University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

2005

Cooperative Brains: Psychological Constraints on the Evolution of Altruism

Jeffrey R. Stevens
University of Nebraska-Lincoln, jstevens5@unl.edu

Marc D. Hauser
Harvard University

9   Cooperative Brains: Psychological Constraints on the Evolution of Altruism

Jeffrey R. Stevens and Marc D. Hauser

Imagine an individual called "hunter" that expends a good deal of energy to capture a gazelle. As the hunter is consuming his small prey, a second individual called "recipient" approaches and begins feeding peacefully alongside the hunter. A few weeks later the roles reverse, such that the previous recipient has now captured a gazelle, and the previous hunter is taking advantage of the recipient's hard work. Could the hunter and recipient be Maasai warriors? Is it equally likely that they are common chimpanzees, African lions, or Nile crocodiles? All of these species hunt gazelle and live in groups, so why would this scenario apply to some species more appropriately than to others? The answer lies in the costs and benefits associated with sharing food with non-kin. Assuming that one individual can consume the entire gazelle, sharing food with the recipient constitutes an altruistic act—the hunter accepts a fitness cost (reduction in food intake) while increasing the fitness of another (increasing the intake of the recipient).

Reciprocal Altruism: Cooperation via Turn Taking

Here we focus on a form of altruistic cooperation with interesting psychological implications: reciprocal altruism (or reciprocity)—the alternation of donor and recipient roles in repeated altruistic interactions. When Trivers (1971) first introduced the concept of reciprocal altruism, he outlined necessary prerequisites such as a large benefit to the recipient and a small cost to the donor, many opportunities for cooperative interactions, and the ability to detect cheaters. Humans appear to satisfy these requirements quite nicely.[1] Intuitively, reciprocity seems fundamental to human social interactions, and experiments confirm this intuition, demonstrating its prevalence across different economic contexts (Fehr & Fischbacher, 2003; Fehr & Henrich, 2003; Gintis, Bowles, Boyd, & Fehr, 2003; McCabe, 2003), as well as across cultures (Henrich et al., 2001). In fact, some argue that it is so integral to human society that we have evolved specialized cognitive mechanisms to facilitate reciprocal interactions

including especially, the detection of cheaters (Cosmides & Tooby, 1992) and punishment (Fehr & Gächter, 2000).

Given that reciprocity is common in humans, emerges fairly early in development (Harbaugh, Krause, & Liday, 2002; Harbaugh, Krause, Liday, & Vesterlund, 2003), and that the prerequisites appear trivial, should we expect to see it in nonhuman animals? In this chapter, we address this question by developing the following argument. First, we argue that the prerequisites for reciprocal altruism have been underestimated. A careful dissection reveals a host of underlying mechanisms that may be necessary for both initiating a reciprocal relationship and for maintaining it over the long haul. Second, we argue that some of the essential psychological ingredients for reciprocation include numerical quantification, time estimation, delayed gratification, detection and punishment of cheaters, analysis and recall of reputation, and inhibitory control. For example, reciprocal altruism requires inhibitory control in order to suppress the temptation to cheat: once B has received from A, B must inhibit the temptation to defect in order to return the favor and maintain a stable cooperative relationship. Reciprocal altruism also requires, by definition, patience: if A gives to B, A must wait some period of time before B returns the favor. And reciprocal altruism may require quantification: A and B may quantify the resources exchanged, potentially across different currencies, in order to evaluate whether the exchange was fair. If reciprocity is, indeed, as cognitively complex as we suggest, then we must anchor our theoretical predictions about adaptive function in realistic constraints imposed by neural and psychological design features. Third, because of these limitations, we predict that reciprocal altruism will be rare in the animal kingdom, and when it appears, will represent a relatively minor force in the evolution of social organizations. Returning to our opening paragraph, although reciprocal altruism represents a theoretical solution to the problem of altruism between unrelated individuals, does the crocodile, lion, or chimpanzee have the cognitive wherewithal for reciprocity?

We begin by critically discussing the empirical evidence for reciprocity in animals. For some, we will appear highly critical. We feel this level of analysis is necessary in order to show what we know about cooperation in animals and what we have yet to learn; hopefully, the fact that we are equally critical of our own work as that of others, will make the exercise seem fair. This discussion leads to the conclusion that animals can maintain stable cooperation for mutual, simultaneous benefits, but rarely if ever sustain stable, reciprocally altruistic relationships that entail delayed benefits. We next turn to an explanation of this conclusion, focusing on a suite of cognitive constraints. Finally, we turn to a brief discussion of some of the neurophysiological substrates that might support reciprocation in humans, and use this evidence to speculate about the neural correlates of cooperation in animals.

Reciprocal Altruism: Theoretical Concerns

After Trivers's initial investigation, the concept of reciprocity remained all but untouched until Axelrod and Hamilton's (1981) description of "tit-for-tat" as a possible reciprocal strategy that allows for stable cooperation in the Prisoner's Dilemma (Flood, 1958; Rapoport & Chammah, 1965). In the Prisoner's Dilemma, two individuals each have the opportunity to cooperate with or defect against each other, resulting in four possible fitness payoffs for each player (figure 9.1). Mutual cooperation results in a moderate reward (*R*), but mutual defection leads to very low payoffs for both players (*P*). When one cooperates and the other defects, the defector receives the largest possible reward (*T*) and the cooperator receives the smallest possible reward (*S*). Therefore, the optimum strategy (Nash equilibrium) for playing a single-shot game is to defect, because defection results in a higher payoff than cooperation regardless of the opponent's choice. Axelrod and Hamilton, however, suggested that stable cooperation can emerge if the game is played repeatedly, the opening move is nice (cooperative), and from that point on, each player copies the other's moves. This winning strategy is a version of reciprocity called tit-for-tat. Following this analysis, a flood of theoretical investigations emerged, some confirming the efficacy of reciprocity, others providing alternative strategies that maintain cooperation (reviewed in Dugatkin, 1997).

Empirical Evidence for Reciprocal Behavior

Following the deluge of theory, reciprocity was invoked to explain many instances of animal cooperation. Here we describe case studies that examined putative reciprocal situations, including blood sharing in vampire bats, cooperative games in blue jays, and food exchange in capuchin monkeys and tamarins. Because of space constraints, we leave out the many other interesting cases that have been described, including the

Against player 2:

Payoff to Player 1:		Cooperate	Defect
	Cooperate	*R*	*S*
	Defect	*T*	*P*

Figure 9.1
Fitness payoffs in Prisoner's Dilemma game—payoffs *R*, *T*, *P*, and *S* represent payoffs to the row player against the column player (e.g., *T* is the payoff to a defector playing against a cooperator). To qualify as a Prisoner's Dilemma, *T* > *R* > *P* > *S*. Therefore, regardless of the opponent's choice, defection results in larger payoffs, but mutual cooperation is more profitable than mutual defection.

exchange of grooming for alliance support in vervet monkeys (Seyfarth & Cheney, 1984) and reciprocal grooming in antelope and many primates (Hart & Hart, 1992; Muroyama, 1991).

Blood Sharing in Vampire Bats

Vampire bats can live for almost 20 years, spending much of their time in large, stable social groups where there are multiple opportunities to interact with the same individuals. A vampire bat's survival depends critically on the consumption of blood. If an individual goes for more than 60 hours without a blood meal, it dies. On any given day, therefore, an individual must either obtain its own meal or convince another bat to regurgitate some of its undigested blood. This suite of attributes makes vampire bats ideal subjects for studies of reciprocal altruism (Wilkinson, 1984).

Wilkinson (1984) observed more than 100 regurgitations in a wild population of vampire bats. Because blood is valuable, giving it up represents a cost—an act of altruism. Of the cases observed, most were between mother and infant. These were not examined in any detail because there's no puzzle: regurgitating to your offspring makes sense since you share half of your genes with them; there is no expectation of reciprocation here. Of the remaining regurgitations, only 20 percent were between more distant relatives or nonrelatives, and of these, most occurred among bats that frequently spent time together. However, since many of these individuals were genetic relatives (half at the level of grandparent-grandchild), it seems that regurgitation is largely motivated by kinship, with an extremely small proportion of cases among genetically unrelated bats. Nonetheless, given that some regurgitations were delivered to non-kin, these cases require some explanation. There are two possibilities: either some bats made mistakes, failing to recognize their kin and thus accidentally giving blood to non-kin, or they purposefully gave blood to non-kin with the expectation that they would receive blood back in the future.

To better understand what motivates regurgitations among non-kin, and to clarify whether giving is contingent upon receiving, Wilkinson (1984) conducted a simple experiment with eight unrelated vampire bats. Over many days, he removed one bat from the colony before feeding while providing the other bats with two hours of access to blood. He then returned the now starving bat to the blood-satiated bats. The pattern of blood sharing was clear: individuals regurgitated blood to those who had regurgitated to them in the past.

