selection may be crucial to the capacity for commitment. If groups routinely give rise to emergent forces of natural selection, this can shape otherwise unaccountable individual behavior tendencies. For instance, if individuals simply avoid those who have betrayed

them, the net effect on an individual prone to defection is, as Hobbes said, exclusion from the group and probable death. It is as if the group had decided to shape its future membership, but no foresight is needed. The effect emerges from the uncoordinated actions of many individual agents. Groups, of course, often do make decisions about who is welcome and who is not, and they sometimes act in unison to punish a deviant. Once commitments are established, such actions are especially easy to explain. Such sequences readily give rise to coevolution, in which social structures shape individual traits that in turn create slightly different social structures (Cavalli-Sforza and Feldman 1981; Lumsden and Wilson 1981; Boyd and Richerson 1985; Durham 1991; Dunbar, Knight, and Power 1999). Many psychological traits seem to exist to make culture possible, such as the tendency to conformity and identification with the group leader, and the susceptibility of children to indoctrination (Boyd and Richerson 1985; Barkow 1989a). Simon has suggested that the benefits of conforming to cultural expectations have shaped a tendency to "docility," a tendency to learn and follow social norms that can result in genuine altruism (Simon 1990). His argument depends on bounded rationality and the assumption that individuals cannot accurately distinguish what is good for them from what is good for their group, but it has the virtue of explaining prosocial traits as a result of social selection (West-Eberhard 1987; Simon 1990; Jason, Brodie, and Moore 1999).

The benefits of using commitment strategies effectively are themselves likely to be potent selection forces that emerge from the structure of a complex social environment. Furthermore, the functions of the newest part of our brain, the frontal lobes, closely match the abilities needed to use commitment strategies. The frontal lobes seem crucial in calculating the tradeoffs between the short-term costs of giving up options and long-term benefits that may or may not be obtained. Many of these calculations involve social capital, with resulting inordinate complexity (Alexander 1975; Humphrey 1976). The frontal lobes also are essential to inhibiting the pursuit of short-term goals in order to fulfill commitments in pursuit of long-term goals. They are deeply involved in the social calculations needed to decide whether another person will fulfill a commitment. This decision is made best by an empathic identification with the other in order to anticipate what he or she is likely to do. To use commitment successfully requires not only higher cognition, but also a theory of mind (Baron-Cohen 1995), intuitive empathy, and a capacity for "mind reading" (Krebs and Dawkins 1984).

Our ability to read fiction, indeed perhaps our tendency to crave and enjoy stories, may well arise because of the huge advantages of getting into the mind of another in order to anticipate whether he or

she is likely to follow through on commitments (Carroll 1995; Turner 1996). We watch Sophocles' play with understanding and horror as Antigone reveals her determination to bury the body of Polynices, despite the king's edict. As she keeps her commitment, and the king keeps his, tragedy unfolds. As we watch, we consider the importance and the risks of keeping our commitments. Likewise, Odysseus overcomes giant obstacles and forgoes the pleasures of various isles to make his way home to his wife. Penelope, the very model of faithfulness, has been using her wiles to stave off the rapacious suitors. When Odysseus finally arrives home, we understand his state of mind perfectly as he takes his bloody revenge. One wonders if the rise of commitment was a key to the growth and power of Greek civilization.

To be explicit about this line of speculation, it seems possible that capacities for commitment have given such substantial fitness advantages that they could have shaped high intelligence, language ability, empathy, a theory of mind, an ability to inhibit impulses, and our fascination with fiction. Furthermore, once commitments become established in a group, they create emergent forces of selection that further speed development of these capacities in a runaway race for social intelligence. Expulsion of those who do not keep commitments will quickly shape tendencies to honesty and loyalty, even if that results in tendencies to act in ways that are genuinely altruistic, or self-destructively spiteful.

Such a scenario could address many long-standing problems. Most important, it offers a possible explanation for how natural selection could shape a moral capacity for acting according to rules instead of calculated self-interest. The parallel benefits of threats and promises offer a deep link between the origins of good and evil. Commitment also offers a possible explanation for why we humans are such extraordinarily emotional and apparently irrational creatures. Many commitments work only if they are based on emotion or beliefs that are outside the range of reason. This offers a way to understand our human fascination with mythology, religion, and ideology, why we spend so much effort constructing social realities, and how those social realities cycle back into tangible costs and benefits that further shape human minds by means of natural selection.

