Chapter 14

The Future of Commitment

RANDOLPH M. NESSE


In 1883 Lord Kelvin mounted a major challenge to Darwin’s theory. He claimed that the time required for evolution was inconsistent with the laws of physics. “Essential principles of thermodynamics have been overlooked. . . . It is quite certain that the solar system cannot have gone on, even as present, for a few hundred thousand or a few million years, without an irrevocable loss . . . of a very considerable proportion of the entire energy initially in store” (Kelvin, Tait, and Darwin 1888, 468–69). He reasoned that the sun’s heat must come mainly from the friction of meteors, because “No other action, except by chemical action can be conceived . . . [and this] would only generate about 3,000 years of heat” (493). He concluded, “The inhabitants of the earth cannot continue to enjoy the light and heat essential to their life for many million years longer, unless sources now unknown to us are prepared in the great storehouse of creation” (494). Many found his argument convincing. He was, after all, Britain’s preeminent scientist, famous for engineering the first successful transatlantic cable. Furthermore, his conclusion was based on the established, rigorous science of physics.

In just a few years, however, it became clear that he was wrong. Fossil and astronomical evidence showed that the earth was far older than 100,000 years. By the end of the nineteenth century Curie and others had discovered radioactivity. By 1934 it was apparent that nuclear reactions fuel the sun and keep the earth’s core molten. We now know that the sun and the earth were formed about 5,000 million years ago and that the universe is about 13,000 million years old. Kelvin’s logic was fine, but because he assumed that all energy was mechanical or chemical, his result was off—by a factor of 120,000.
In recent decades two principles—kin selection and reciprocity—have come to dominate explanations for cooperation. The advance they represent cannot be overestimated. They are as central for understanding social behavior as the laws of thermodynamics are for understanding physics. Together with a gene-centered view of evolution, they provide a new, solid foundation for understanding relationships and social behavior. They have revolutionized the study of animal behavior. While political and emotional objections have slowed their application to human behavior, their utility is enormous. They are providing new insights into the workings of relationships, families, and social groups.

It is increasingly clear, however, that kin selection and reciprocity are not sufficient to explain all social phenomena. The oldest and most significant evidence is the very existence of mental mechanisms for guilt and sympathy (O’Connor 2000). While they can be interpreted as manipulations or devices for managing exchange relationships, they often give rise to actions far different from those predicted by rational choice theory (RCT). More recently, hundreds of studies in behavioral economics find that nonrational decisions are ubiquitous (Fehr and Falk 1999). While some of these deviations from RCT merely reflect the limitations of generally effective mental mechanisms (Gigerenzer, Todd, and ABC Research Group 1999), others are better explained by evolved social predispositions that likely give superior outcomes. Further evidence comes from psychological studies showing that close relationships are harmed if attention is called to the favors exchanged. This book offers many more examples, especially in parts III and IV.

The basic principle of sociobiology remains clear and incontrovertible: behavior regulation mechanisms for all species are shaped, inevitably and necessarily, to induce actions that tend (in the long run, on the average, in the natural environment) to increase the frequency of that individual’s genes in future generations. It is increasingly obvious, however, that maximizing reproductive success often requires keeping promises and fulfilling threats even when that requires profound sacrifices of personal short-term interests. That natural selection has shaped special mental capacities to make this possible, including a capacity for commitment, seems likely.

What We Have Learned

The chapters in this book offer an overview of commitment from the perspectives of several different disciplines. A single-authored volume could be more consistent, but would fail to capture the diverse points of view that are brought together here. Early on in this project,
I found this diversity dizzying and spent weeks trying to reconcile different positions. While that effort was worthwhile, premature attempts to resolve all the apparent contradictions are unjustified. This is a new area of work and it needs to avoid premature closure that glosses over important differences. While much effort remains to further unpack the idea of commitment and to understand how it works, we are far ahead of where we started. Like natural selection, commitment is a simple idea whose elaborations are subtle and complex. The level of complexity is not overwhelming, though; it is just about right for a core concept with wide explanatory power.