There are four reasons why we want to express caution in accepting the vampire bat case as evidence of reciprocal altruism, even though many authors have trumpeted these observations as some of the best evidence to date (Dugatkin, 1997; Hauser, 2000). One: the number of naturally observed cases is small and can be explained as errors of recognition as opposed to reciprocation among non-kin. Though regurgitations are given to unrelated animals, these are infrequent, and there is no evidence that indi-

viduals recognize the recipients as non-kin as opposed to kin. Wilkinson did not conduct any tests to show that bats recognize their kin, and if so, to what degree of relatedness. The consequence of contingent regurgitation may benefit non-kin, but the payoffs and mechanisms may have evolved for kin, occurring among non-kin as a by-product. Two: the number of experimental cases is also small, and might reflect an artificial outcome, an exchange that is more important in captivity when bats have less certainty with respect to their next meal. Three: even if we accept these few cases, it is not at all clear whether reciprocal altruism among non-kin plays a significant or trivial role in individual survival. The fact that individuals need blood to survive is clear. Whether or not they depend upon reciprocation with non-kin to survive is a different issue. It may well be that individuals would survive fine without it, relying on their own skills, and the good nature of their relatives. Four: only one study has ever attempted to replicate these findings, even though 20 years has elapsed since their original publication. Denault and McFarlane (1995) observed regurgitations among vampire bats, but the degree of relatedness was close to the level of grandparent-grandchild, thereby allowing kinship to account for the pattern of altruistic behavior.

Experimental Games in Blue Jays

A second approach to testing for reciprocal altruism in animals comes from work on captive blue jays trained to peck keys in one of two classical economic games. In most instances of cooperative behavior in animals, the fitness costs and benefits of cooperation remain unclear. This ambiguity makes it difficult to determine whether the animals face a true Prisoner's Dilemma or whether some other benefits to cooperation actually reduce the temptation to cheat. To circumvent these uncertainties, Stephens and colleagues (Clements & Stephens, 1995; Stephens, McLinn, & Stevens, 2002; Stevens & Stephens, 2004) have devised ecologically artificial but economically relevant situations in which blue jays play cooperative games. Their paradigm involves placing pairs of jays in adjacent operant chambers in which each individual simultaneously chooses to cooperate or defect by pecking a key (figure 9.2A). Once both subjects pecked a key, each received a specific number of pellets associated with the game's payoff matrix. For example, in a Prisoner's Dilemma matrix, when both cooperated each received three pellets (R), when both defect each received one pellet (P), and when one cooperated and the other defected, the defector received five pellets (T) and the cooperator received no pellets (S). In the second game, called "Mutualism," the payoffs for mutual cooperation were higher than for all other possibilities, and the payoffs for mutual defection were lower than for one player defecting and the other cooperating. When the jays played a Prisoner's Dilemma game, they rapidly defected. In contrast, when the jays switched to a game of Mutualism they not only cooperated but maintained this pattern over many days. That jays switch strategies as a function

(a)
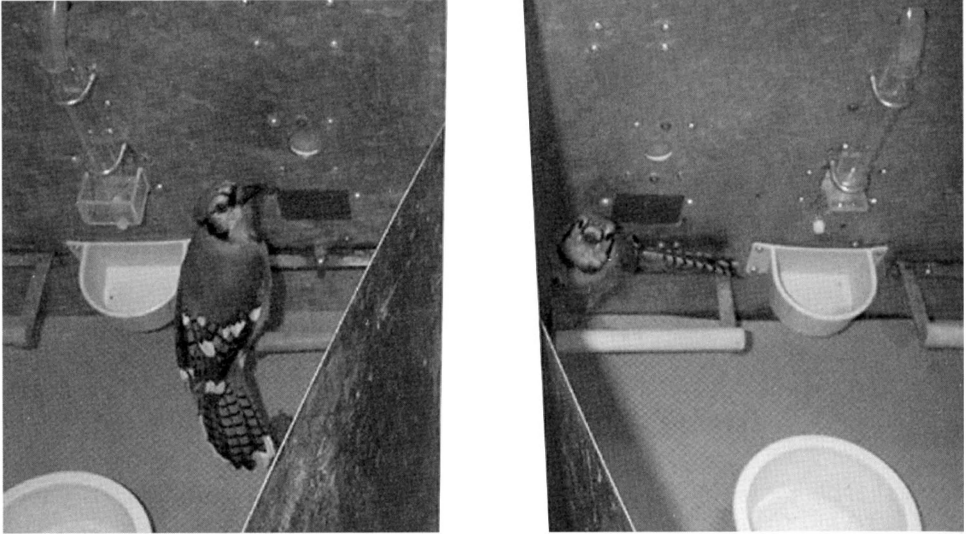
(b)
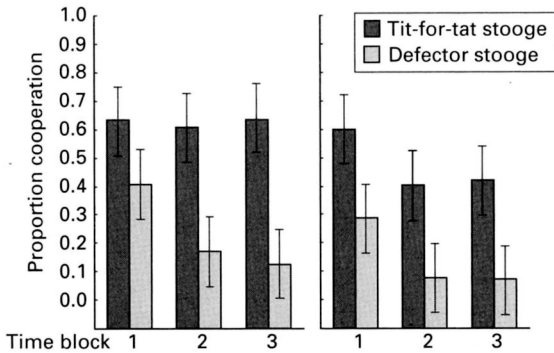
Figure 9.2
Stephens et al. (2002) tested cooperation in an operant experiment with blue jays (*Cyanocitta cristata*) playing Prisoner's Dilemma games. (*A*) Freely behaving subjects played against either tit-for-tat or defector stooges and food rewards were either dispensed after every trial or accumulated over four trials. (*B*) Whereas cooperation decreased over time in all other conditions, cooperation was maintained when subjects played against a tit-for-tat stooge and when food accumulated. This suggests that cooperation depends on opponent strategy as well as temporal discounting. (Adapted from Stephens et al., 2002 with permission from the American Association for the Advancement of Science.)

of the game played shows that their responses are contingent upon the payoffs associated with each game.

Clements and Stephens's results show how cooperation depends upon the relative costs and benefits of different strategies. When cooperation yields the highest possible payoff, jays do not defect. Mutual cooperation is the only reasonable option. In contrast, when there is a temptation to defect, as defined by the Prisoner's Dilemma, then jays are incapable of maintaining a cooperative relationship.

To determine if other conditions might enable cooperation among jays in the Prisoner's Dilemma, Stephens and colleagues ran a second experiment (Stephens, McLinn & Stevens, 2002), this time targeting a potential constraint on the evolution and stability of reciprocal altruism: the temptation to take an immediate benefit outweighs the benefits of waiting for a larger payoff. Numerous studies of animals and humans, discussed more fully below, reveal that waiting for a payoff devalues this item's worth. A small payoff now is better than a large payoff later. This trade-off between time and value is called *discounting*, and is a central idea in economic models of choice. In the original jay work, pecking brought an immediate payoff of some amount. In the new study, payoffs accumulated. To obtain food, each pair of jays had to play several rounds with their partner before obtaining the payoffs, thereby removing the immediate temptation. In addition, the jays played against a partner that either always defected or played tit-for-tat. The jays cooperated only when the food rewards accumulated and when playing against a tit-for-tat partner (figure 9.2B). They solved the repeated Prisoner's Dilemma.

Clements and Stephens concluded their original paper on jays as follows: "[T]here is no empirical evidence of non-kin cooperation in a situation, natural or contrived, where the payoffs are known to conform to a Prisoner's Dilemma" (p. 533). The follow-up studies with jays led Stephens and colleagues to a different conclusion, but one that is consistent with the idea that animals are incapable of maintaining reciprocal relationships under natural conditions: "Our work suggests that the timing of benefits can be the difference between stable cooperation and cooperation that erodes to mutual defection" (p. 2218). The authors also point out that "the experimental machinations required to stabilize cooperation . . . are special" (p. 2218). In other words, nature may never provide animals with the right conditions for reciprocally stable relationships.

Food Sharing in Capuchins

A third example comes from a social primate—the capuchin monkey. Capuchins live in multimale, multifemale social groups with a polygynous mating system. They are a highly dexterous species, and in the wild, hunt in groups and often share food. De Waal attempted to capitalize on their apparent social intelligence by conducting a series of experiments on cooperation. In the first experiment (de Waal & Berger, 2000),

female capuchins had to work for food, either on their own or with another unrelated individual. The task was simple: pull a rod to bring a cup of food within reach. When there were two capuchins and therefore two rods, each individual had to pull at the same time in order to bring the cups within reach. When the experimenter placed food in both cups, both capuchins pulled. Although their joint action is cooperative, it can more readily be explained as selfish, with each individual pulling for herself. When the experimenter placed food in only one cup, the individual lined up with the food almost always pulled whereas the other individual pulled less frequently. Importantly, however, when the player facing an empty cup pulled, she was more likely to obtain food from the other capuchin than when she failed to help. Individuals with access to the food cup rarely handed food to helpers. Instead, they approached helpers and allowed them to grab pieces of food through the partition as it fell to the ground.

Allowing another individual to *take* food is psychologically different from *giving* food. For one, tolerated taking is more difficult to interpret with respect to the motivations or intentions of the possessor. On some occasions, perhaps the possessor did not intend to have food taken but was simply not swift enough to stop the action. Giving, in contrast, is clear-cut, and represents a cost of physical exertion as well as a reduction in one's own resources. In almost 10,000 observations of food transfer, less than 1 percent involved giving. Nonetheless, this experiment shows that capuchins are more likely to tolerate food-taking by an individual who helped them pull in the past.