While plausible and consistent with much evidence, these speculations are nowhere near being facts. Taken together they make an intriguing story, but we do not yet even know how important the capacity for commitment really is, nor do we know how people use commitment strategies differently within and among cultures. That capacities for commitment have been shaped by natural selection seems likely, but this needs much more exploration. This book re-

views prior work and summarizes the state of the field from diverse perspectives in order to call attention to the importance of commitment and pose clear questions that will advance our understanding.

This chapter was greatly improved by comments and suggestions from Stephanie Brown, Terry Burnham, Dov Cohen, Lee Dugatkin, Alexander Field, Alan Gibbard, Lillian Gleiberman, Martin Gold, Jack Hirshleifer, Matt Keller, Philip Kulkulski, Peter Railton, Thomas Schelling, Natalie Smith, two anonymous reviewers, and members of the University of Michigan Evolution and Human Adaptation Program Weekly Seminar.

References

Akerlof, George A. 1983. "Loyalty Filters." *American Economic Review* 73: 54–63.

Alcock, John. 1997. *Animal Behavior: An Evolutionary Approach*, 6th ed. Sunderland, Mass.: Sinauer.

———. 2001. *The Triumph of Sociobiology*. New York: Oxford University Press.

Alexander, Richard D. 1974. "The Evolution of Social Behavior." *Annual Review of Systematics* 5: 325–83.

———. 1975. "The Search for a General Theory of Behavior." *Behavioral Science* 20: 77–100.

———. 1982. "Biology and the Moral Paradoxes." *Journal of Social and Biological Structures* 5: 389–95.

———. 1985. "A Biological Interpretation of Moral Systems." *Zygon* 20(1): 3–20.

———. 1987. *The Biology of Moral Systems*. New York: Aldine de Gruyter.

Atran, Scott. 1990. *Cognitive Foundations of Natural History: Towards an Anthropology of Science*. New York: Cambridge University Press.

Axelrod, Robert. 1984. *The Evolution of Cooperation*. New York: Basic Books.

———. 1986. "An Evolutionary Approach to Norms." *American Political Science Review* 80: 1095–1111.

———. 1997. *The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration*. Princeton Studies in Complexity. Princeton, N.J.: Princeton University Press.

Axelrod, Robert, and William D. Hamilton. 1981. "The Evolution of Cooperation." *Science* 211: 1390–96.

Badcock, Christopher R. 1986. *The Problem of Altruism*. Oxford: Blackwell.

Barash, David 1977. *The Biology of Altruism, Sociobiology and Behavior*. New York: Elsevier.

———. 1979. *Sociobiology: The Whisperings Within*. New York: Harper and Row.

———. 1982. *Sociobiology and Behavior*, 2d ed. New York: Elsevier.

Barkow, Jerome H. 1989a. *Darwin, Sex, and Status: Biological Approaches to Mind and Culture*. Toronto: University of Toronto Press.

———. 1989b. "The Elastic Between Genes and Culture." *Ethology and Sociobiology* 10: 111–29.

Baron-Cohen, Simon. 1995. "Mindblindness: An Essay on Autism and the Theory of Mind." Cambridge, Mass.: MIT Press.

Beck, Aaron T. 1976. *Cognitive Therapy and the Emotional Disorders*. New York: International Universities Press.

Becker, Gary. 1991. *A Treatise on the Family*. Cambridge, Mass.: Harvard University Press.

Behe, Michael J. 1996. *Darwin's Black Box: The Biochemical Challenge to Evolution*. New York: Free Press.

Bingham, Paul M. 1999. "Human Uniqueness: A General Theory." Quarterly Review of Biology 74(2): 133–69.

Binmore, Ken. 1993. "Bargaining and Morality." In *Rationality, Justice and the Social Contract: Themes from Morals by Agreement*, edited by D. Gauthier and R. Sugden. Ann Arbor, Mich.: University of Michigan Press.

———. 1994. *Game Theory and the Social Contract*. Cambridge, Mass.: MIT Press.

Boehm, Christopher. 1999. "The Natural Selection of Altruistic Traits." *Human Nature* 10(3): 205–52.

