

The Evolution of Social Capital

On the origins of the social contract

by Katharine Betts

Evolutionary science offers a general explanation for the origins of human beings, but what of the origins of human society? Is there any general theory to explain how society began? The great theorists of the social contract, Thomas Hobbes, John Locke and Jean-Jacques Rousseau, provide us with accounts which, though they differ in a number of important respects, all rest on the theory that at some stage in the past isolated individuals met and decided to form a society. We were once alone, solitary and living in a state of nature, but at some specific juncture, we came together and entered into a social contract with other people.

If we take Locke's version of this story of the origins of society, it is evident that it was devised as a justification for the second English revolution of 1688. He wanted to show that the sovereign, James II, had violated the terms of the social contract and that the people therefore had a right to dissolve his government. Locke's work became a key source of argument for the American revolution nearly 100 years later, and shortly after, for the French revolution. His account of the social contract was clear and influential. But it is wrong.

If to exist in a state of nature is to exist without society, we know that people never were in such a condition. The long dependence of the human infant requires, at least, a society of two — care giver and child. This society must endure for ten to

twenty years and in most cases it is larger than two. Successful human reproduction requires the successful maintenance of human societies. For anatomically modern humans, and doubtless for our hominid ancestors as well, there can never have been a state of nature before society. We are social creatures and we evolved with society. But how could this begin?

Selfish genes and spontaneous cooperation

Evolutionary theory is based on the idea of self-interested competition. Some evolutionary theorists have thought that this competition takes place between groups and others that it is competition between individuals. The current theory, popularized by Richard Dawkins in *The Selfish Gene*, is that it is the genes themselves which lock horns in the struggle to survive.¹ Usually the implicitly selfish orientation of the genes leads to overtly selfish behavior on the part of the individuals who carry them. This is because individuals (microbes, plants, animals, people) which look after their own interests are more likely to propagate their genetic material than individuals which do not.

How could a cooperative society emerge from the chaotic self-interested striving of the most minute particles of living matter? We could argue that individuals who become part of cooperating groups may have better reproductive success than those who do not, but it is hard to see how cooperative behavior could ever have become established. Individuals who paused to help others would have less time and energy to look after their own concerns and would thus lose out to those who were more single minded in their self-interest.² Mutual cooperation may be mutually beneficial but it is hard to see how it can get started. If a would-be cooperator offers a favor to an individual programmed by selfish genes (as all living things must be), the creature who receives the favor

Katharine Betts, Ph.D., teaches at Swinburne University of Technology, Australia. The author of *Ideology and Immigration* (Melbourne University press, 1988), Dr. Betts is a specialist on Australian immigration, population and multiculturalism matters. Her e-mail address is kbetts@swin.edu.au.

should just take it and run.

If societies did not come into being as the result of an original binding agreement between previously scattered individuals, how was it that they emerged at all? This is another way of stating the problem of collective action. How can the best outcomes for a group of individuals be achieved when the collective interests of the group and the individual's interests conflict?³

We should start the search for an answer by focusing on cooperation rather than altruism. It is by no means axiomatic that cooperation is always, or even mostly, unselfish and altruistic in the sense of having detrimental consequences for an individual's personal (and genetic) interests.⁴ It is true that, on some occasions, commitment to group projects may require terrible individual sacrifices but in most cases cooperation pays. While evolutionary, and social, theorists should acknowledge the possibility of individuals exhibiting genuine heroism and self-sacrifice (even if it means a feckless disregard for their genetic posterity),⁵ here we can set our sights on the lower target of simple cooperation. For present purposes *cooperation* is not a synonym for altruism, it is just cooperation. And its opposite is not selfishness but just *non-cooperation*.

Economics and social capital

Moralists, whether they work from Darwinism or from Christian notions of original sin, usually assume that we are born selfish and must be trained rigorously if we are to stifle our innate preference for criminal egotism. The picture of our innate human nature from this perspective looks bleak, but it need not be painted entirely in black. It is an evolutionary truism that the genes which replicate themselves live in bodies which take care that they should do so. We can hold on to this and still see how individual interests can flourish within cooperating groups.