To explore the contingency part of the capuchin's interactions, as well as the role of food quality, de Waal and colleagues ran other experiments (de Waal, 1997). Individuals were more likely to tolerate food taking when lower quality food items were at stake. Among female-female pairs, individual A was more likely to allow individual B to take food if on the previous run, individual B allowed A to take food. This relationship or correlation accounted for less than 10 percent of the variation in behavior, suggesting that many other factors influence whether or not two females tolerate food-taking; for example, since de Waal did not observe the players before or after the game, we do not know if tolerated food-taking was *repaid* in some other currency such as grooming. Moreover, if two males or a male and female played this game, then tolerated food-taking was not at all contingent on prior food-taking.

A second complication associated with the capuchin studies is that the analyses focus on tolerated food-taking, independently of how much food was taken. Although de Waal reports that the food possessor generally "ate the lion's share," it might well be important to know how much food the nonfood possessor obtained in each of the conditions, as opposed to whether or not it received any food. One might imagine, for instance, that the amount of food taken depends on its quality, on the amount obtained in previous runs, on the food possessor's hunger level, expectations about the number of games to be played in the future, and so forth.

De Waal's work shows that capuchins will tolerate food taking from others, and that this behavior has something to do with the help received on the rod-pulling task. Capuchins clearly cooperate. There are, however, three reasons why we believe that the capuchin work falls short of the required evidence for reciprocation. One: although there is some evidence for reciprocated food exchange, it happens infrequently and is restricted to female-female pairs. Two: Female-female cooperation in captivity may be the by-product of kin selection in nature. Although the females in de Waal's groups were genetically unrelated, in nature, females living in a social group are typically kin. Among most primate species, including capuchins, females stay in their natal groups for life, whereas males emigrate; this leads to groups consisting of closely related females and distantly related males. Reciprocation among female-female pairs could, therefore, be an artificial by-product of selection for kin interactions. It is unclear whether it plays any role in natural groups of capuchins, and if it does, whether it is dwarfed by cooperation among kin. Three: because there is little cost to pulling the rod, and food exchange occurs most frequently when food quality is poor (costs of exchange are low), it is not clear that this task involves altruistic actions; neither the pulling by the helper, nor the tolerated taking of low quality food by the owner, are costly. These three points lead, we suggest, to the conclusion that reciprocal altruism is a weak force in capuchin social relationships.

Food-Giving in Tamarins

A final example comes from another social primate—the cotton-top tamarin. Unlike capuchins that live in large social groups, characterized by polygamous mating behavior, tamarins live in small groups characterized by monogamy. Within groups, which consist of the breeding pair and typically, one to two generations of offspring, older offspring help rear the younger ones. Part of the help comes in the form of food sharing (Feistner & Price, 1990; Roush & Snowdon, 2001). To explore the possibility of reciprocal altruism in tamarins, Hauser and colleagues designed a series of experiments focused on the problem of food exchange (Hauser, Chen, Chen, & Chuang, 2003). Each experiment set up a game between unrelated tamarins in which one animal—the actor—could pull a tool to give food to an unrelated recipient without getting any food for self; pulling the tool is thus altruistic (figure 9.3). Why would unrelated tamarins give each other food?

In the first test, an experimenter trained two tamarin stooges to function as players with diametrically opposite roles: one acted as a unilateral altruist, always pulling the tool to give food to its partner, and the other acted as a unilateral defector, never pulling the tool. The reason for training was simple: if tamarins give food to others based on previous acts of kindness, then they should give most to the altruist and least or nothing to the defector. Supporting this prediction, tamarins pulled the tool most often for the altruist and infrequently for the defector. This shows two things:
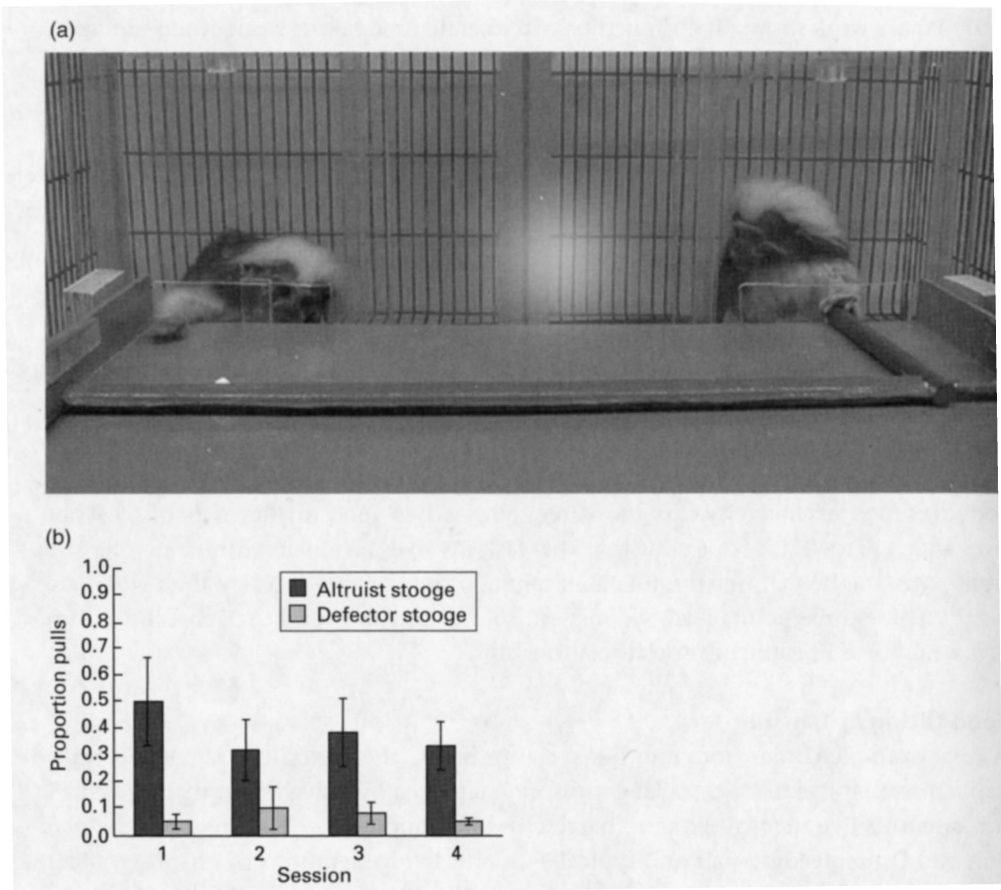
Figure 9.3

Hauser et al. (2003) tested an altruism game in cotton-top tamarins (*Saguinus oedipus*). (*A*) Tamarins had the option to pull a tool, thereby giving food to an unrelated tamarin but receiving nothing in return. Subjects played against an altruist stooge that always pulled or a defector stooges that never pulled. (*B*) The tamarins consistently pulled more often for the altruist stooge than the defector stooge, suggesting that their propensity to cooperate is contingent on their partner's behavior. (Redrawn from Hauser et al., 2003 with permission from the Royal Society of London.)

tamarins give food to unrelated others, and do so based on contingent acts of giving in the past. Is this reciprocal altruism? Not yet. Perhaps tamarins feel more generous when they eat more? When the altruist plays, she gives food on every trial. Her partner must feel good. When a tamarin feels good, it is more likely to pull the tool and give food back. What looks like reciprocation based on an altruistic act of food giving is actually a dumb by-product of feeling good—feeling sated.

To test the "feel good" explanation, Hauser and colleagues ran other experiments, this time leaving the stooges out, and using untrained pairs instead. In one game, player A could pull the tool to obtain a piece of food for self and simultaneously deliver three pieces to player B. On the next trial, player B could pull the tool to give two pieces to player A, but obtain nothing for self. Given these payoffs, reciprocal pulling would pay as each player would obtain three pieces of food after a complete round. Animals in the player A role should always pull out of selfish interest to get food; they did. If feeling good motivates giving food, then player B should cooperate and pull because player A always gives them food—player A looks like a unilateral altruist from one perspective. If player B notices, however, that player A is pulling out of selfish interest—nothing like the unilateral altruist—then player B shouldn't cooperate. Animals in the player B role don't cooperate. Feeling good isn't enough to set recip-rocation in motion. And for food giving to count, it can't be an accidental by-product of selfish behavior. It must be given with an altruistic intent. Although we may be equally happy to acquire a $100 bill from a person who hands it to us and from someone who accidentally drops it from his wallet while running by, most of us would be inspired to cooperate with the first and not with the second. Tamarins see the world in the same way.