The chapters in part I describe and define commitment in wonderful depth and detail. I will not summarize them here, but instead will draw a few crucial distinctions from those chapters that begin the process of unpacking the concept of commitment. Nine questions can help to classify a commitment:

- Whom is the commitment intended to influence: the self, other individuals, a group, or a spiritual being?
- Is the commitment a preemptive action that leaves the next move to the other, or is it a statement of intent regarding some future action?
- Does the commitment represent a threat to harm, or a promise to help?
- Does the commitment bring benefits now with costs later, or costs now with benefits later?
- Are the effects of the commitment probabilistic, or is the outcome highly predictable?
- Does the commitment make options impossible, or only more costly?
- Is fulfilling the commitment contingent on some event or some action by others, or is it noncontingent?
- What controls incentives after the commitment is made: the situation itself, a third party, a group, or does the individual who made the commitment continue to control incentives?
- Why would others believe the commitment: incentives in the situation itself, incentives controlled by a third party, the benefits of reputation, or internal motives?

A full exposition on each of these questions would (and no doubt eventually will) fill other books. Here I will only briefly comment on some of them. The first question identifies whom the commitment is intended to influence. Most commitments are intended to influence others, but people make some commitments to influence themselves.
They anticipate that they will not be able to restrain their impulses in a future situation, so they take action now to make certain options are either impossible or too costly (Elster 1979; Burnham and Phelan 2000). If you ask your friends to bind you to the mast, or you sign yourself into a locked drug rehabilitation ward, some options are eliminated outright (question 6). If you tell your friends about your commitment to lose twenty pounds, your option to give up your diet is not impossible, but it does become embarrassing. People often voluntarily take actions to influence their own behavior by eliminating options that are tempting in the short term but undesirable in the long term.

The second distinction is between preemptive actions that leave the next move to the other, versus commitments that refer to future behavior, a difference emphasized by Hirshleifer (ch. 4, this volume). This is important because the former is relatively simple and is tractable in game theory. The latter kind of commitment—to future actions—is the main focus of this book.

Another major difference is between threats and promises. The idea of commitment highlights their deep similarities, and the correspondingly profound possibility that our capacities for good and evil have parallel origins. Yet many of the phenomena involved are distinct. Promises usually obtain benefits now based on a commitment to provide costly help later, and they tend to be contingent on the other person fulfilling a corresponding promise. In contrast, threats are not mutually agreed on, and they tend to be contingent on another person's lack of compliance. To consider promises and threats together in this book has been essential, but their many important differences suggest that separate treatments often will be preferable.

The fourth distinction is whether the commitment requires costs now for benefits later or vice versa. When a small person stands up to a bully, the costs come quickly while the benefits, if any, are only in the long run. Conversely, when you promise to stay with someone in sickness and in health, some benefits to the relationship come soon but the costs are later and uncertain, as are the benefits if you are the one who gets sick. When benefits come first, the problem is why someone will follow through and pay the costs later. When the costs come first, the problem is how to motivate accepting costs now, especially if the benefits are by no means certain but only possible.

The final two distinctions are the basis for the categories provided in the introductory chapter. They call attention to the core problem of commitment: Why would others believe that someone would do something that is not in his or her self-interest? The global distinction here is between commitments that are secured and those that are not. A secured commitment is like a loan secured by collateral; tangible costs and benefits make it worthwhile for the person to fulfill his or
her commitment. When a commitment is unsecured, however, no tangible incentives enforce fulfillment. Such subjective commitments are at center stage here.

The chapters in part II consider the role of subjective commitment in animal behavior. They find related phenomena, especially of preemptive actions, but no compelling examples of subjective commitment. As all three authors note, however, we have not yet looked hard. Having commitment in mind may allow us to see phenomena we have missed, especially those mediated by reputation. It may well turn out that commitment strategies are possible only for species that can accurately signal their future intent. The vigorous claw snapping of vulnerable molting shrimp, described by Adams, seems to represent a mere bluff. The chimpanzee "quiet signals" described by Silk may well qualify as commitments. The guppies described by Dugatkin certainly recognize and influence each other, but this seems to reflect reciprocity at most. Overall, the exercise of looking for commitment strategies in animals motivates useful attempts to operationalize the concept in behavioral terms, but the task remains difficult. The fact that three distinguished ethologists do not find definitive evidence for clear-cut subjective commitments in animals is valuable information. Each of the authors describes strategies for going about looking more carefully for commitment and related phenomena. Perhaps their suggestions eventually will allow us to see phenomena we have not looked for previously.