There is, of course, one well-developed model of human behavior which argues that individual selfishness promotes the common good, and that is market economics. Economists are keen to point out that self-interest, regulated by competition, can

generate complex systems of behavior which organize the factors of production (land, labor and economic capital) in such a way as to produce and distribute wealth for the wider good. It is not, said Adam Smith, from my butcher's benevolence that I expect my dinner, but from his self-interest.

But Smith's insight does not take us far enough. We still do not know how societies developed to the point where money and markets became possible. And the combination of the factors of production with egocentric greed can only explain part of the variability between different

nations. This is because economics provides a limited vision of the relations among people and between people and their land. The invisible hand of the market can regulate the production of marketable goods and direct their distribution, but it can also squander natural capital (such as plants, animals, clean air

and water, mineral resources, and the capacity to absorb wastes) and, as contemporary, global "casino capitalism" demonstrates,⁶ it can denude social capital.

Social capital is important. The term has been newly-coined by political analysts but, now that they have named it, most of us can recognize it. *Social capital* stands for the bonds of trust and obligation between human beings which allow non-coercive cooperation.⁷ Without it, there is no social foundation for any economy, market or otherwise, and no common resources to draw on to protect the natural environment. Smith did not take it into account in *The Wealth of Nations* but Matt Ridley argues that his earlier work, the *Theory of Moral Sentiments*, may be central to our understanding of it.⁸

We need to understand how societies can develop without any overt plans or conscious organization, how cooperative behavior between selfish individuals can evolve. If we do, we will know more about the origins of social capital and how our remote ancestors could have used it to solve the problem of collective action.

Game theory and the problem of collective action

Work done by modern game theorists on the

***"We need to understand
how ... cooperative
behavior between selfish
individuals can evolve."***

ancient problem now known as *prisoner's dilemma* provides a new key to the problem of collective action. It can help us imagine how cooperating social groups might have emerged from a pre-human Hobbesian war of all against all. It is not surprising that game theorists conclude that, over the long term, individuals do better if they are part of cooperating groups, and that they do not need to know this fact for it to be true. The interesting part is that these theorists can also show how such groups can spontaneously come together in the first place. But to understand the story we need to grasp the structure of the prisoner's dilemma.

Many social transactions which underpin cooperation involve bargaining. You do this for me and I'll do that for you. There is nothing unselfish about this. However, unless the favors can be evaluated at their face value and exchanged on the spot, there is always the possibility that one of the partners will pocket his gains and renege on the obligation to reciprocate. They will not only fail to cooperate, they will defect. The prisoner's dilemma is an archetypal situation where two people could each either honor their promises to each other to cooperate or defect on their obligations. Darwin used the phrase *natural selection* to emphasize the absence of deliberate intentions in the origins of species. The social organization which now allows humans to make plans and sign binding contracts had to begin among pre-human creatures who were probably incapable of forming long-term goals.

The dilemma is this. The attempts of the two individuals to cooperate and honor their promises (implicit or explicit) are risky. This is because cooperation requires a partner if it is to be successful while defection is unilateral. If a situation mirrors that of the prisoner's dilemma, the rewards for the individual depend on what the partner does. Mutual cooperation is hard to ensure but it would suit the individual well. However, the best possible outcome for the individual in any one encounter is that he or she exploits the partner's trust and defects while the partner who has been betrayed cooperates. (The books robustly declare that, in this situation, the betrayed partner has earned the "sucker's payoff".) If, however, the partner anticipates treachery and defects as well, neither do

very well. They would both have been better off cooperating. The worst situation for the individual is that he or she cooperates but is deceived. The dilemma is this: if you attempt to cooperate you may do quite well but you run the risk of being betrayed and failing bitterly. In contrast, if you defect you may do very well indeed and you only run the risk of moderate failure.

"The social organization which now allows humans to make plans and sign binding contracts had to begin among pre-human creatures who were probably incapable of forming long-term goals."

The logic of the situation was first developed to analyze the options facing a pair of prisoners, each of whom would benefit if he or she confessed (defected) and implicated the other. This is because the defector would go free while the betrayed confederate would serve a substantial sentence. If neither confessed they would both serve short sentences, and if both confessed they would both serve moderate sentences. The tension lies in the fact that the best common outcome (neither defect, both cooperate) is not as good for either individual as is unilateral betrayal.