Tamarins give food to unrelated others. Giving is not the simple by-product of feeling good. Giving depends on whether food was given in the past, and how it was given. Although these findings address many of the essential ingredients of Trivers's account, a closer look at the patterns of giving reveal the signature of an unstable system. When the tamarins play against the unilateral altruist, they don't respond with unilateral cooperation as one might expect given the level of generosity. They cooperate less than 50 percent of the time, and as each game progresses, the amount of food giving drops. This decline represents the signature of most games of cooper-ation developed by economists. If we repeatedly cooperate to achieve some goal—hunting a gazelle, let's say—then it pays to defect on the last interaction because the relationship is ending. But if I think through this logic right before the last opportu-nity to interact, then I will surely think about defecting on the second to last oppor-tunity, and the third to last, and so on. Cooperation unravels. We see this same unraveling with the tamarins. A further sign of vulnerability comes from the experi-ments with the untrained pairs. Here, cooperation ends if one of the players defects on two consecutive opportunities to pull, a pattern that happens often. Like the jays,

tamarins can maintain some level of cooperation under some restricted conditions. Overall, however, it is an unstable system. And although tamarins naturally give food to each other in the wild and in captivity, providing a certain level of ecological validity, tamarins in nature will rarely have the opportunity to reciprocate with unrelated animals, with the exception of their mates and possibly a sneaky mating with a neighbor. Thus, although tamarins have evolved some of the necessary psychological ingredients for reciprocal altruism—detecting cheaters, calculating contingencies, distinguishing between accidental and intentional actions—it is unlikely that these mechanisms evolved for reciprocal altruism. They may, however, have evolved to solve kin-based interactions.

How Common Is Reciprocity?

Despite these examples and the enormous theoretical interest in cooperation and reciprocity in the Prisoner's Dilemma, very little empirical evidence supports the theory for nonhuman animals. Most instances of putative reciprocity have either not been replicated or can be accounted for by simpler mechanisms. For example, one of the first observations of reciprocity involved alliance formation in olive baboons (Packer, 1977). Packer found that when soliciting help to separate a female from a mating pair, males tended to choose the same partners and alternate which partner solicited help to gain access to the female. A similar analysis on the closely related yellow baboon failed to find this reciprocal relationship (Bercovitch, 1988), and another study suggested that coalition formation is not a Prisoner's Dilemma at all (Noë, 1990). Another case of reciprocity involved predator inspection in fish (Dugatkin, 1991; Milinski, 1987). Milinski and Dugatkin have contended that some species of fish take turns potentially risking their safety by approaching predators. However, others suggest that this behavior can be explained by simpler mechanisms such as group cohesion (Lazarus & Metcalfe, 1990; Stephens, Anderson, & Benson, 1997). None of these point-counterpoint cases settles the issue. They do, however, leave us with doubts concerning the significance of reciprocity for animal social relationships.

Given the theoretical feasibility and the ubiquity of human reciprocity, why do we not find much evidence of nonhuman animals reciprocating (Hammerstein, 2003)? We propose two hypotheses: (1) researchers have not used the appropriate species and/or methodology to find reciprocity or (2) reciprocity is too cognitively demanding for most, if not all nonhuman animals. The first hypothesis is unlikely for a number of reasons. First, theory predicts that reciprocity could apply to a variety of species from bacteria to primates—theoretically, there are few constraints limiting which species should be able to reciprocate. Because of this, cooperation has been investigated in a large number of species including invertebrates, fish, birds, and many mammal species. Second, researchers have used a number of methodological tech-

niques to investigate cooperation, ranging from observation to natural experiments to highly controlled laboratory experiments. The breadth of species and techniques used suggests that the lack of evidence for reciprocity is not because we have not looked carefully.

We propose that reciprocity is a deceptively simple-sounding strategy that is intuitively appealing but cognitively complex. Reciprocity may, in fact, require a complicated suite of cognitive abilities that may limit its utilization by many animal species.

Cognitive Constraints on Cooperation

Cognitive abilities are clearly important in constraining animal behavior, and a resurgence of interest in integrating proximate and ultimate questions proves this point (Krebs & Davies, 1997). To investigate how these constraints influence cooperation and reciprocity, we must break this difficult problem down into its component parts. Here we examine a suite of cognitive abilities necessary to implement reciprocal strategies. Although reciprocity may tap cognitive abilities such as memory, cheater detection, social learning, and theory of mind, we focus on three constraints supported by a considerable amount of evidence: inhibitory control, temporal discounting, and numerical discrimination.

Inhibitory Control

Consider the following deal. In my left hand I am holding a $100 bill and in my right, a $1 bill. I say nothing at all, but give you an inviting wink that suggests you should reach for the money. Your instinct is surely to reach for the $100 bill. When you do, I pull this hand back and offer you the $1 bill. What would happen the second time around? Would you reach for the $100 bill again, or switch strategies and see what happens? Puzzled by the first outcome, you might reach again for the $100, and then rediscover that you are offered the $1 bill. Soon enough, you would switch, reaching for the $1 bill and obtaining the $100 bill. You would now have a hunch about what is going on. You would surmise that the best strategy is to pick the amount you don't want to get the amount you want. You have acquired a new rule to solve this task. But you have accomplished this task by overriding an old rule. We believe that this kind of problem, which requires resolving conflict and inhibitory control, is a core component of reciprocation, and may provide one explanation for why animals have difficulty.

The psychologist Sally Boysen (see chapter 10, this volume) ran this experiment with adult chimpanzees using food treats rather than dollar bills. Boysen assumed that chimpanzees were highly motivated for food and would choose the hand with less food to obtain the hand with more food. Precisely the opposite occurred. The choosers reached for the larger quantity of food and consistently received the smaller quantity.

We can explain this pattern of choice in two ways: either the choosers are exceptionally altruistic, intending to give away the larger stash of food, or they are incapable of controlling their desire to reach for the larger amount of food. The inhibitory challenge here is motivational, as the chimpanzees have evolved brains and stomachs designed to maximize the amount of food consumed. But Boysen's "reverse contingency" task requires the chimpanzees to overcome either an overlearned or innate response to reach for larger quantities of food over smaller quantities. If two simultaneously available patches of food differ only in terms of their quantities, no animal would feed in the patch with less food. But this is precisely what Boysen's task demands. Apparently, these chimpanzees are incapable of overriding the desire to reach for more food.

An alternative explanation is that Boysen's assumption that the chimpanzees were motivated for food was invalid. After all, they are fed every day, and even if they point to the larger quantity of food, they are still rewarded with the smaller quantity. Would they eventually learn to point to the smaller quantity of food if they received nothing at all after pointing to the larger quantity? Can other species solve this task? When other species—Japanese macaques, squirrel monkeys and cotton-top tamarins—are tested on Boysen's task, all fail, picking the larger quantity of food and getting stuck with the smaller quantity (Anderson, Awazu, & Kazuo, 2000; Kralik, Hauser, & Zimlicki, 2002; Silberberg & Fujita, 1996). But if the experimenter imposes a cost, withholding all food when subjects pick the larger quantity, macaques and squirrel monkeys eventually learn to pick the smaller quantity; tamarins stick with the losing strategy, picking the larger and getting nothing at all. This suggests that part of the chimpanzee's failure to point to the smaller quantity is due to the lack of costs associated with pointing to the larger quantity.

To solve Boysen's task, individuals must first inhibit the impulse to reach for the larger quantity and second, reach for the smaller but less desirable quantity. The primates' failure appears to be due to their underlying motivation for more food. This hypothesis is strengthened by a second set of experiments. Boysen ran a different version of the original task with chimpanzees that already knew the Arabic symbols from 1 to 6. This time, instead of choosing between one food treat versus four treats, the chimpanzees chose between a card with the number "1" written on its face and a card with the number "4"; each card covered up the corresponding number of food treats. The chimpanzees quickly learned to pick the number 1 card and received four treats, indicating chimpanzees can learn a rule like, "Point to the one you don't want to get the one you want." Therefore, it seems as though the difficulty of this task results from the chimpanzees' strong motivation to reach for food rather than an inability to learn the reverse-contingency rule.

Although Boysen's task does not pose a cooperative dilemma, it does set up an inhibitory problem that individuals must solve in order to stabilize cooperation. The

chooser must reach for the undesirable over the desirable food quantity, and then wait a return. The first move can be likened to giving away food, the second step to waiting for a reciprocated act (see the following discussion). The first move is costly to self, the second is beneficial. In Boysen's task, the cost appears too great, the inhibitory system too weak.

Temporal Discounting

A related topic to the inhibition problem is temporal discounting (also considered in studies of delayed gratification, impulsivity, and rate maximization)—a devaluing of future rewards. Discounting often results in a preference for smaller, immediate rewards over larger, delayed rewards. For example, imagine that a monkey encounters an unripe fruit. Should it consume the fruit now or wait for it to ripen (Kacelnik, 2003)? Waiting would yield a higher fitness benefit (more sugars are available), but the future is uncertain—another monkey may eat it; winds may knock it into a stream below; a fungus may infest it, spoiling a perfectly good fruit. This uncertainty may have provided a strong adaptive benefit for a preference for immediacy.

Given that the future is uncertain, should all organisms discount in the same way? Although impulsivity is probably universal among animals, the *rate of discounting*— that is, how quickly animals devalue food over time—varies widely across species, ages, and even context. In experiments that estimate discounting rate, subjects are presented with two stimuli: one associated with a small, immediate reward and the other with a large, delayed reward. The discounting rate is "titrated" by incrementally increasing the delay-to-large until subjects are indifferent between choosing the large delayed reward and the small, immediate reward. Therefore, researchers can find indifference points between immediate and delayed rewards over a range of small and large reward amounts. Pigeons and rats both discount future rewards quite highly (Mazur, 1987; Richards, Mitchell, de Wit, & Seiden, 1997)—sometimes devaluing a reward by up to 50 percent in the first second of delay!