Part III, in contrast, provides abundant examples of commitment and subjective commitment in humans. Much of what goes on in honor societies seems to illustrate commitment strategies. Similarly, many phenomena involved in group enforcement of norms seem to fit into a commitment framework, and the selection forces that emerge from such groups offer a powerful route by which natural selection could shape capacities for commitment. The existence of the moral passions offer the strongest evidence we have for such a possibility. These chapters provide a rich starting point for anyone who wants to find detailed examples of the role of commitment in human life.

Part IV further documents the importance of commitment in social groups, including law, psychotherapy, and religion. Aspects of law that are otherwise difficult to explain look to be based on commitments. Psychotherapy can be understood in part as an investigation of how people's beliefs about commitments influence their relationships. Those who have experienced inconsistent or exploitative relationships have an opportunity to find new ways to relate to others. The consternation patients experience in a relationship that is committed but based on monetary payment is explained by a commit-
ment framework. This framework goes further yet, and suggests that some aspects of human groups seem to be shaped by the effects of commitments, including defined group boundaries, initiation rituals, ideologies, and costly signals of loyalty to the group. Conversely, many aspects of groups seem to exist largely to facilitate commitments to cooperate and to limit the utility of coercive threats.

Some economists challenge the very existence of subjective commitments on the straightforward grounds that they are effective only if people believe that others will do things that are not in their interests (Field 2001). If people are assumed to act in their own self-interest, then beliefs in such commitments usually will be false and therefore tendencies to believe such commitments should be eliminated by natural selection. Emotions should do nothing to change this assessment. In short, subjective commitments require actions that decrease fitness, so people should not believe them. Commitments not believed cannot influence people and therefore should not exist. This critique calls attention to the need for explicit accounts of why subjective commitments are not made useless by deception. The general theoretical answer, as outlined in the first three chapters and throughout the book, is that individual decisions that seem inconsistent with RCT can nonetheless give a long-term advantage via reputation, and via the sum payoffs from many commitments with probabilistic payoffs. A more specific answer comes from observation: subjective commitments are ubiquitous, they are believed, and they do influence people. If you know that a person has taken spiteful revenge in the past, you anticipate this and adjust your behavior irrespective of what RCT says about the payoffs in the current situation. Whether by reputation or other means, anything that convinces people that a commitment likely will be kept will influence behavior. Emotional displays and unreasonable behavior may not be reliable signals, and they certainly can be deceptive, but they do establish that the actor's behavior cannot be readily predicted on the basis of simple self-interest.

Many subjective commitments are enforced by pledges of reputation that make the cost of reneging very costly indeed—assuming, of course, that the person has a reputation to begin with. We distrust strangers partly because we know nothing about their reputation. Yet if we know a person's character has integrity, then we are influenced by his or her threats or promises. This makes reputation a valuable resource; people who lack it are socially feeble. Fear of losing reputation thus can motivate fulfilling even costly commitments. This is the simplest explanation as to why people believe subjective commitments—it also is a potent explanation for how natural selection could have shaped tendencies to fulfill commitments, even when doing so requires actions that considered alone directly decrease fitness. This
can explain how natural selection could have shaped moral faculties for fulfilling commitments and following rules even when that is very costly. In the social life of humans, the power of reputation is profound.

Some subjective commitments are maintained even when that would not be justified by costs to reputation. I remain unsure whether these reflect a whole additional route to commitment, or whether they are better studied as phenomena made possible by commitment mechanisms shaped by natural selection. Certainly many of us rely on such commitments for most of our important relationships. When you promise to stay married to someone, the commitment is largely emotional. If you defend a friend who has been unfairly treated by your company, your loyalty is likely to be costly. We care deeply about subjective commitments; they are a central theme in religion and great literature. Finally, what we believe about commitments shapes the social realities of our lives (Miller 1999). In short, commitments are extremely important to human social life.