The literature usually assumes the prisoners are criminals. (Possibly the moral dimensions of their problem would become clearer if we assumed they were green activists imprisoned for their noble deeds by a tyrannical regime, hell-bent on environmental destruction.) For convenience we will call them Peter and Jane. They are put in separate cells and urged to confess. They had worked as a cooperating partnership before, but now they cannot communicate. Will the partnership endure in the face of the temptation to betray? If one confesses but the other does not, the defector will go free while the person who is betrayed gets a 20-year sentence. If both confess, both get 16 years. If neither confesses, both get eight years.

The same logic applies to each of them, but if we look at the situation from Peter's point of view we can see that his best outcome would be that he

confesses while Jane does not. Failing that, his next best outcome would be that he refuses to confess and so does Jane. But if he adopts that strategy he risks the worst outcome for himself which is that he refuses to confess and Jane betrays him. She defects and goes free, leaving him to serve the 20 years.

Because this is game theory, the participants are often called players and, as each player can either defect or cooperate, there are four possible outcomes. Table 1 shows the sentences applying in each. It also assigns a points score to these outcomes which allows us to generalize from this situation to similar problems of collective action which do not involve jails and criminals.

While it only represents a "community" of two, the prisoner's dilemma offers a way of analyzing a situation where there is conflict between an individual's immediate self interest and the community's broader collective interests. Quadrant 1 in the table shows that cooperation yields the highest collective score (6), while unilateral defection in either Quadrant 2 or 3 offers the highest possible individual score (5).

The logic of the dilemma may be easier to understand in the more familiar situation of bargaining in the market place. Peter is selling a second-hand car and Jane is paying for it by check. On-the-spot examination cannot reveal whether the car is really a good car or whether the check is really a good check. Peter's best outcome is that he sells a worthless car and Jane gives him a good check. (He's the defector, Jane wants to be a cooperator but, sad to say, is taken for a fool.) Peter's next-best outcome is that he sells a good car and gets a good check. (They both cooperate.) Next best, for Peter, is that he sells a bad car and

Jane gives him a bad check. (They both defect.) Worst of all, for Peter, is that he sells a good car and gets a worthless check. He wanted to be a cooperator but ended up betrayed. However, the best situation for the community of Peter and Jane together is the second one, that of mutual cooperation.

What of the past? Could such dilemmas have confronted our ancestors before checking accounts and complex consumer goods were devised? Biologists have shown that a variety of creatures (fish, birds, vampire bats) face problems of prisoner's

dilemma,¹⁰ so there is no reason to exempt pre-humans. If one hominid has an extra spear to barter and another has an extra axe, they can each evaluate the soundness of the other's offering there and then and make an exchange without further ado. The problem arises when promises of future behavior are involved. Perhaps they wish to enter into an arrangement to share the risks of failing to catch any game. (Those with strong stomachs can see how the vampire bats manage such arrangements today.)¹¹ The hunters could each search for game in different parts of the forest and implicitly agree to share their catch in the evening. This will maximize the odds that one of them will at least catch something. But defection is possible. One of the hunters could indeed snare a small animal but he could eat it on the spot. Full of assumed disappointment he returns in the evening empty-handed. He has nothing to put into the common pot, but the faithful partner honors his side of the agreement and divides his own catch with his well-fed and treacherous associate. But if both cooperate, they can truly spread the risks of hunting, while if both defect they will sometimes eat well but will often have to endure hunger.

Table 1
Prisoner's dilemma: four possible sets of outcomes

1. <i>Best joint outcome</i> Both cooperate. Jane gets 8 years and scores 3 points; Peter gets 8 years and scores 3 points.	3. <i>Best outcome of all for Jane, worst for Peter.</i> Jane defects and goes free; Peter cooperates, gets 20 years, and scores 0 points.
2. Jane cooperates, gets 20 years, and scores 0 points. Peter defects, goes free, and scores 5 points.	4. <i>Second-best joint outcome.</i> Both defect. Jane gets 16 years and scores one point; Peter gets 16 years and scores one point.