Boster, Franklin J., Jose I. Rodriguez, Michael G. Cruz, and Linda Marshall. 1995. "The Relative Effectiveness of a Direct Request Message and a Pregiving Message on Friends and Strangers." *Communication Research* (Special Issue), *Communication and Social Influence* 22(4): 475–84.

Boyd, Robert, and Peter J. Richerson. 1985. *Culture and the Evolutionary Process*. Chicago: University of Chicago Press.

Bradie, Michael. 1994. *The Secret Chain: Evolution and Ethics*. Suny Series in Philosophy and Biology. Albany: State University of New York Press.

Brickman, Philip, and Dan Coates. 1987. "Commitment and Mental Health." In *Commitment, Conflict, and Caring*, edited by C. B. Wortman and R. Sorrentino. Englewood Cliffs, N.J.: Prentice-Hall.

Burnham, Terry. Forthcoming. "Testosterone and Negotiations: An Investigation into the Role of Biology in Economic Behavior." Unpublished paper. Harvard University, Cambridge, Mass.

Burnham, Terry, and Jay Phelan. 2000. *Mean Genes: From Sex to Money to Food, Taming Our Primal Instincts*. Cambridge, Mass.: Perseus Publishing.

Buss, David M. 1989. "Sex Differences in Human Mate Preferences: Evolutionary Hyotheses Tested in 37 Cultures." *Behavioral and Brain Sciences* 12: 1–49.

Byatt, Antonia Susan. 2000. "The Darwin Wars." *Prospect* (November).

Campbell, Donald T. 1975. "On the Conflict Between Biological and Social Evolution and Between Psychology and the Moral Tradition." *American Psychologist* 30: 1115–22.

———. 1983. "Two Distinct Routes Beyond Kin Selection to Ultrasociality: Implications for the Humanities and the Social Sciences." In *The Nature of Prosocial Development*, edited by D. L. Bridgeman. New York: Academic Press.

Caplan, Arthur L. 1978. *The Sociobiology Debate: Readings on the Ethical and Scientific Issues Concerning Sociobiology*. New York: Harper and Row.

Caporael, Linda R., Richard M. Dawes, John M. Orbell, and Alphons J. C. Van de Kragt. 1989. "Selfishness Examined: Cooperation in the Absence of Egoistic Incentives." *Behavior and Brain Sciences* 12(4): 683–739.

Carroll, Joseph. 1995. *Evolution and Literary Theory*. Columbia, Mo.: University of Missouri Press.

Cavalli-Sforza, Luca L., and Marcus W. Feldman. 1981. *Cultural Transmission and Evolution: A Quantitative Approach*. Monographs in Population Biology, 16. Princeton, N.J.: Princeton University Press.

Cosmides, Leda, and John Tooby. 1989. "Evolutionary Psychology and the Generation of Culture, Part II. Case Study: A Computational Theory of Social Exchange." *Ethology and Sociobiology* 10(1–3): 51–98.

———. 1992. "Cognitive Adaptations for Social Exchange." In *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*, edited by J. H. Barkow, L. Cosmides, and J. Tooby. New York: Oxford University Press.

Cronin, Helena. 1991. *The Ant and the Peacock: Altruism and Sexual Selection from Darwin to Today*. New York: Cambridge University Press.

Daly, Martin, and Margo Wilson. 1987. "Children as Homicide Victims." In *Child Abuse and Neglect: Biosocial Dimensions*, edited by R. J. Gelles and J. B. Lancaster. New York: Aldine de Gruyter.

Darwall, Stephen L., Allan Gibbard, and Peter Albert Railton. 1997. *Moral Discourse and Practice: Some Philosophical Approaches*. New York: Oxford University Press.

Darwin, Charles. 1859. *On the Origin of Species by Means of Natural Selection*. London: John Murray.

Dawkins, Richard. 1976. *The Selfish Gene*. Oxford: Oxford University Press.

———. 1982. *The Extended Phenotype: The Gene as the Unit of Selection*. San Francisco: W. H. Freeman.

———. 1989. *The Selfish Gene*, new ed. Oxford and New York: Oxford University Press.

de Waal, Frans B. M. 1982. *Chimpanzee Politics: Power and Sex Among Apes*. New York: Harper and Row.

———. 1984. "Deception in the Natural Communication of Chimpanzees." In *Deception: Perspectives on Human and Nonhuman Deceit*, edited by R. W. Mitchell and N. S. Thompson. New York: SUNY Press.