What Commitment Needs Now

Work of several kinds is needed to develop and study commitment: scholarship, modeling, measurement, and experimental and field studies. This book starts with the seminal work of Schelling, Frank, and Hirshleifer, and ranges extraordinarily widely, but it has hardly begun to incorporate relevant work from many important areas. Social psychology, for instance, has developed a body of relevant research on trust. Personality researchers have studied individual differences related to commitments. Also in psychology, the concept of emotional commitments to goals and close relationships have been explored in depth (Klinger 2000). These concepts are somewhat different than the notion of commitment described here, but the interface has much to offer. In economics and political science a growing literature investigates trust and social capital, as well as the influence of loyalty (see other volumes in this series). Behavioral economics is such a huge field now that considerable work will be needed to determine which results can be interpreted sensibly in terms of commitment (Fehr and Schmidt 2000). Anthropology and sociology also must involve similar lines of thinking and research. Certainly commitment is a major theme in centuries of work by moral philosophers. To try to bring these different lines of thought together to see what they have already said about phenomena related to commitment is terribly important. I hope that the scholars in these other fields will understand that a book such as this cannot hope to cover all the relevant contributions; it is just a beginning.
Studies of commitment would benefit greatly from a rigorous model. Verbal descriptions simply won’t do. We need defined variables and postulated relationships among them that will allow us to determine how commitment strategies can work and under what circumstances. Why this has not been accomplished more than it has already (see Hirshleifer, ch. 4 herein, and Frank 1988, appendix) may be all too apparent. One reason is the need for many variables. Another is the need to accommodate not only general reputation, but also individualized relationships (Axelrod 1997). A model based on observed patterns of human relationships might have the following characteristics. Imagine a multiplayer game in which each agent can, on each move, invest any proportion of his or her reserves in a joint venture with any other individual or group. Each player privately seals cash in an envelope and gives it to a central authority, who sums the amounts contributed to each venture and then adds interest of perhaps 10 percent before distributing the sweetened pot equally among the investors. To include randomness, the payoff would vary, perhaps with a standard deviation of 15 percent of the total, with occasional larger deviations. Posting of outcomes offers a translucent indication of others’ cooperativeness, but owing to the random variation, players can never be sure if the results from a given investment resulted from partners being generous or stingy or from the random variations. On some rare moves the blind would be unexpectedly broken and all investments would be revealed. Some cost would be needed for getting information about others’ actions and reputations. To do well in such a game an agent must invest in mutual ventures. Yet this exposes the investor to the risk that others will cheat slyly, by putting in less than an equal share, or that they will defect by investing nothing. Another problem is how to get others to be willing to invest with you. This game combines characteristics of a public goods game with aspects of individual and generalized reciprocity models. I would be pleased and not at all surprised to learn that someone has tried already to operationalize something along these lines. Many game theorists have addressed issues and models close to these (Axelrod 1986; Skyrms 1996; Gintis 2000; Nowak et al. 2000).

What strategies would work best in such a game? Would steady intermittent slight cheating yield a superior return, or would others notice the inferior rate of return, and act on that information and on information from the occasions that reveal the actual investments? It might well work better to invest more than one promises in order to avoid suspicion and preserve relationships with good partners. Just as in human society and many economic models, no one strategy is likely to dominate. I think it is likely, however, that what many people do in actual life would probably work in the game: increase in-
vestments slowly in ventures with a few people and with groups who appear generous and trustworthy and whose actions and reputations can be monitored closely. If many individuals play this strategy, those who make investments on the basis of RCT would be isolated. Those who are caught cheating are likely to be excluded from the game, or at least forced to play with others who have equally bad reputations.