Humans have a much lower discounting rate (Rachlin, Raineri, & Cross, 1991); it is by no means constant, however. In fact, impulsivity changes with age—children are much more impulsive than adults (Green, Myerson, & Ostaszewski, 1999). Mischel and colleagues have actually followed children to adulthood, measuring their impulsivity longitudinally (Mischel, Shoda, & Rodriguez, 1989). They have found that, although impulsivity decreases with age, impulsive children tend to develop into impulsive adults. In addition, impulsivity at a young age provides a reasonable predictor of future intelligence, social responsibility, resistance to temptation, and response to stress.

Discounting is not necessarily a static parameter that applies to any choice situation. Rather, it can change choice preferences in different situations, thus the discounting rate is context-dependent. For example, blue jays are usually quite impulsive

birds, preferring immediate to delayed rewards (Stephens & McLinn, 2003). In the autumn, however, jays switch from consuming every acorn they encounter to caching them behind tree bark or under leaf litter. This example of context-specific discounting is common across a number of bird and mammal species. The economic paradigm used to assess discounting also has profound effects on choice. Stephens and McLinn (2003) found that in three potentially equivalent economic scenarios,[2] blue jays acted impulsively in one and exhibited more self-control in the other two. The time from choice to reward greatly influenced discounting behavior in the jays even when the overall reward rate was the same. The extreme variation found in human discounting rates (Frederick, Loewenstein, & O'Donoghue, 2002) may also be attributed to our sensitivity to the economic context.

Many psychologists consider the iterated Prisoner's Dilemma as a type of discounting problem (Green, Price, & Hamburger, 1995; Rachlin, 2000). Individuals can choose between the immediate reward of defecting (and gaining only P fitness units) or the long-term reward of cooperating (and gaining R fitness units). Cross-sectional data on human discounting and cooperation agree with this perspective. After using a titration experiment to establish individual discounting rates, Harris and Madden (2002) found that these discounting rates correlated with cooperation levels—less impulsive individuals cooperated more frequently.

Although Axelrod and Hamilton (1981) included a discounting parameter in their original formulation of tit-for-tat, their parameter only considered the probability of future interactions. Experimental evidence indicates that sheer repetition is not enough to circumvent the discounting problem; other methods must be used to mitigate discounting. As mentioned earlier, Stephens and colleagues (2002) offered repeated Prisoner's Dilemma games to blue jays, using the same general approach as in the Clements and Stephens's studies (figure 9.2). This time, however, they altered the delay to payoff. They only found cooperation when the payoffs accumulated over several trials. This accumulation technique reduced impulsivity in a discounting game as well, suggesting that the jays could only cooperate when their natural discounting tendencies were reduced. Baker and Rachlin (2002) also reduced discounting in pigeons by decreasing the time between Prisoner Dilemma trials. When these intertrial intervals were short, the pigeons cooperated more frequently.

Recent theoretical and empirical evidence indicates that temporal discounting can have profound influences on cooperation. This interface between biology, economics, and psychology provides rich opportunities in which to ask important questions about the nature of social choices, including the mechanisms that both facilitate and constrain them. Beyond the theoretical insights we have sketched, these findings also imply that discounting must be considered when designing appropriate tests of cooperation.

Numerical Discrimination

A challenge for reciprocal interactions lies in quantifying the economics of the entities given and returned, and evaluating whether the exchange was equitable. For example, if an altruist gives four apples and receives one back, this is not equitable, and natural selection should eliminate this poor decision maker from the population. Do animals count or quantify in these ways? If they do, then individuals can at least assess equitable returns. If they don't, then either individuals are satisfied with some return, regardless of amount, or they are open to defectors' giving back less than a fair amount. The simple answer is that animals can count small numbers less than four with precision, and large numbers greater than four with only an approximate sense, with evidence for these two systems anchored in a wealth of behavioral and neurobiological studies (see chapters 6, 7, and 8).

Rats and pigeons can be trained to press a key for food. They can also be trained to press the same key a number of times for food. However, animals make more errors as the required number of presses increases (reviewed in Boysen & Capaldi, 1993). When an experimenter requires a rat to press a key four times for food, it is usually dead on, pressing exactly four most of the time, and on occasion, pressing three or five times. In contrast, when the target number is 30, sometimes the rat presses 30 times, but often it presses somewhere between 20 and 40 times. We observe these patterns when the rat (or pigeon) has to count the number of presses, light flashes, or tones, and when the task changes from counting to waiting a particular period of time before pressing. What these studies show is that animals can count, but only approximately so. In terms of reciprocal exchanges, animals can quantify, approximately, what was given and what was received. But does this map on to anything in the natural world?

There are at least three naturally occurring situations where number would appear to matter: aggressive competitions within and between groups, foraging for food, and reciprocal exchanges of resources in either the same or different currencies. In lions and chimpanzees, two species that attack and kill foreign intruders, individuals attend to the number of competitors. In lions, playback experiments in Tanzania show that females respond more aggressively to one foreigner calling than to three foreigners calling (McComb, Packer, & Pusey, 1994). In chimpanzees, a group of males is more likely to approach and kill a foreign male if the attackers have a three to one advantage (Wilson, Hauser, & Wrangham, 2001). Within groups of dolphins, lions, and many primate species, two or three individuals will form coalitions to defeat either a single dominant individual or a smaller coalition. Although these coalitions involve small numbers, they nonetheless require some capacity to count the number of competitors. And this capacity emerges in a naturally occurring, evolutionarily significant context.

Studies of foraging in animals show that individuals attempt to maximize the rate of energetic returns, picking patches with more over less food (Stephens & Krebs, 1986). Since estimates of rates of return depend on quantifying the amount of food consumed over time, we can ask whether animals count the pieces, guesstimate the volume, or time the foraging periods in a patch. Hauser and colleagues (2000) showed lone rhesus monkeys two opaque empty boxes and then put two pieces of apple into one box and one piece of apple into the other. Subjects consistently picked the box with two apples; they also picked the box with three apples over two, and four apples over three. But when they were presented with five versus four apple pieces, some animals picked four and some five. Without training, rhesus monkeys can count the number of pieces of food, and spontaneously discriminate four from three, but not larger numbers. But there's a problem. Perhaps the monkeys aren't counting at all. Perhaps they are timing how long it takes to load up the box with four apples versus three apples. Since it takes longer to place four pieces than three pieces, their internal timer will tell them to pick the box with four. Although timing is an impressive calculation, and one that plays an important role in the lives of all animals, it is not the same calculation as counting. But a simple experiment shows that number, not time, controls their response. If you place four pieces of apple into one box and three pieces of apple plus a rock into the other, equating time and the number of objects, rhesus pick the box with four pieces of apple. Number, not time, is responsible for the rhesus monkeys' preferences.

Based on an overwhelming number of carefully controlled experiments, it is now fair to say that animals have a number sense. It is a capacity that, in our opinion, consists of two naturally available systems (for a recent review of this literature, see Hauser & Spelke, in press; for a different perspective, see chapter 6, this volume). One allows animals to count up to about four with precision; the second allows them to approximate number, but without any limits on magnitude. Humans, including both young infants and mature adults, also have these two systems. But we have an additional system that relies on language, and in particular, words for numbers. This third system allows us to discriminate any two numbers with precision.

Returning to the problem of reciprocal altruism, we can now give a more specific answer to the question of limits or constraints on reciprocation, especially with respect to quantificational abilities. If and when animals engage in a bout of reciprocal altruism, they will either be limited to small numbers of objects in cases where the exchange must be precise (a banana for a banana), or they will be freed from this constraint where approximate exchanges are tolerated. The same prediction holds for cases where the currency is time, such as the duration of a grooming bout. If one antelope grooms another for 10 minutes, the groomer will most likely accept—as fair exchange, that is—a reciprocated grooming bout of between 8 and 12 minutes. A cheater who shoots for a slightly shorter bout would, in the long run, win. And

this selfish victory might well cause the demise of a potentially cooperative society.

Neural Correlates of Cooperation

Given the cognitive constraints discussed, implementing reciprocal cooperative strategies may prove difficult for many animal species. In cases in which we do see reciprocation, the question remains: are reciprocators integrating all of these domain-general abilities or are they tapping domain-specific cognitive adaptations for cooperation? In other words, are some animals specialized to reciprocate? One way to answer these questions is to delve into the brain to search for neural correlates of cooperative behavior. Although there are studies examining the neural correlates of the cognitive components such as inhibition (Hauser, 1999; Roberts & Wallis, 2000), temporal discounting (Manuck, Flory, Muldoon, & Ferrell, 2003), and numerical competence (Nieder, Freedman, & Miller, 2002; Sawamura, Shima, & Tanji, 2002), there are no studies exploring the neurobiology of cooperation in nonhuman animals. For this reason, here we focus on neuroeconomics—the neurobiology of economic decision making in humans—with the hope that it will shed some light on the possible neural correlates in animals, and minimally, open the door to research in this area. The logic is basically this: if we can document the necessary and sufficient circuitry underlying human cooperation and reciprocal interactions, then this provides one way in which one species solved the problem of reciprocal altruism. Although other animals may solve this problem by means of other circuitry, if animals are incapable of maintaining reciprocally stable relationships, then understanding which part of the circuitry is missing or deficient may help explain why.