Multiple players, the commons, and the free-rider

Many situations where mutual co-operation is the best collective outcome but unilateral defection offers a higher individual payoff involve more than two people. For example, Garrett Hardin's unmanaged commons¹² is a variant of prisoner's dilemma, but with multiple players. Any one herder's best interest would be served if he alone ran his herd on the commons while everyone else refrained. (He "defects," they "cooperate" and are thoroughly taken down.) The individual herder's

can work at a national level too. One country invests in reducing carbon-dioxide emissions, others go on as usual or even produce more. One country restrains its fishing fleet, others move in and increase their catch. One country avoids nuclear industries, others dump radioactive waste at sea. One country cuts its birth rate, others do not. Some of these defectors may even demand that the lower-birth-rate countries accept their surplus population. In each case the defecting countries "win," leaving others with the fool's payoff and the world as a whole loses.

"Without some rudimentary society to produce and nurture basic social skills, no meeting and discussion would have been possible."

next-best option would be full cooperation; no one grazes his cattle on the commons or, more realistically, limited grazing is shared equally. (This, of course, would be the best outcome for the community of herders and the commons would cease to be unmanaged.) Next best from the individual's point of view is that everyone runs his cattle to the limit — everyone "defects." The individual at least gets something before the final disaster hits. The worst outcome, from the point of view of the individual, would be that everyone else runs his cattle except him. He had hoped to play the role of cooperator but ended up being taken for a fool.

We can imagine other variants. One person, decides not to drive a car in the hope of reducing pollution and congestion. He hopes to play the role of "cooperator" but risks putting valuable personal resources into alternatives (cycling or walking if there is no public transportation) while others speed past, enveloping him in exhaust fumes. Another person limits her family to one or two children. Another spends a lot of time sorting the garbage for recycling. Whether they are idealistic simpletons or useful leaders of a new trend depends on what other people do. The dilemma involving attempts at cooperation risks being met by ruthless exploitation

The free-rider problem is a slightly different form of the problem of collective action.¹³ For example, Helen and Jim do not wish to have their child vaccinated against whooping cough. They are afraid that the vaccine might have side effects. They reason that every one else will have their children vaccinated, so the likelihood of whooping cough being spread is small. If the other parents cooperate, and Helen and Jim defect, they (or, rather, their child)

can gain a free ride at the expense of others. But if most other parents plan the same defection, the situation will be similar to "all defect" in prisoner's dilemma and there will be a lot of infants, and doubtless many parents, all crammed into Quadrant 4 of Table 1 and all sharing a new whooping cough epidemic. There is, however, a difference between the free-rider dilemma and the prisoner's dilemma. Helen and Jim might prefer to be the one-and-only family of defectors, but they would be better off as unilateral cooperators (having their child vaccinated even though no-one else's child was) than they would be in the situation where no-one gets their children vaccinated. There's no "sucker's pay off" for being a unilateral cooperator, just a mild reaction to a vaccine in return for immunity to a disease that has everyone else in its grip. From the individual's point of view, it can be important to work out whether a particular problem of collective action is a prisoner's-dilemma problem or a free-rider problem. But from the community perspective, everyone having their children vaccinated (everyone happy and secure in Quadrant 1), is still the best joint outcome, and free-riders pose a threat to the common good in the same way that defectors do.

The importance of continuity, history and memory

Do people (and countries) who behave well always risk being exploited if a situation has a prisoner's-dilemma structure? If they do, solving the problem of collective action must always mean talking with other "players" and enforcing mutually-agreed rules, establishing the kind of social contract Locke imagined. But if this is true, we are left with the insoluble problem of how human beings ever developed the language and ideas which they would have needed in order to hold their first meeting and set their ground rules. Without some rudimentary society to produce and nurture these basic social skills, no meeting and discussion would

“Moral values rewarding cooperation and punishing defection are more likely to emerge and survive in stable communities where people know each other...”

have been possible. A pre-Lockean society could only have arisen spontaneously, but the prisoner's dilemma seems to render such spontaneous cooperation impossible because it seems to show that defection is always a better risk than cooperation.

When the mathematicians first began working on the problem in the early 1950s, this was the depressing conclusion which they reached.¹⁴ But it turned out to be true only if we take it one game at a time. The outcome is different if we look at repeated or "iterative" games. As computers developed and mathematicians used them to try different approaches to the problem, the results changed. We can now see that it is possible for cooperating groups to emerge spontaneously (if they are small and if the individuals involved know each other well).