Economics and sociobiology often have been criticized for promoting simplistic models of human behavior. This criticism is justified. Like other organisms, we humans are designed to maximize reproductive success, but how we do so may be too complicated to incorporate into a tractable model. Adding other crucial factors that strongly influence people’s cooperation behaviors, such as kinship, ethnicity, social groups, ideologies, and other factors, would complicate matters further. Real social life is so complex that it may have been a potent force shaping human intellect (Alexander 1974; Humphrey 1976) and, perhaps, capacities for making and assessing commitments.

This brings us to the general issue of measurement. If you were going to describe and measure commitments, how would you do it? The nine distinctions at the beginning of this chapter are only a start. Specific examples are everywhere waiting for careful observation and description: spouses caring for their sick partners, jealous lovers threatening murder, disgruntled employees seeking revenge, and employees who stay in a low-paying job out of loyalty to a committed employer. Identifying and measuring commitment is awkward in that it is only one factor in a relationship that also may well include reciprocity, kinship, mutualisms, and other influences. Nonetheless, analysis of the role of commitment in actual relationships is certainly a good starting point.

Such detailed observations and descriptions of commitment in natural settings will provide a foundation for measurements that make possible studies of variations between subjects, within subjects, and across cultures. The study of personality variations in tendencies to commitment is primary, even though it is no substitute for the study of how situations influence commitments. Even a moment’s attention reveals the vast diversity of human capacities for commitment. A certain percentage of people are sociopaths who have no feelings to motivate loyalty or principled behavior. Having no personal experience of such feelings, they are likely to interpret other people’s commitments as manipulations: sociopaths cannot benefit from commitments except by using con strategies to gain the confidence of those they fleece. On the other extreme are those socially sensitive souls who spend every conscious minute (and many others also) worrying about
whether they have done their duty, whether they have offended anyone, and what people will think of them. The simple fact of this wide variation is remarkable. Individuals in other species also have personalities that persist across the lifespan, but to observe such a huge range of variation in such a core trait is remarkable. Also important is the genetic contribution to these traits. Behavioral genetic studies of twins show that 40 percent or more of the tendency to social sensitivity and sociopathy arises from genetic differences (Bouchard 1994). It is important to try to discover whether this variation is the result of random genetic drift, whether it reflects frequency-dependent selection for different strategies that offer the greatest benefits when rare (Mealey 1995), or whether we are in the midst of a transition in which genetic tendencies to capacities for commitment-based cooperation are spreading but not yet completely dominant.

Studies are needed to find out how a person’s commitments vary depending on the situation and the actions of others. Our major life commitments are too stable to be readily studied, yet most of us have a relationship with say, our grocer, based on some combination of mutualism and reciprocity, although even here are hints of commitment. Your grocer may try to instill feelings of loyalty by occasional free offers, and by generously replacing anything you feel is defective, whether it is or not. Retail corporations seem generally to have settled on quite lenient return policies that give the consumer the benefit of the doubt. They have learned that trying to argue over a dollar here and there will lose them customers in the long run. In fact one angry customer may cost them ten other customers. Generous return policies may be an example of principled moral behavior that requires accepting many short-term losses. Some unscrupulous consumers exploit these policies, but corporations must accept this cost in order to maintain a good reputation. In the world of work, employers always are trying to get their employees to believe that the company is committed to them. Sometimes this is true, as when the employees of Malden Mills were kept on after the factory burned. When employees are deceived, however, the repercussions are dramatic. Some say this accounts for a period of low employee morale at Northwest Airlines, where employees apparently perceived that their sacrifices in bad times would make them partners in good times. No such luck. When the economy revived, only the bosses and shareholders reaped big rewards. The sense of betrayal of an apparent commitment seems to have aroused persisting anger that is turned readily into job actions. Even beyond such examples, an evolutionary perspective has much to offer business management (Nicholson 2000).

Variations in commitment also could be addressed using experimental methods. Some of this already has been done, for instance in
the experiments by Clark and Mills showing that people become more distant, not closer, when a friend suggests an explicitly reciprocal exchange (Mills and Clark 1994; Clark and Mills 1993). Similar methods could manipulate the circumstances in groups, so as to foster or inhibit the use of commitment strategies. Probably much along these lines has been done, and is ready to be brought into the commitment framework.