Neuroeconomics of Cooperation

The emerging field of neuroeconomics integrates the latest technology of functional neuroimaging and neuronal recordings with classical experimental economics to determine how the brain makes economic decisions. Recently, economists and neuroscientists have collaborated to perform functional neuroimaging on human subjects playing several different cooperative games. In all of these games, individuals can choose to behave selfishly or altruistically. Because each game is either sequential or repeated, recipients of the altruistic option have the opportunity to reciprocate—therefore, these are potentially games of reciprocity. Because this is a nascent field, the data are limited to a few studies, and the causal relationship between brain area and function are still unclear.

Rilling and colleagues (2002) used functional magnetic resonance imaging (fMRI) to scan subjects that played repeated sequences of the Prisoner's Dilemma game shown in figure 9.4A. Subjects played against three partners: a freely behaving human, a

(a)

Prisoner's Dilemma

(b)

Trust game

(c)

Ultimatum game
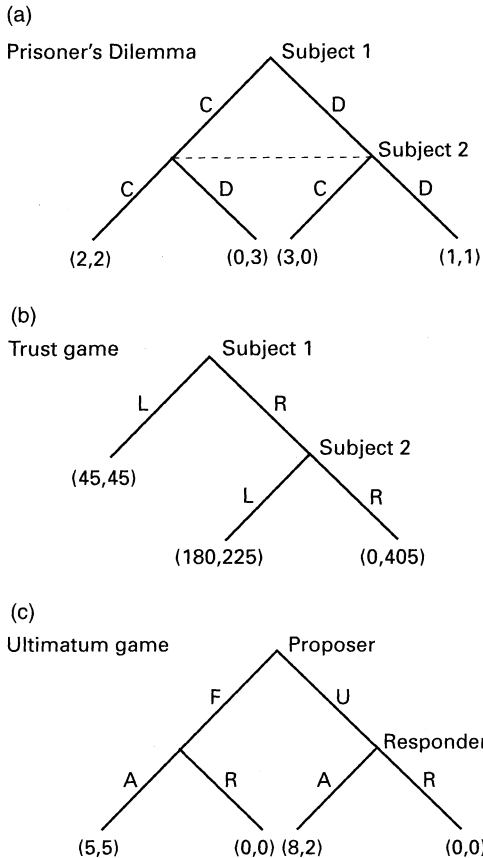
Figure 9.4

Extended form of games in neuroeconomics experiments depicting the sequential choices of the players. The nodes of the trees represent choice points for particular players. Dashed lines connecting nodes indicate that the player does not know her current position in the decision tree. Payoffs are indicated in parentheses for subject 1 and subject 2, respectively. (*A*) In the Prisoner's Dilemma, both subjects simultaneously choose either to cooperate (C) or defect (D). In this example, subject 1 chooses C or D, and subject 2 chooses not knowing subject 1's choice. The equilibrium strategy is mutual defection for a one-shot game. (*B*) In the trust game, subject 1 can choose left (L) and quit the game with a small payoff or choose right (R) to continue the game. Subject 2 can then reciprocate subject 1's trust by choosing L to receive a moderate payoff or break subject 1's trust by choosing R to receive a large payoff. The equilibrium strategy is for subject 1 to choose L and quit the game. (*C*) For the ultimatum game, the proposer submits either a fair (F) or unfair (U) offer. The responder can either accept (A) that offer and distribute the earnings or reject (R) the offer, discarding the earnings for both players. Proposing an unfair offer and accepting any nonzero offer are the equilibrium strategies. Iterated play can lead to stable cooperation and reciprocation in all three games.

human stooge that began cooperating but defected after three consecutive mutually cooperative moves, and a computer playing tit-for-tat. Against both human partners, subjects initially cooperated but reduced cooperation in later trials. When playing the computer, subjects initially defected, then increased cooperation, only to defect in the final trials. Mutual cooperation with human partners was associated with activation in both the anteroventral striatum and the orbitofrontal cortex (Brodmann Area 11) more than the other three possible outcomes; both of these areas have been implicated as playing a significant role in reward (Damasio, 1994; Rolls, 1999). Only the orbitofrontal cortex, however, was activated by mutual cooperation with a computer partner. Reciprocating a partner's previous cooperation increased activation in the anterior caudate, the postcentral gyrus (BA 1/3), the anterior cingulate cortex (BA 25), and the anteroventral striatum. Again, these areas play a role in reward assessment, and significantly, in resolving conflict when there are alternative moves. In our opinion, these data do not provide any additional explanatory power with respect to the psychology and economics of decision making. They do, however, provide new insights into the neural correlates, and these are important with respect to both comparative and developmental data: Are similar areas activated in animals playing the same game? What about human children at different ages?

In the Prisoner's Dilemma, both subjects choose simultaneously and therefore do not know their partner's choice. In a closely related "trust" game, subjects move sequentially, so that the second player to choose can reciprocate the first player's kindness. For example, McCabe and colleagues (2001) imaged subjects playing the trust game shown in figure 9.4B. Subjects played against a human partner and against a computer playing a known, probabilistic strategy. Although they do not describe the behavioral results, previous studies show that the first player cooperates in about 50 percent of the trials, and the second player reciprocates in 75 percent of those trials (McCabe & Smith, 2000). No data were presented on contrasting behavioral responses to human and computer partners. Subjects that cooperated on at least one third of the trials showed increased activation in the occipital lobe (BA 17, 18), the parietal lobe (BA 7), the thalamus, the middle frontal gyrus, and the frontal pole (BA 10) when playing against a human rather than a computer. The authors suggest that because of selective activation in the prefrontal cortex specifically when cooperating with a human partner, the trust game may recruit theory of mind modules such as shared-attention mechanisms. Although intriguing, these results leave several issues unresolved. For example, does the trust game require a theory of mind? Would autistics, who show a deficit in theory of mind type tasks, make similar decisions as do nonautistics in the trust game? What about normal children, below the age of approximately five years, who have yet to acquire a full-blown theory of mind? Given that the prefrontal cortex also plays a significant role in inhibition (see above), and that one of the problems facing both autistics and young children with respect to theory of mind

tasks is an inhibitory one, could prefrontal activation reflect inhibitory mechanisms as opposed to theory of mind systems? Why do the other areas of the brain show high levels of activation only against human partners?

A commonly analyzed cooperative game—the Ultimatum Game—examines how individuals value fairness. In this game, the first player (the proposer) is given an amount of money to split between herself and the second player (the responder). After proposing a split, the responder can either accept the offer or reject it, thereby preventing either player from receiving any money (figure 9.4C). Economic theory predicts that the proposer's offer should be the smallest possible amount to the responder, and the responder should accept any positive offer. After all, both players benefit—why look a gift horse in the mouth? Surprisingly, in experimental game situations, proposers tend to offer approximately 50:50 splits, and responders often reject offers lower than 20 percent of the stake (reviewed in Camerer, 2003).

Sanfey and colleagues (2003) scanned subjects playing ultimatum games against human and computer partners. The human and computer partners acted only as proposers, playing a fixed strategy of offering either fair 5:5 splits of ten dollars or offering unfair splits of 9:1, 8:2, or 7:3. Subjects acted as responders, accepting the fair offer and increasing rejection as the offers became more unfair; however, they treated human and computer proposers differently by rejecting more unfair offers from human partners. Unfair offers from human partners were associated with activation in the bilateral anterior insula, dorsolateral prefrontal cortex (BA 46), and anterior cingulate cortex (BA 24/32); again, these areas are significantly involved in emotional regulation, choice and in resolving conflict. There was greater activation in the anterior insula for unfair offers from human partners than computer partners. Insular activation was also associated with rejecting unfair offers; the authors contend that this activation corresponds to negative emotional states such as pain, distress, anger, and especially disgust. They also state that activation in the dorsolateral prefrontal cortex indicates a conflict between its "executive control" functions and the insular emotional reactions. At the level of function, these results are not surprising. Playing the Ultimatum Game involves resolving a conflict between selfishly keeping the larger proportion of the initial pot and being generous; it also involves emotion, especially on the part of the recipient who obtains either a fair or unfair offer. It was therefore expected that the circuitry underlying decision making, conflict resolution, and emotion would activate in this task. Nonetheless, this study pinpoints some of the necessary substrates for cooperation, opening the door to both comparative and developmental studies.

Cooperation and the Brain

Since neuroeconomics is in its infancy, there is no clear computational theory predicting how cooperation is processed and represented in the brain (but see McCabe, 2003). This is evident by the piecemeal findings of these first studies of the neuroe-

conomics of cooperation. In all of the games presented here, economic theory predicts selfishness, but experimental results show that people tend to cooperate well above expected levels. Despite these similarities in behavior, the neuroeconomic studies reviewed here all cite different cognitive components in their results: reward-center processing, executive control, emotional centers, and conflict-resolution areas. Perhaps these disparities simply reflect the growing pains associated with integrating neuroscience and economics, especially the psychology of decision making and choice (Glimcher, 2003).