The importance of continuity seems clear from our own experience. If Peter only intends to sell one car, and if he expects he will never see Jane again, it is in his interest to pass off one that is worthless. But if he is starting a second-hand car business and hopes to build up a loyal clientele the situation changes. In a similar way, if the hunter wishes to

have an enduring partnership it is in his interest, not only to cooperate, but to demonstrate to his potential collaborator that this indeed is what he will do. It is in his interest to show that he is trustworthy, a pre-human who keeps his implicit promises, a collaborator worth taking into partnership. Moral values rewarding cooperation and punishing defection are more likely to emerge and survive in stable communities where people know each other and most "games," or interactions, are iterative.¹⁵ While we know this from our own experience, it is important to show how it is that selfish, unreflecting, pre-human entities could implicitly arrive at this conclusion for themselves. Computer simulations can now demonstrate that we are right to believe that keeping promises within small communities offers higher rewards than treachery and defection. But how can this be explained in a Darwinian world where selfishness should rule?

Matt Ridley's book, *The Origins of Virtue*, provides a lively account of the computer-based research which has been done on the problem: my discussion of it is drawn from him.¹⁶ The simulations he describes show that in one-off games of prisoner's dilemma, defection always pays (the defector may score 5, but is certain of scoring at least 1 and of avoiding zero). But in repeated games this certain victory evaporates. Continual defection can turn out to be a losing strategy. Tournaments of prisoner's dilemma have been played out in which individual programs designed by different programmers play against each other, scoring points for defection and cooperation according to the logic of the game. (The points normally allocated are those set out in Table 1.) With this scoring system, programs with different strategies can accumulate points and the robustness of each strategy be tested. For example, a program which is set up in such a way that it will always defect will do well against programs set up to always cooperate, because it will routinely score 5 to the other's 0. Soon, however, the strategies which always cooperate will be driven from the field defeated. (Think of the points as food. On a continual diet of zeros the program which always cooperates will starve to death. In contrast the well-fed can survive and, if they were really living creatures, propagate their kind.) But after the habitual cooperators have been

exploited to the point of extinction, the strategies which always defect will only meet others of their own ilk. Their days of plenty will be over because now they will only score 1 point in each encounter.

As the tournament continues, program competes with program, sometimes using cooperation as a strategy and sometimes defection, and electronic circuits hum with artificial life. But the results are easier to visualize if we pretend that the programs are real creatures, perhaps rather limited human beings. Here we are, a collection of strangers meeting for the first time in a field of opportunities where we can make promises about future cooperation, and then honor them or not as we please. If a person meets a particular stranger only once and can then move on to a different field in a foreign country, the hit-and-run-strategy of Always Defect will always pay. The defector may score 5 if the stranger is naive, but the worst he will get is 1. Not for him the humiliation and impoverishment of attempts at cooperation which are exploited and fail; as a callous opportunist the defector is bound to keep on winning. But what strategy works out best if no one can move on? What happens if we are all going to stay in the same field and have a series of encounters with its other inhabitants?

The best long-term strategy would be “always cooperate, so long as the other person is also a cooperator.” Such a strategy would mean that we would consistently score 3, and so would our partners, and that there would be no losers. If we could only ensure this outcome, we could all stay safely in Quadrant 1. But how can we know if another person is playing the same way? He may make promises of cooperation but why should we believe him? Many people with different strategies populate this field; some are trustworthy cooperators, some are predatory defectors. Others have variable strategies; sometimes they defect, sometimes they cooperate. How do we know what they will do to us? How should we behave toward them? As potential partners come together they can promise and cooperate, or they can meet and cheat. After many encounters the *deus ex machina* (or programmer) says “enough” and counts the score, asking not which group or team has won, but which individual (or which program). Who in this game of life won the most points and acquired the

potential to leave the most descendants?

Initial research showed that the winner was a simple, variable program: the tit-for-tat strategy. Tit-for-tat begins every encounter as a cooperator but afterward always responds in the same way as its partner. It rewards cooperation with cooperation and defection with defection. This set of rules — start by being nice but then give back what you get — seemed to defeat all comers. The implication drawn has been that, given enough time, all the

“...some are trustworthy cooperators; some are predatory defectors....How do we know what they will do to us? How should we behave toward them?”

other strategies would perish from lack of points and the field would be totally populated by cheerful tit-for-tatters. These simulations seemed to show how it was that voluntary cooperation could spontaneously evolve in a selfish world. Not only would a creature which played tit-for-tat survive and prosper, it would also leave more descendants than others which were less well-behaved. And assuming that its strategies were passed on to its progeny (either culturally or genetically), its kind would become widespread.