———. 1989. *Peacemaking Among Primates*. Cambridge, Mass.: Harvard University Press.

Dembski, William A. 1998. *Mere Creation: Science, Faith & Intelligent Design*. Downers Grove, Ill.: InterVarsity Press.

Dennett, Daniel. 1995. *Darwin's Dangerous Idea*. New York: Simon & Schuster.

Diamond, Jared M. 1997. *Guns, Germs, and Steel: The Fates of Human Societies*. New York: Norton.

Dugatkin, Lee Alan. 1997. *Cooperation Among Animals: An Evolutionary Perspective*. Oxford Series in Ecology and Evolution. New York: Oxford University Press.

Dunbar, Robin I. M., Chris Knight, and Camilla Power. 1999. *The Evolution of Culture: An Interdisciplinary View*. New Brunswick, N.J.: Rutgers University Press.

Durham, William H. 1982. "Interactions of Genetic and Cultural Evolution: Models and Examples." *Human Ecology* 10(3): 289–323.

———. 1991. *Coevolution: Genes, Culture, and Human Diversity*. Stanford, Calif.: Stanford University Press.

Elster, Jon. 1979. *Ulysses and the Sirens: Studies in Rationality and Irrationality*. London: Cambridge University Press.

———. 2000. *Ulysses Unbound*. London: Cambridge University Press.

Emmons, Robert A. 1999. *The Psychology of Ultimate Concerns: Motivation and Spirituality in Personality*. New York: Guilford Press.

Fehr, Ernst, and Armin Falk. 1999. *Homo Reciprocans and Homo Cooperation*. Zurich: University of Zurich Press.

Fehr, Ernst, and Simon Gächter. 2000. "Cooperation and Punishment in Public Goods Experiments." *American Economic Review* 90(4): 980–94.

Fiske, Alan Page. 1991. *Structures of Social Life*. New York: Free Press.

Flinn, Mark V., and Richard D. Alexander. 1982. "Culture Theory: The Developing Synthesis from Biology." *Human Ecology* 10(3): 383–400.

Frank, Robert H. 1988. *Passions Within Reason: The Strategic Role of the Emotions*. New York: Norton.

———. 1992. "The Role of Moral Sentiments in the Theory of Intertemporal Choice." In *Choice over Time*, edited by G. Loewenstein. New York: Russell Sage Foundation.

Frank, Robert H., Thomas Gilovich, and Dennis Regan. 1993. "Does Studying Economics Inhibit Cooperation?" *Journal of Economic Perspectives* 7(Spring): 159–71.

Fukuyama, Francis. 1995. *Trust : Social Virtues and the Creation of Prosperity*. New York: Free Press.

Gauthier, David. 1986. *Morals by Agreement*. New York: Oxford University Press.

Ghiselin, Michael T. 1969. *The Economy of Nature and the Evolution of Sex*. Berkeley: University of California Press.

Gibbard, Allan. 1990. *Wise Choices, Apt Feelings: A Theory of Normative Judgment*. Oxford: Oxford University Press.

Gintis, Herbert. 2000a. *Game Theory Evolving*. Princeton, N.J.: Princeton University Press.

———. 2000b. "Strong Reciprocity and Human Sociality." *Journal of Theoretical Biology* 206: 169–79 (Available at Gintis's personal web site).

Goodall, Jane. 1986. *The Chimpanzees of Gombe: Patterns and Behavior.* Cambridge, Mass.: Belknap Press.

Gould, Steven J., and Richard C. Lewontin. 1979. "The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme." *Proceedings of the Royal Society of London* 205: 581–98.

Hamilton, William D. 1964. "The Genetical Evolution of Social Behavior I and II." *Journal of Theoretical Biology* 7: 1–52.

———. 1975. "Innate Social Aptitudes of Man: An Approach from Evolutionary Genetics." In *Biosocial Anthropology*, edited by R. Fox. New York: Wiley.

Hartup, Willard W., and Nan Stevens. 1999. "Friendships and Adaptation Across the Life Span." *Current Directions in Psychological Science* 8(3): 76–79.

Henrich, Joseph. 2000. "Does Culture Matter in Economic Behavior? Ultimatum Game Bargaining Among the Machiguenga." *American Economic Review* 90(4): 973–79.