Although neuroimaging studies can provide interesting correlates of behavior, causality is difficult to infer. Neuronal recordings of candidate brain centers in non-humans can offer more direct assessment of neuronal activity and can provide an evolutionary framework for understanding the cooperative brain. One potentially fruitful avenue might be through the mirror neuron system located in the premotor cortex of macaques and humans (see chapter 11 for specific details). This part of the brain is equipped with neurons that fire when an individual performs a particular action or when the subject sees an individual perform the exact same action. These neurons fire for action and perception. We tentatively propose that they could provide a necessary, but not sufficient piece of circuitry for reciprocal altruism. For example, there are mirror neurons that only respond when a hand grasps a piece of food, rotates around a piece of food, displaces food, or releases food. Other neurons fire only when a hand grasps an object with index finger and thumb, and not at all when a pair of pliers, held by a hand, grasps the object in the same way. Some neurons even fire when the complete trajectory of the action is concealed, thereby causing the animal to infer or predict the intended action relative to the target object. Together, the fine coding of these mirror neurons suggests that the premotor cortex provides a warehouse of motor commands, a library of action manuals. Given the symmetry between action and perception in the tamarin task, it is possible that the mirror neuron system was engaged. Actor A pulls the tool and gives food to animal B. As animal B watches, A's pull triggers a matched response in B, thereby beginning the mirror neuron loop. This explanation might provide the most parsimonious explanation for the origins of this system. Others have argued that it evolved for imitation and theory of mind, two capacities that macaques lack. It would therefore be of interest to run the tamarin reciprocation experiment while individuals are in a scanner. The strong prediction is that seeing one's partner pull will activate the mirror neuron system as the perceiver's action system clicks into gear.

Conclusions

Cooperation is quite common in both human and nonhuman societies (Dugatkin, 1997). We argue, however, that most instances of animal cooperation can be attributed to selfish benefits or indirect benefits via helping kin. True altruistic cooperation

maintained by reciprocity is rare if not absent among animals, despite its ubiquity in humans. We propose that cognitive constraints on animal inhibition, temporal discounting, numerical discrimination, memory, cheater detection, punishment, theory of mind, and other components may limit the ability of many species to implement and maintain reciprocally altruistic strategies. In particular, animals have difficulty inhibiting the tendency to choose large amounts of food when available. This inhibition problem could pose a challenge for making altruistic decisions that require forgoing large rewards for smaller rewards. In addition, animals often highly discount future rewards. The extreme preference for immediacy exhibited by many species makes waiting for reciprocated rewards very difficult. Finally, precise numerical competence in animals is restricted to small quantities—larger quantities are estimated. When exchanging rewards in cooperative situations requires precision, the quantities in question may be limited to small numbers. These and other faculties may be necessary components of our capacity to reciprocate. If correct, then comparative research must illuminate which components are shared with other animals, which are unique to humans, and why certain components evolved in our species and no other. Although the crocodile, lion, chimpanzee, and Maasai warrior may all cooperate during a hunt, only the Maasai may engage a uniquely human, domain-specific specialization for cooperation.

We can now return to a question raised earlier on: What kinds of cognitive specializations, if any, are required for reciprocation? One approach to answer this question is to assess what happens in the brain when individuals are placed in cooperative games. Although we know little about how animals make decisions in these games, the burgeoning field of neuroeconomics is elucidating the role of the human brain in reciprocal games. Neuroimaging studies of the Prisoner's Dilemma, trust games, ultimatum games, and investment games implicate various areas of the prefrontal cortex, cingulate cortex, and striata among others. While interesting, these studies provide only correlational inferences concerning the relationship between brain activation and decision-making behavior. The timing of activation—whether it occurred at the time of the decision or is a consequence of the decision—is unclear. The real power of these studies lies in their implications for comparative and developmental questions. Does activation correlate with decision making in nonhuman animals playing similar games? Does activation change over the lifespan of an individual human? That is, as these brain centers come online in children, how does their decision making change? These types of questions get at the heart of a cognitive theory of reciprocity.

These cognitive and neurobiological analyses provide interesting insights into the economics, psychology, and evolution of altruistic cooperation. Ultimately, understanding the nature of human cooperation will require cooperation among disciplines.

Acknowledgments

Notes

1. A growing group of economists and anthropologists have argued that Triversian reciprocity is actually selfish as the initial altruistic act is made with the explicit expectation that the recipient will return the favor. Humans appear to have evolved a different, and apparently unique form of cooperation called *strong reciprocity*, defined recently by Gintis and colleagues (2003) as a "predisposition to cooperate with others and to punish those who violate the norms of cooperation, at personal cost, even when it is implausible to expect that these costs will be repaid either by others or at a later date."

2. By "potentially equivalent economic scenarios" we mean that the overall intake rate for the animals is the same for each scenario given a particular strategy. For example, if the animal chooses the small reward every time, it would receive the same amount of food over a given time period in all three scenarios.

References

Anderson, J. R., Awazu, S., & Kazuo, F. (2000). Can squirrel monkeys (*Saimiri sciureus*) learn self-control? A study using food array selection tests and reverse-reward contingency. *Journal of Experimental Psychology: Animal Behavior Processes*, *26*(1), 87–97.

Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. *Science*, *211*(4489), 1390–1396.

Baker, F., & Rachlin, H. (2002). Self-control by pigeons in the prisoner's dilemma. *Psychonomic Bulletin and Review*, *9*(3), 482–488.

Bercovitch, F. (1988). Coalitions, cooperation and reproductive tactics among adult male baboons. *Animal Behaviour*, *36*(4), 1198–1209.

Boysen, S. T., & Capaldi, E. J. (1993). *The Development of Numerical Competence: Animal and Human Models*. Hillsdale, NJ: Erlbaum.

Camerer, C. (2003). *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton, NJ: Princeton University Press.

Clements, K. C., & Stephens, D. W. (1995). Testing models of non-kin cooperation: Mutualism and the Prisoner's Dilemma. *Animal Behaviour*, *50*(2), 527–535.

Cosmides, L., & Tooby, J. (1992). Cognitive adaptations for social exchange. In J. H. Barkow, L. Cosmides, & J. Tooby (Eds.), *The Adapted Mind: Evolutionary Psychology and the Generation of Culture* (pp. 163–228). Oxford: Oxford University Press.

Damasio, A. R. (1994). *Descartes' Error: Emotion, Reason and the Human Brain*. New York: Putnam.

De Waal, F. B. M. (1997). Food transfers through mesh in brown capuchins. *Journal of Comparative Psychology, 111*(4), 370–378.

De Waal, F. B. M., & Berger, M. L. (2000). Payment for labour in monkeys. *Nature, 404*(6778), 563.

Denault, L. K., & McFarlane, D. A. (1995). Reciprocal altruism between male vampire bats, *Desmodus rotundus. Animal Behaviour, 49*(3), 855–856.

Dugatkin, L. A. (1991). Dynamics of the Tit for Tat strategy during predator inspection in the guppy *Poecilia reticulata. Behavioral Ecology and Sociobiology, 29*(2), 127–132.

Dugatkin, L. A. (1997). *Cooperation Among Animals: An Evolutionary Perspective*. New York: Oxford University Press.

Fehr, E., & Fischbacher, U. (2003). The nature of human altruism. *Nature, 425*(6960), 785–791.

Fehr, E., & Gächter, S. (2000). Cooperation and punishment in public goods experiments. *American Economic Review, 90*(4), 980–994.

Fehr, E., & Henrich, J. (2003). Is strong reciprocity a maladaptation? On the evolutionary foundations of human altruism. In P. Hammerstein (Ed.), *The Genetic and Cultural Evolution of Cooperation* (pp. 55–82). Cambridge, MA: MIT Press.

Feistner, A. T. C., & Price, E. C. (1990). Food-sharing in cotton-top tamarins (*Saguinus oedipus*). *Folia Primatologica, 54*(1–2), 34–45.

Flombaum, J. I., Junge, J. A., & Hauser, M. D. (in press). Rhesus monkeys (*Macaca mulatta*) spontaneously compute addition operations over large numbers. *Cognition.*

Flood, M. M. (1958). Some experimental games. *Management Science, 5*(1), 5–26.

Frederick, S., Loewenstein, G., & O'Donoghue, T. (2002). Time discounting and time preference: A critical review. *Journal of Economic Literature, XL*(2), 351–401.

Gintis, H., Bowles, S., Boyd, R., & Fehr, E. (2003). Explaining altruistic behavior in humans. *Evolution and Human Behavior, 24*(3), 153–172.

Glimcher, P. W. (2003). *Decisions, Uncertainty, and the Brain*. Cambridge, MA: MIT Press.

Green, L., Myerson, J., & Ostaszewski, P. (1999). Discounting of delayed rewards across the life span: Age differences in individual discounting functions. *Behavioural Processes, 46*(1), 89–96.