Some commentators then considered the prisoner’s dilemma solved, but others were less happy. As a model for real life it was unsatisfactory. Suppose one tit-for-tat strategy meets another? This should be the beginning of a happy life-time of mutually rewarding cooperation. But suppose the second strategy makes a random mistake. It once, and only once, fails to meet its obligations and, because of this error, replies to cooperation with defection. Its partner then flips over to a retaliatory stance and defects in its turn. A lethal feud ensues. Because of one error, both are grievously impoverished.

In the face of this problem, refinements were introduced. A new program called “Generous” was introduced. Generous played tit-for-tat except that, occasionally and randomly, it responded to defection with cooperation. This allowed the

program the chance of escaping from dreadful feuds which had only been instigated by mistake. Generous proved to be a great success, more robust than the original tit-for-tat. Other refinements produced programs with memories, strings of computer code which could recall the behavior of other chance-met strings of code. The programs became more human-like, and each program could draw on its recollections of the behavior of other programs. For example, Generous could learn to recognize other Generous programs and begin each new encounter with cooperation. Generous could also, being forewarned, avoid giving others with a known history of defection the benefit of the doubt. But these memory-enhanced programs achieved their best results when they could be selective. A new strategy, the Discriminating Cooperator, was developed.¹⁸ This program did not greet Always Defect with defection. Discriminating Cooperator ostracized it. Indeed, Discriminating Cooperator refused to play with any other program which did not have a known history of reliable cooperation. In this way opportunists who always, often or sometimes defected could be isolated. Because of this, unreliable opportunists could now only play with each other. Routinely breaking their agreements for meager rewards, they battered one another other down in the ill-rewarded badlands on the edge of the field, while the circle of virtuous and trustworthy cooperators prospered at the center.

The creation of Discriminating Cooperator might seem like a happy ending to the story of virtual life behind the screen. But as the Discriminating Cooperator programs proliferated, driving out the defectors, a more naive type of program, Always Cooperate, grew in strength among them. After all, Discriminating Cooperator would not defect on Always Cooperate so why should Always Cooperate not prosper too? The two strategies continued to interact amiably and profitably but, slowly, Always Cooperate gained in numbers. All would have been well were it not for the risk that, at some point, a remnant member of the Always Defect gang would return from the fringe and, rather than meeting the firm but fair response of Discriminating Cooperator, it could encounter the naive optimist, foolish Always Cooperate. Inasmuch as these simpletons had increased in numbers, the Always Defect gangsters

would make large inroads. (Even if Discriminating Cooperator had built walls to keep the bad programs out, mutation would mean that, from time to time, one or two of them would appear wolf-like in the midst of the circle of cooperators.) Each time they reappear the newly-arrived crew of Always Defect rules for a while, preying on Always Cooperate, until the defectors are once again reduced by the steady attrition of Discriminating Cooperator discovering them and isolating them.

If we bring our human interpretations to this shadowy artificial world, we can see the Always Cooperate strategy as a free rider. It "wants to be nice" and is able to do so by exploiting the benign environment produced by the more tough-minded Discriminating Cooperator. As long as the Always Cooperate programs are protected from villainy, they can enjoy their niceness. But once evil reappears among them, the curtain comes down on the pleasant fable of universal cooperation which they have been telling to each other.

From artificial life to human societies

These simulations show how cooperative behavior can emerge between very limited creatures, and how a little memory goes a long way in solving the problem of collective action. But there is no final ending, happy or otherwise, in these computer simulations — only a continual oscillation. Does this situation mirror contemporary human interactions? It may. But the size of the cooperating group and the permeability of its boundaries matter.