Henrich, Joseph, and Francisco Gil-White. 2000. "The Evolution of Prestige: Freely Conferred Status as a Mechanism for Enhancing the Benefits of Cultural Transmission." *Evolution and Human Behavior* 22: 1–32.

Hirshleifer, Jack. 1987. "On the Emotions as Guarantors of Threats and Promises." In *The Latest on the Best: Essays on Evolution and Optimality*, edited by J. Depré. Cambridge, Mass.: MIT Press.

———. 1999. "There Are Many Evolutionary Pathways to Cooperation." *Journal of Bioeconomics* 1: 73–93.

Hobbes, Thomas. 1996 [1651]. *Leviathan.* Cambridge: Cambridge University Press.

Humphrey, Nicholas K. 1976. "The Social Function of Intellect." In *Growing Points in Ethology*, edited by P. G. Bateson and R. A. Hinde. London: Cambridge University Press.

———. 1999. "Varieties of Altruism." *Social Research* 64: 199–209.

Hurd, James P. 1996. "Investigating the Biological Foundations of Human Morality." *Symposium Series*, vol. 37. Lewiston, N.Y.: E. Mellen Press.

Huxley, Thomas Henry. 1897. *Evolution and Ethics and Other Essays*. 6 vols. *The Works of Thomas Henry Huxley*, vol. 1. New York: D. Appleton.

Huxley, Thomas Henry, and Julian Huxley. 1989 [1893]. *Evolution and Ethics.* Princeton, N.J.: Princeton University Press.

Irons, William. 1996. "Morality, Religion and Human Evolution." In *Religion and Science: History, Methods, Dialogue*, edited by W. M. Richardson and W. J. Wildman. New York: Routledge.

———. 1998. "Adaptively Relevant Environments Versus the Environment of Evolutionary Adaptedness." *Evolutionary Anthropology* 6(6): 194–204.

Jason, B. Wolf, Edmund D. Brodie III, and Allen J. Moore. 1999. "Interacting Phenotypes and the Evolutionary Process II. Selection Resulting from Social Interactions." *American Naturalist* 153: 254–66.

Jones, Diane Carlson. 1991. "Friendship Satisfaction and Gender: An Examination of Differences in Contributors to Friendship Satisfaction." *Journal of Social and Personal Relationships* 8: 167–85.

Kagen, Jerome. 1984. *The Nature of the Child*. New York: Basic Books.

Katz, Leonard. 2000. *Evolutionary Origins of Morality: Cross Disciplinary Perspectives*. Devon: Imprint Academic.

Kerr, Michael E., and Murray Bowen. 1988. *Family Evaluation: An Approach Based on Bowen Theory*. New York: Norton.

Kitcher, Philip. 1985. *Vaulting Ambition: Sociobiology and the Quest for Human Nature*. Cambridge, Mass.: MIT Press.

———. 1993. "The Evolution of Human Altruism." *Journal of Philosophy* 90(10): 497–516.

Klinger, Eric. 1998. "The Search for Meaning in Evolutionary Perspective and Its Clinical Implications." In *Handbook of Personal Meaning: Theory, Research & Application*, edited by P. T. P. Wong and S. F. Prem. Mahwah, N.J.: Lawrence Erlbaum.

———. 2000. "Commitment." In *Encyclopedia of Psychology*. New York: American Psychological Association and Oxford University Press.

Knack, Stephen, and Philip Keever. 1997. "Does Social Capital Have an Economic Payoff? A Cross-Country Investigation." *Quarterly Journal of Economics* 112(4): 1252–88.

Kohn, Alfie. 1990. *The Brighter Side of Human Nature: Altruism and Empathy in Everyday Life*. New York: Basic Books.

Konner, Melvin. 1999. "Darwin's Truth, Jefferson's Vision: Sociobiology and the Politics of Human Nature." *American Prospect* (45): 30–38.

Krebs, Dennis. 1970. "Altruism: An Examination of the Concept and a Review of the Literature." *Psychological Bulletin* 73: 258–303.

Krebs, Dennis, Kathy Denton, and Gillian Wark. 1997. "The Forms and Functions of Real-Life Moral Decision-Making." *Journal of Moral Education* 26(2): 131–46.