Green, L., Price, P. C., & Hamburger, M. E. (1995). Prisoner's dilemma and the pigeon: Control by immediate consequences. *Journal of the Experimental Analysis of Behavior, 64*(1), 1–17.

Hammerstein, P. (2003). Why is reciprocity so rare in social animals? A protestant appeal. In P. Hammerstein (Ed.), *Genetic and Cultural Evolution of Cooperation* (pp. 83–94). Cambridge, MA: MIT Press.

Harbaugh, W. T., Krause, K., & Liday, S. G. (2002). Children's bargaining behavior: differences by age, gender, and height. Working paper. Department of Economics, University of Oregon.

Harbaugh, W. T., Krause, K., Liday, S. G., & Vesterlund, L. (2003). Trust in children. In E. Ostrom & J. Walker (Eds.), *Trust and Reciprocity: Interdisciplinary Lessons from Experimental Research* (pp. 302–322). New York: Russell Sage Foundation.

Harris, A. C., & Madden, G. J. (2002). Delay discounting and performance on the prisoner's dilemma game. *Psychological Record, 52*(4), 429–440.

Hart, B. L., & Hart, L. A. (1992). Reciprocal allogrooming in impala, *Aepyceros melampus. Animal Behaviour, 44*(6), 1073–1083.

Hauser, M. D. (1999). Perseveration, inhibition and the prefrontal cortex: A new look. *Current Opinion in Neurobiology, 9*(2), 214–222.

Hauser, M. D. (2000). *Wild Minds: What Animals Really Think.* New York: Henry Holt.

Hauser, M. D., Carey, S., & Hauser, L. B. (2000). Spontaneous number representation in semi-free-ranging rhesus monkeys. *Proceedings of the Royal Society of London, Series B, 267*(1445), 829–833.

Hauser, M. D., Chen, M. K., Chen, F., & Chuang, E. (2003). Give unto others: Genetically unrelated cotton-top tamarin monkeys preferentially give food to those who altruistically give food back. *Proceedings of the Royal Society of London, Series B, 270*(1531), 2363–2370.

Hauser, M. D., & Spelke, E. S. (in press). Evolutionary and developmental foundations of human knowledge: A case study of mathematics. In M. Gazzaniga (Ed.), *The Cognitive Neurosciences III.* Cambridge, MA: MIT Press.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., et al. (2001). In search of Homo economicus: Behavioral experiments in 15 small-scale societies. *American Economic Review, 91*(2), 73–78.

Kacelnik, A. (2003). The evolution of patience. In G. Loewenstein, D. Read, & R. Baumeister (Eds.), *Time and Decision: Economics and Psychological Perspectives on Intertemporal Choice* (pp. 115–138). New York: Russell Sage Foundation.

Kralik, J. D., Hauser, M. D., & Zimlicki, R. (2002). The relationship between problem solving and inhibitory control: Cotton-top tamarin (*Saguinus oedipus*) performance on a reversed contingency task. *Journal of Comparative Psychology, 116*(1), 39–50.

Krebs, J. R., & Davies, N. B. (1997). *Behavioural Ecology: An Evolutionary Approach* (4th ed.). Oxford: Blackwell Science.

Lazarus, J., & Metcalfe, N. B. (1990). Tit for Tat cooperation in sticklebacks: A critique of Milinski. *Animal Behaviour, 39*(5), 987–989.

Manuck, S. B., Flory, J. D., Muldoon, M. F., & Ferrell, R. E. (2003). A neurobiology of intertemporal choice. In G. Loewenstein, D. Read, & R. Baumeister (Eds.), *Time and Decision: Economics and Psychological Perspectives on Intertemporal Choice* (pp. 139–172). New York: Russell Sage Foundation.

Mazur, J. E. (1987). An adjusting procedure for studying delayed reinforcement. In M. L. Commons, J. E. Mazur, J. A. Nevin, & H. Rachlin (Eds.), *Quantitative Analyses of Behavior: The Effect of Delay and of Intervening Events on Reinforcement Value* (vol. 5, pp. 55–73). Hillsdale, NJ: Erlbaum.

McCabe, K. A. (2003). A cognitive theory of reciprocal exchange. In E. Ostrom & J. Walker (Eds.), *Trust and Reciprocity: Interdisciplinary Lessons from Experimental Research* (pp. 147–169). New York: Russell Sage Foundation.

McCabe, K. A., Houser, D., Ryan, L., Smith, V. L., & Trouard, T. (2001). A functional imaging study of cooperation in two-person reciprocal exchange. *Proceedings of the National Academy of Sciences, 98*(20), 11832–11835.

McCabe, K. A., & Smith, V. L. (2000). A comparison of naive and sophisticated subject behavior with game theoretic predictions. *Proceedings of the National Academy of Sciences, 97*(7), 3777–3781.

McComb, K., Packer, C., & Pusey, A. E. (1994). Roaring and numerical assessment in contests between groups of female lions, *Panthera leo. Animal Behaviour, 47*(2), 379–387.

Milinski, M. (1987). Tit for Tat in sticklebacks and the evolution of cooperation. *Nature, 325*(6103), 433–435.

Mischel, W., Shoda, Y., & Rodriguez, M. L. (1989). Delay of gratification in children. *Science, 244*(4907), 933–938.

Muroyama, Y. (1991). Mutual reciprocity of grooming in female Japanese macaques, *Macaca fuscata. Behaviour, 119*(3–4), 161–170.

Nieder, A., Freedman, D. J., & Miller, E. K. (2002). Representation of the quantity of visual items in the primate prefrontal cortex. *Science, 297*(5587), 1708–1711.

Noë, R. (1990). A veto game played by baboons: A challenge to the use of the Prisoner's Dilemma as a paradigm for reciprocity and cooperation. *Animal Behaviour, 39*(1), 79–90.

Packer, C. (1977). Reciprocal altruism in *Papio anubis. Nature, 265*(5593), 441–443.

Rachlin, H. (2000). *The Science of Self-Control.* Cambridge, MA: Harvard University Press.

Rachlin, H., Raineri, A., & Cross, D. (1991). Subjective probability and delay. *Journal of the Experimental Analysis of Behavior, 55*(2), 233–244.

Rapoport, A., & Chammah, A. N. (1965). *Prisoner's Dilemma: A Study in Conflict and Cooperation.* Ann Arbor: University of Michigan Press.

Richards, J. B., Mitchell, S. H., de Wit, H., & Seiden, L. S. (1997). Determination of discount functions in rats with an adjusting-amount procedure. *Journal of the Experimental Analysis of Behavior, 67*(3), 353–366.

Rilling, J., Gutman, D., Zeh, T., Pagnoni, G., Berns, G., & Kilts, C. (2002). A neural basis for social cooperation. *Neuron, 35*(2), 395–405.

Roberts, A. C., & Wallis, J. D. (2000). Inhibitory control and affective processing in the prefrontal cortex: Neuropsychological studies in the common marmoset. *Cerebral Cortex, 10*(3), 252–262.

Rolls, E. T. (1999). *The Brain and Emotion*. Oxford: Oxford University Press.

Roush, R. S., & Snowdon, C. T. (2001). Food transfer and development of feeding behavior and food-associated vocalizations in cotton-top tamarins. *Ethology, 107*(5), 415–429.

Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2003). The neural basis of economic decision-making in the ultimatum game. *Science, 300*(5626), 1755–1758.

Sawamura, H., Shima, K., & Tanji, J. (2002). Numerical representation for action in the parietal cortex of the monkey. *Nature, 415*(6874), 918–922.

Seyfarth, R. M., & Cheney, D. L. (1984). Grooming, alliances and reciprocal altruism in vervet monkeys. *Nature, 308*(5959), 541–543.

Silberberg, A., & Fujita, K. (1996). Pointing at smaller food amounts in an analogue of Boysen and Berntson's (1995) procedure. *Journal of the Experimental Analysis of Behavior, 66*(1), 143–147.

Stephens, D. W., Anderson, J. P., & Benson, K. E. (1997). On the spurious occurrence of Tit for Tat in pairs of predator-approaching fish. *Animal Behaviour, 53*(1), 113–131.

Stephens, D. W., & Krebs, J. R. (1986). *Foraging Theory*. Princeton, NJ: Princeton University Press.

Stephens, D. W., & McLinn, C. M. (2003). Choice and context: Testing a simple short-term choice rule. *Animal Behaviour, 66*(1), 59–70.

Stephens, D. W., McLinn, C. M., & Stevens, J. R. (2002). Discounting and reciprocity in an Iterated Prisoner's Dilemma. *Science, 298*(5601), 2216–2218.

Stevens, J. R., & Stephens, D. W. (2004). The economic basis of cooperation: trade-offs between selfishness and generosity. *Behavioral Ecology, 15*, 255–261.

Trivers, R. L. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology, 46*(1), 35–57.

Wilkinson, G. S. (1984). Reciprocal food sharing in the vampire bat. *Nature, 308*(5955), 181–184.

Wilson, M. L., Hauser, M. D., & Wrangham, R. W. (2001). Does participation in intergroup conflict depend on numerical assessment, range location or rank for wild chimpanzees? *Animal Behaviour, 61*(6), 1203–1216.