Ridley emphasizes the importance of numbers,¹⁹ and it is a point that Hardin's analysis of Hutterite communities, in a different context, makes.²⁰ Around about 150 people seems to be the optimal size if we want each member of a social group to know the others well and thus, through social pressure backed by the threat of ostracism, reinforce the norm of cooperation. (Indeed Aristotle made the same point in the fourth century B.C. when he argued that the ideal state should not be so large that its citizens could not know each other.)²¹ Robin Dunbar's research provides more detail. He endorses the finding that 150 is the approximate limit to the number of other people whom we can know well, but also argues that the maximum number whom we can know to the point of feeling deep personal empathy toward them is around 14 while, on average, we can recall around 1500 to 2000 faces.²²

The uneasy relationship between discriminating and indiscriminating cooperators and anti-social predators modeled by the game theorists may be applicable to societies of many millions, but the basic principle of spontaneous mutual aid which they outline cannot apply to large groups. When societies are so large that we cannot hope to know more than a fraction of the members personally, all that game theory can suggest is that defection is more likely and, by implication, that the need for overt rules is more pressing.

It is not just that large groups contain too many people for us to recognize.

Social interaction within them also becomes very much more complex. Thomas Malthus is rightly famous for arguing that human numbers can grow geometrically while resources, if they grow at all, can *not* grow geometrically. His argument is focused on the pressure of population on natural resources. But could exponential growth also threaten social capital? Of course, as competition for natural resources becomes more acute, we would expect social conflict to increase. But do the numbers themselves make a difference to our capacity to manage conflict?

Jared Diamond's research shows that social capital is strained by growth, irrespective of the pressure on natural capital. This because as a population grows the potential number of encounters which an individual may have with other individuals grows at even faster rate. A group of two people contains only one possible dyad, a group of three contains three, a group of four contains six, a group of five contains nine, and a group of six contains 15. As Diamond puts it:

*Relationships within a band of 20 people involve only 190 two-person interactions (20 people times 19 divided by 2), but a band of 2,000 would have 1,999,000 dyads. Each of those dyads represents a potential time bomb that could explode in a murderous argument...*²³

And one-to-one encounters are far from being the only form which social relations take. If large groups are to survive they need rules and they need

specialists with power to enforce them.

But why should a small group become a larger one? While natural increase alone can produce growth in the size of any one group, a group which grew too large could always bifurcate, as indeed the Hutterite groups do. Diamond argues that the threat of war is the principle cause of the development of larger groups. Such threats either foster defensive mergers between allies or, through conquest, force mergers between enemies. Either way, conflict results in numerous smaller groups becoming fewer larger ones, and it is this, he claims, which leads small bands to give

way to larger tribes and then to chiefdoms and then to states.²⁴

Lessons from game theory for social capital

Trust and obligation between known individuals can explain cooperation in small groups. However, these values cannot alone provide a basis for larger numbers of human beings to live in a helpful and cooperative way. They cannot by themselves support the modern nation state and they certainly cannot provide a basis for a world without borders, based on universal sharing and cooperation. But the game theorists have shown how cooperative behavior can evolve among groups of self-interested individuals and how the problem of collective action can promote a desire to be known as trustworthy. This a good beginning. Social theorists less interested in evolution and more concerned with the problems of contemporary societies coined the term *social capital* to describe these very feelings of mutual trust and obligation, arguing that this social capital consists of networks supported by these values. It is built on a predisposition to trust other people and a sense of obligation to behave as someone who may be safely trusted by others. Here game theory, evolution and political analysis meet.

The task now is to preserve the social virtues in a world where numbers of human beings and the complexity of human affairs threaten natural and social capital. Apart from the direct effects of growing human numbers, both types of capital have

“[Malthus’] argument is focused on the pressure of population on natural resources. But could exponential growth also threaten social capital?”

two enemies: people and groups who defect, and the naive who are prepared to cooperate with anyone, even habitual defectors.

As human beings we are not prisoners of our genes and we are not bound by the limits of a computer program. We can understand the problem of collective action and use this understanding to check our growth. We can also use it to restrain, reform and educate those who defect and to enlighten the naive optimists who do not understand the possibility of defection. We can try to persuade more people and more governments to become thoughtful, discriminating cooperators. To the degree that this strategy succeeds, social capital will flourish and more of us can become realistic optimists.

TSC

NOTES

¹ R. Dawkins, *The Selfish Gene*, Oxford University Press, Oxford, 1976 and 1989.

² See G. Hardin, *Living Within Limits: Ecology, Economics and Population Taboos*, Oxford University Press, New York, 1993, p.228.