Krebs, John, and Richard Dawkins. 1984. "Animal Signals: Mind-Reading and Manipulation." In *Behavioral Ecology: An Evolutionary Approach*, edited by J. R. Krebs and N. B. Davies. Sunderland, Mass.: Sinauer.

Lewontin, Richard C. 1982. *Human Diversity*. New York: Scientific American Library.

Lewontin, Richard C., Steven Rose, and Leon J. Kamin. 1984. *Not in Our Genes: Biology, Ideology, and Human Nature*. New York: Pantheon Books.

Lockard, Joan S., and Delroy L. Paulhus, eds. 1988. *Self-Deception: An Adaptive Mechanism?* Englewood Cliffs, N.J.: Prentice Hall.

Lumsden, Charles J., and Edward O. Wilson. 1981. *Genes, Mind, and Culture*. Cambridge, Mass.: Harvard University Press.

Lydon, John E., David W. Jamieson, and John G. Holmes. 1997. "The Meaning of Social Interactions in the Transition from Acquaintanceship to Friendship." *Journal of Personality & Social Psychology* 73(3): 536–48.

Masters, Roger D. 1983. "The Biological Nature of the State." *World Politics* 35(2): 161–93.

Maxwell, Mary. 1990. *Morality Among Nations: An Evolutionary View*. Suny Series in Biopolitics. Albany: State University of New York Press.

Maynard Smith, John. 1964. "Group Selection and Kin Selection." *Nature* 201: 1145–47.

———. 1982. *Evolution and the Theory of Games*. London: Cambridge University Press.

———. 1998. "Too Good to Be True." *Nature* 400(6741): 223.

Maynard Smith, John, and Eörs Szathmáry. 1995. *The Major Transitions in Evolution*. Oxford: W. H. Freeman Spektrum.

McClennen, Edward F. 1990. *Rationality and Dynamic Choice: Foundational Explorations*. New York: Cambridge University Press.

Mealey, Linda. 1995. "Sociopathy." *Behavioral and Brain Sciences* 18(3): 523–99.

Midgley, Mary. 1994. *The Ethical Primate: Humans, Freedom, and Morality*. London: Routledge.

Mill, John Stuart. 1848. *Principles of Political Economy*. London: John W. Parker.

Miller, Geoffrey F. 2000. *The Mating Mind: How Sexual Choice Shaped the Evolution of Human Nature*. New York: Doubleday.

Mills, Judson R., and Margaret S. Clark. 1982. "Communal and Exchange Relationships." In *Annual Review of Personality and Social Psychology*, edited by L. Wheeler. Beverly Hills, Calif.: Sage.

———. 1994. "Communal and Exchange Relationships: Controversies and Research." In *Theoretical Frameworks for Personal Relationships*, edited by Ralph Erber and Robin Gilmour. Hillsdale, N.J.: Lawrence Erlbaum.

Mitchell, Robert W., and Nicholas S. Thompson. 1986. *Deception: Perspectives on Human and Nonhuman Deceit*. New York: State University of New York Press.

Moore, David W. 1999. *Divided on Origins of Human Species*. Gallup Organization Reports.

Nesse, Randolph M. 1990a. "Evolutionary Explanations of Emotions." *Human Nature* 1(3): 261–89.

———. 1990b. "The Evolutionary Functions of Repression and the Ego Defenses." *Journal of the American Academy of Psychoanalysis* 18(2): 260–85.

———. 1999. "The Evolution of Hope and Despair." *Journal of Social Issues* 66(2): 429–69.

Nesse, Randolph M., and Allen T. Lloyd. 1992. "The Evolution of Psychodynamic Mechanisms." In *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*, edited by J. H. Barkow, L. Cosmides, and J. Tooby. New York: Oxford University Press.

Nisbett, Richard E., and Dov Cohen. 1996. *Culture of Honor: The Psychology of Violence in the South, New Directions in Social Psychology*. Boulder, Colo.: Westview Press.

Nitecki, Matthew H., and Doris V. Nitecki. 1993. *Evolutionary Ethics*. Suny Series in Philosophy and Biology. Albany: State University of New York Press.

Nowak, Martin A., and Karl Sigmund. 2000. "Evolution of Indirect Reciprocity by Image Scoring." *Nature* 393: 573–77.

———. 1998. "Shrewd Investments." *Science* 288(5467): 819.