Finally there is the urgent need to look at commitment across cultures, and in different subcultures. My anthropologist colleagues tell me that the very idea of guilt is hard for people in some cultures to comprehend. Today’s newspaper has a story about how politicians in certain countries tend to expect other nations to act ruthlessly in their self-interest, and how befuddled they are when they encounter anything else. If these cross-cultural differences prove as substantial as they seem, this makes it seem that the capacity for commitment is more a creation of particular cultures than a universal human capacity shaped by natural selection. Of course, the mental and emotional predispositions that make commitment possible may be shaped by natural selection for other purposes, perhaps limited to a few close relationships. Some cultures may use these capacities to create large-scale commitments that give big advantages to their members. Also, the capacity for commitment possibly has been shaped for large-scale social situations and lies in wait, ready to be expressed when doing so is likely to be worthwhile.

**Did Natural Selection Shape a Capacity for Commitment?**

Our core question is whether natural selection shaped mental mechanisms that facilitate making and assessing commitments. While the evidence is not yet ironclad, there are good reasons to think the answer is yes. Commitments are central and ubiquitous in social life. We are preoccupied with our reputation and the knowledge about the reputations of others. Our reputations determine whether others believe our commitments, and we need to know if we should trust commitments made by others. We have emotions and moral capacities that are otherwise hard to explain. Our close relationships explicitly avoid the appearance of reciprocal exchange. Social institutions, especially religions, seem to be designed to foster commitments; and our distinctive human capacities for empathy, foresight, intelligence, and love and hate appear to be just what is needed to make and assess commitments. These general factors, combined with the detailed explications throughout the book, offer convincing evidence that we have mental capacities for commitment. This does not necessarily
mean that natural selection shaped our capacities for commitment specifically due to the direct benefits of making commitments. Indeed quite possibly these capacities were shaped somewhat indirectly by social selection. While the capacity for commitments gives benefits, strong selection is likely against people who lack such a capacity, because they will have fewer allies and will be vulnerable to exclusion from the group.

Threats may prove easier to study both in the field and in the lab. Also, the reasons why people respond to threats may be different from why they respond to promises. Imagine walking down a dark street at night, and suddenly someone pokes something in your back and says, "Give me your wallet or I will shoot you." Your wallet, like that of most people, has only $50 in it, and the prison term for armed robbery is twenty years, so you could quickly calculate that the thief must be bluffing. Even if the chance of being caught is small, the benefits to the thief just aren’t worth it. It especially is not worthwhile to kill you, since that would set off a manhunt that most likely would lead to capture and years spent in prison. So if you were like some economists, you might conclude that the threat must be a bluff since the cost-benefit ratio is so wildly irrational. Yet most of us, less rational but perhaps longer-lived, would freeze and give up our money. Perhaps we are calculating that the value of the wallet is nothing compared to even a small chance of being killed. This is a reflection of the smoke detector principle—that is, the tendency of normal systems to express inexpensive protective defenses, such as cough or anxiety, whenever there is even a small chance that danger is actually present (Nesse, forthcoming). These systems normally give many false alarms, but a net advantage in the long run. Such exigencies well might make threats work successfully, even thoroughly irrational threats such as murder or suicide.

A parallel question that requires further study is why people believe that others will keep commitments. Parts III and IV offer substantial evidence that people do make and keep and believe commitments to cooperate; but what are the advantages? Several have not been fully explored. First, while reciprocity provides benefits only when you have other rewards to offer in turn, commitment provides help when you need it most—when you are sick or excluded from the group or when you are simply down and out. A person who can call on help in such critical times has a huge advantage over those who can get help only when they have something to offer in return. Such help, of course, usually depends on the other person’s expectation that you would do the same for him or her. This requires careful attention to that person’s needs, ensuring that the relationship does not appear to be based merely on reciprocal exchange. Those who can
believe in others' commitments are more likely to be able to make credible commitments of their own.