³ See M. Ridley, *The Origins of Virtue*, Viking, London, 1996, pp.56, 175.

⁴ For the distinction between the motives inspiring altruistic acts and the consequences of these acts see Hardin, 1993, op. cit. p.225. Hardin restricts his discussion to consequences and this is my intention as well. However I have dodged the word *altruism* and replaced it with *cooperation*.

⁵ Genuine altruism, where neither the protagonist nor his or her genes gain any advantage, is always possible. It is just not an evolutionarily stable strategy.

⁶ The term "casino capitalism" has been coined by Susan Strange. See H.-P. Martin and H. Schuman, *The Global Trap*, Zed Books, London, 1997, p.194.

⁷ For more on social capital see R. Putnam, *Making Democracy Work: Civic Traditions in Modern Italy*, Princeton University Press, Princeton, New Jersey, 1993, especially pp.167-169, 176; and F. Fukuyama, *Trust: The Social Virtues and the Creation of Prosperity*, Hamish Hamilton, London, 1995, especially pp.10, 99-104.

⁸ See the discussion of Robert Frank's analysis, Ridley, 1996, op. cit., pp.132-147.

⁹ See *ibid.*, pp.51-66.

¹⁰ See for example, *ibid.*, pp.61-66, 79-84; Dawkins, 1976, op. cit., pp.75-90 (or chapter 12 of the 1989 edition); and D. C. Dennett, *Darwin's Dangerous Idea: Evolution and the Meanings of Life*, Allen Lane The

Penguin Press, London, 1995, pp.252-261.

¹¹ Ridley, op. cit., pp.62-63.

¹² The theme of Hardin's famous essay "The Tragedy of the Commons" is summarized in Hardin, 1993, op. cit., pp.217-218.

¹³ Free-rider situations are described and labeled *chicken games* in I. McLean, *Public Choice: An Introduction*, Basil Blackwell, Oxford, 1987, p.48.

¹⁴ Ridley provides a history of our understanding of the prisoner's dilemma. See Ridley, 1996, op. cit., pp.54-62.

¹⁵ Fukuyama discusses this problem under the heading of "transaction costs"; if you are dealing with strangers whom you do not know and cannot trust, the amount of effort you have to put into striking a secure bargain rises. See Fukuyama, 1995, op. cit., pp.27-28, 151, 200-202, 352.

¹⁶ See Ridley, 1996, op. cit., especially pp.67-84.

¹⁷ See Dawkins, 1976 (1989), op. cit.; and Dennett, 1995, op. cit., pp.253-257, 479-480.

¹⁸ Ridley, following his sources, calls this program Discriminating Altruist. The term *discriminating altruism* has been given specific meaning in Hardin's work. See Hardin, 1993, op. cit., pp.225-236. (Ridley's *Origins of Virtue* is useful and insightful, but he is ungenerous to Hardin. He treats Hardin's work superficially and does not refer to his ground-breaking theory of discriminating versus promiscuous altruism.) For more on discriminating altruism (cooperation) and prisoner's dilemma, see Hardin, 1993, op. cit., and Dennett, 1995, op. cit., pp.479-481.

¹⁹ See Ridley, op. cit., pp.43, 69, 180-183.

²⁰ Hardin, 1993, op. cit., pp.266-267.

²¹ Aristotle, *The Politics* (translated by H. Rackham), William Heinemann: The Loeb Classical Library, London, 1932, Book VII, Chapter IV, pp.557,559.

²² R. Dunbar, *Grooming, Gossip and the Evolution of Language*, Faber and Faber, London, 1996, pp.69-79.

²³ J. Diamond, *Guns, Germs and Steel: The Fates of Human Societies*, W. W. Norton and Company, New York, 1997, pp.267-281. See also F. Fukuyama, *The End of History and the Last Man*, Avon Books, New York, 1992, pp.73-74.

²⁴ *Ibid.*, pp.271-281. Diamond's case that large groups cannot rely on informal organization is strong. In contrast, Ridley argues that states themselves are the cause of wars and we should reduce their authority, relying more on the informal world of local, face to face groups. For all his sophistication he implies that these groups should be borderless and that he aspires to a human world arranged "like a water-color painting, not a mosaic of human populations." See 1996, op. cit., pp.168-169, 264.