Oyama, Susan. 1989. "Innate Selfishness, Innate Sociality." *Behavioral and Brain Sciences* 12(4): 717–18.

———. 2000. *Evolution's Eye*. Durham, N.C.: Duke University Press.

Parfit, Derek. 1984. *Reasons and Persons*. Oxford: Clarendon.

Petrinovich, Lewis F. 1995. *Human Evolution, Reproduction, and Morality*. New York: Plenum Press.

Radcliffe-Richards, Janet. 2000. *Human Nature After Darwin: A Philosophical Introduction*. New York: Routledge.

Rawls, John. 1971. *A Theory of Justice*, vol. HP 27, a Harvard Paperback. Cambridge, Mass.: Belknap Press of Harvard University Press.

Richards, Robert John. 1987. *Darwin and the Emergence of Evolutionary Theories of Mind and Behavior*. Chicago: University of Chicago Press.

———. 1993. "Birth, Death and Resurrection of Evolutionary Ethics." In *Evolutionary Ethics*, edited by M. H. Nitecki and D. V. Nitecki. Albany, N.Y.: State University of New York Press.

Richerson, Peter J., and Robert Boyd. 1999. "Complex Societies: The Evolutionary Origins of a Crude Superorganism." *Human Nature* 10(3): 253–89.

Ridley, Matt. 1997. *The Origins of Virtue: Human Instincts and the Evolution of Cooperation*. New York: Viking.

Rose, Hilary, and Steven Rose, eds. 2000. *Alas, Poor Darwin: Arguments Against Evolutionary Psychology*. New York: Harmony Books.

Rose, Michael R., and George V. Lauder, eds. 1996. *Adaptation*. San Diego: Academic Press.

Roth, Alvin E., Vesna Prasnikar, M. Okuno-Fujiwara, and S. Zamir. 1991. "Bargaining and Market Behavior in Jerusalem, Ljubljana, Pittsburgh and Tokyo: An Experimental Study." *American Economic Review* 81(5): 1068–95.

Rue, Loyal D. 1989. *Amythia: Crisis in the Natural History of Western Culture*. Tuscaloosa: University of Alabama Press.

Runciman, W. G., John Maynard Smith, and Robin I. M. Dunbar, eds. 1996.

Evolution of Social Behaviour Patterns in Primates and Man. New York: British Academy/Oxford University Press.
Ruse, Michael. 1982. *Darwinism Defended: A Guide to the Evolution Controversies*. Reading, Mass.: Addison-Wesley.
———. 1986. *Taking Darwin Seriously: A Naturalistic Approach to Philosophy*. New York: Blackwell.
———. 1989. *The Darwinian Paradigm*. London: Routledge.
Schelling, Thomas C. 1960. *The Strategy of Conflict*. Cambridge, Mass.: Harvard University Press.
———. 1978a. "Altruism, Meanness, and Other Potentially Strategic Behaviors." In *Choice over Time*, edited by G. Loewenstein and J. Elster. New York: Russell Sage Foundation.
———. 1978b. "Egonomics, or the Art of Self-Management." *American Economic Review: Papers and Proceedings* 68: 290–94.
———. In press. "Commitment." In *New Palgrave Dictionary of Economics and the Law*, edited by P. Newman. New York: Macmillan.
Schwartz, Barry. 1986. *The Battle for Human Nature: Science, Morality and Modern Life*. New York: Norton.
Segerstråle, Ullica C. O. 2000. *Defenders of the Truth*. New York: Oxford University Press.
Simon, Herbert A. 1990. "A Mechanism for Social Selection and Successful Altruism." *Science* 250: 1665–68.
Singer, Peter. 2000. *A Darwinian Left: Politics, Evolution, and Cooperation*. New Haven: Yale University Press.
Skyrms, Brian. 1996. *Evolution of the Social Contract*. New York: Cambridge University Press.
Smith, Adam. *The Theory of the Moral Sentiments*. 1984 [1759]. Edited by D. D. Raphael and A. L. Macfie. Indianapolis: Liberty Classics.
Smuts, Barbara B. 1999. "Multilevel Selection, Cooperation, and Altruism: Reflection on Unto Others." *Human Nature* 10(3): 311–27.
Sober, Elliott, and David Sloan Wilson. 1998. *Unto Others: The Evolution and Psychology of Unselfish Behavior*. Cambridge, Mass.: Harvard University Press.
Sperber, Dan. 1996. *Explaining Culture: A Naturalistic Approach*. Oxford: Blackwell.
Sugden, Robert. 1993. "The Contractarian Enterprise." In *Rationality, Justice and the Social Contract: Themes from Morals by Agreement*, edited by D. P. Gauthier and R. Sugden. Ann Arbor, Mich.: University of Michigan Press.
Tooby, John, and Leda Cosmides. 1984. *Friendship, Reciprocity and the Banker's Paradox*. Institute for Technical Studies Report 1.
———. 1989. "Evolutionary Psychology and the Generation of Culture, Part I: Theoretical Considerations." *Ethology and Sociobiology* 10(1–3): 29–50.
———. 1992. "The Psychological Foundations of Culture." In *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*, edited by J. H. Barkow, L. Cosmides, and J. Tooby. New York: Oxford University Press.
———. 1996. "Friendship and the Banker's Paradox: Other Pathways to the Evolution of Adaptations for Altruism." In *Evolution of Social Behaviour Patterns in Primates and Man*, edited by J. M. Smith, W. G. Runciman, and R. I. M. Dunbar. London: British Academy/Oxford University Press.