A closely related benefit is having allies who can give political support. Reciprocity envisons two actors with individual and differentiated interests. When two people become partners, however, their interests merge. They try to help and protect each other in every circumstance. Studies of primates offer many hints. Chimpanzees spend much of their time making and preserving alliances that they use to challenge the status hierarchy above and to defend attacks from below (de Waal 1982). An individual without allies is hopelessly weak. Such alliances can be interpreted as trading favors, but observation shows that the unity of interests is much more general; and so it is with humans and their friends. We expect friends to take our side in conflicts. Friends do not carefully and objectively consider all the aspects of the situation before deciding if they will offer support, they just automatically offer it. Indeed, if reproductive success depends on social success, and social success depends on political alliances, then these alliances well might shape social tendencies different from those expected to arise from reciprocity, an idea worth looking into. Notably such alliances can be characterized by mutualisms: neither party has a chance to succeed without the continuing efforts of the other. Moreover, in the case of chimpanzees or, for that matter, that of humans as depicted in Shakespeare's Henry VI trilogy, as soon as the alliance reaches its goal it dissolves and the erstwhile partners now fight as their interests diverge. This is discouraging to a point of view that would like to see commitments as enduring beyond those situations in which they offer benefits.

Another more mundane advantage of basing cooperation on commitment is reduced monitoring and bookkeeping costs. Even when trading baby-sitting, tallying and recording every minute is a drain and a source of potential conflict. A more relaxed attitude invites exploitation, but reaps major efficiency benefits. If each partner strives to offer a little more than expected instead of a little less, this offers manifold advantages. Such agreements of course could be interpreted as a metareciprocity game in which the players commit to repeated cycles of cooperation that eventually become habitual. Yet the emotional predispositions for friendship are powerful, and quite different from those that would be shaped by mere reciprocity. Nonetheless, when a relationship becomes too unbalanced, related emotions usually motivate actions to reestablish fairness or else end the relationship. If the relationship has been based on commitments, the sense of betrayal is likely to be intense.

Social groups give rise to perhaps the most important and intriguing advantage of a tendency to commitment-based cooperation. The benefits of cooperation between individuals have been emphasized
here, as elsewhere, because analyses of individual relationships are more tractable. Yet as Boyd and Richerson and others going back at least to Hobbes have suggested, groups manage to enforce tight norms of behavior by punishing deviants (Boyd and Richerson 1985; Gintis 2000). This raises the question of why being one of those who imposes punishments is advantageous. In many groups, however, you don’t have to be a whistleblower to punish deviants. All you need to do is support those who support you and the group, and to turn away when others want your help. By making alliances with those who help you and ignoring others, individuals acting strictly in their own interests create powerful forces of selection. This is a bit different from cultural group selection increasing the frequency of traits that benefit groups—it is better described as social selection (West-Eberhard 1987; Jason, Brodie, and Moore 1999). When individuals pursue their own interests in social groups, this may create emergent forces of selection that pose serious fitness costs to those who show no capacity for commitment. Social selection works in the short term to shape individual behavior to conform to group norms as well as tendencies to learn the benefits of following social customs and to try to please others in the group. This is what Simon meant by docility (Simon 1990). In the long term such forces of social selection also can be forces for natural selection. Individuals who exclusively pursued their own interests, even by subtle reciprocity strategies, may have done poorly indeed, both because they had few allies in the group and because they were at constant risk of being “cast out of Society” where they likely “perishith” (Hobbes 1996 [1651]).

Even though it is a likely engine of social complexity, commitment can be very simple. All it takes is one person with a modicum of three core human capacities: foresight, language, and moral passions. If that person uses foresight to anticipate how others will respond to changed expectations, language to communicate a commitment, and moral passions to inhibit tendencies that might prevent fulfillment, then others will be influenced and social life is changed. The benefits of thus influencing people could well motivate keeping commitments that seem senseless when considered in isolation. Once many people start using such strategies, a tangled bank of social complexity is born, with personalities of many kinds communicating with escalating subtlety in the elaborately constructed groups and rituals that are the framework of human society.

References