Trivers, Robert L. 1971. "The Evolution of Reciprocal Altruism." *Quarterly Review of Biology* 46: 35–57.
———. 1976. Foreword. In *The Selfish Gene*, by Richard Dawkins. New York: Oxford University Press.
———. 1981. "Sociobiology and Politics." In *Sociobiology and Human Politics*, edited by E. White. Lexington, Mass.: Lexington.
———. 1985. *Social Evolution*. Menlo Park, Calif.: Benjamin/Cummings.
Turner, Mark. 1996. *The Literary Mind.* New York: Oxford University Press.
Von Neumann, John, and Oskar Morgenstern. 1944. *Theory of Games and Economic Behavior*. Princeton, N.J.: Princeton University Press.
Wedekind, Claus, and Manfred Milinski. 2000. "Cooperation Through Image Scoring in Humans." *Science* 288: 850–52.
West-Eberhard, Mary Jane. 1987. "Sexual Selection, Social Competition, and Speciation." *Quarterly Review of Biology* 58: 155–83.
Williams, George C. 1966. *Adaptation and Natural Selection: A Critique of Some Current Evolutionary Thought*. Princeton, N. J.: Princeton University Press.
———. 1992. *Natural Selection: Domains, Levels, and Challenges*. New York: Oxford University Press.
———. 1993. "Mother Nature Is a Wicked Old Witch." In *Evolutionary Ethics*, edited by M. H. Nitecki and D. V. Nitecki. Albany: State University of New York Press.
Williams, George C., and Doris C. Williams. 1957. "Natural Selection and Individually Harmful Social Adaptations Among Sibs with Special Reference to Social Insects." *Evolution* 17: 249–53.
Wilson, David Sloan. 1975. "A General Theory of Group Selection." *Proceedings of the National Academy of Sciences* 72: 143–46.
Wilson, David Sloan, and Elliott Sober. 1994. "Reintroducing Group Selection to the Human Behavioral Sciences." *Behavioral and Brain Sciences* 17(4): 585–607.
Wilson, Edward O. 1975. *Sociobiology*. Cambridge, Mass.: Harvard University Press.
———. 1978. *On Human Nature*. Cambridge, Mass.: Harvard University Press.
———. 1998. *Consilience: The Unity of Knowledge*. New York: Knopf.
Wilson, James Q. 1993. *The Moral Sense*. New York: Free Press.
Wrangham, Richard W., and Dale Peterson. 1996. *Demonic Males: Apes and the Origins of Human Violence*. Boston: Houghton Mifflin.
Wright, Robert. 1994. *The Moral Animal: The New Science of Evolutionary Psychology*. New York: Pantheon Books.
———. 2000. *Non-Zero: The Logic of Human Destiny*. New York: Pantheon Books.
Wynne-Edwards, Vero Copner. 1962. *Animal Dispersion in Relation to Social Behavior*. Edinburgh: Oliver and Boyd.
Zak, Paul J., and Stephen Knack. 2001. "Trust and Growth." *Economic Journal* (April).