PERSPECTIVES: SOCIAL SCIENCE

Shrewd Investments

Martin A. Nowak and Karl Sigmund

There is a well-known story about an elderly academic who made it a point to attend the funerals of colleagues, his reason being that "otherwise, they won't come to mine." This joke reveals a pervasive human trait: Whatever we do, we expect some sort of return. Reciprocation is the basis of human cooperation.

Enhanced online at www.sciencemag.org/cgi/content/full/288/5467/819

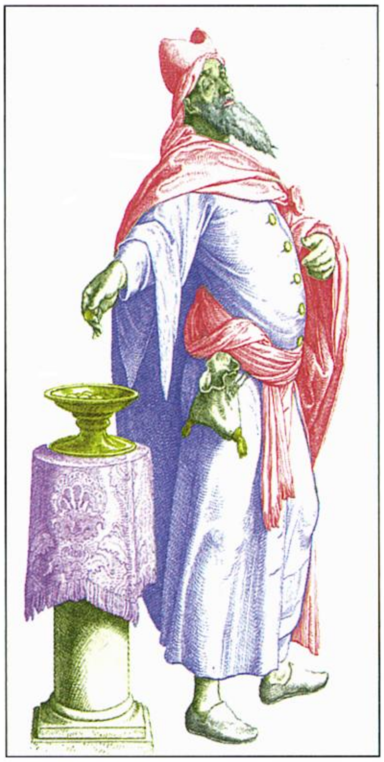
Few altruistic acts are less likely to be returned by the recipient than that of paying last respects. Yet one feels that the old professor has reason to expect that his own funeral will be well attended. He is relying on indirect reciprocity: His act of kindness will be returned not by recipients, but rather by third parties. This type of cooperation figures prominently in human societies, and has been hailed as the "basis of all systems of morality" (*1*). On page 850 of this issue, Wedekind and Milinski examine the part that indirect reciprocity plays in human society. They designed a fascinating psychology experiment to investigate how generosity garners future rewards for donors from third parties, but not from the original recipients (*2*).

An ongoing concern in evolutionary biology is "to take altruism out of altruism," that is, to explain how helping others can emerge in a Darwinian world of "selfish genes." The late W. D. Hamilton (*3*) showed that part of the explanation lies in kin selection (*4*). He proposed that the degree of benefit conferred on an individual directly reflects the kinship of the individual to the donor. Thus, a gene programming an individual to help relatives would benefit those who likely already bear a copy of that gene, and hence would promote its own propagation. Another explanation for why we help others is reciprocal altruism (*5*), which provides an economic, rather than a genetic, underpinning for cooperation. Reciprocal altruism can be divided into direct reciprocity (the receiver returns the favor) and indirect reciprocity (third parties return the favor). With direct reciprocation we expect a return for a favor and feel upset if it is inadequate (although less so with close kin, incidentally).

M. A. Nowak is at the Institute for Advanced Study, Princeton University, Princeton, NJ 08540, USA. E-mail: nowak@ias.edu K. Sigmund is at the Institut fuer Mathematik, Universitaet Wien and International Institute for Applied Systems Analysis, Laxenburg, Austria. E-mail: karl.sigmund@univie.ac.at

CREDIT: MIDNIGHT CALL

With an act of altruism "the shadow of the future"—that is, the expectation of coming gains—has to exceed the cost (*6*). As Trivers points out, "Reciprocal altruism can evolve when two individuals associate long enough to exchange roles frequently as potential altruist and recipient" (*5*). But with indirect reciprocity, this link between donor and recipient is broken. To quote Trivers again, "The donor may receive the return from third parties uninvolved in the initial interaction" (*5*). Such a roundabout way of seeking a return seems threatened by free riders (*7*). With direct reciprocation, a partner that has defected can be immediately punished, with the hope that he or she will learn their lesson and cooperate once more. With indirect reciprocation, where future returns depend on third parties, there seems to be no way to punish free riders. But our obsession with reputation and status, which is widespread in human and animal societies (*8*), points to a reason why indirect reciprocity may work.



Keeping score. The Pharisee makes sure that his generous donations to the Temple are carefully noted to ensure goodwill and rewards in the future, an example of indirect reciprocity.

Wedekind and Milinski (*2*) test how indirect reciprocity operates in a minimalistic experimental setup with 79 first-year Swiss students (unaware of the concept of reciprocal altruism). They designed a game in which the students have the option to donate money to individuals in the group. Each student has a "score" that rises if that individual performs an act of altruism (donating money) and falls if the individual refuses to do so. In such a group, discriminating strategies channel help toward those recipients with high scores, that is, those who have helped others before. Such help is assumed to provide recipients with a benefit—in the experiment, students could donate 1 or 2 Swiss Francs (SFr), but the value of the donation to the recipient was 4 SFr. The number of rounds the game was played was not announced beforehand, and players were unlikely to meet twice. Great care was taken to ensure that the players could not establish links among each other; they knew only the score of the potential recipient, but nothing else. Thus, the students had no way of knowing whether particular co-players had helped them personally in previous rounds.

The results are clear. Helping behavior was widespread, and those who refused to help did so mostly when the co-player had a low score. This shows that even if the chance of individuals meeting twice is negligible, cooperation can be established, provided players have an opportunity to keep tally of their co-players' scores. Of course it is likely that further experiments will reveal additional factors that influence our strategies for indirect reciprocation, such as past experiences, group size, average payoff, and cultural background. But basically, Wedekind and Milinski have shown that indirect reciprocity works as long as there is some way to keep score of giving, and depends on the altruism of third parties.

At first glance, this seems unsurprising. By assisting another individual, you buy a higher score and are more likely to be helped in the future, "a trading of concrete for abstract benefits" (*9*)—the concrete benefit being the payoff and the abstract benefit the score. This makes it look as if, for entirely selfish reasons, you should give whenever you have an opportunity to do so. But here is the twist: Theoretical models (*10*) show that if payoff is counted in terms of Darwinian fitness (that is, number of offspring), then a society of such indiscriminate altruists will quickly be invaded, and indeed subverted, by free riders. A stable level of cooperation can persist only if there is a sufficient number of individuals who are prepared to refuse help to those with a low score. Refusing help lowers the score of such discriminators, however, and diminishes their likelihood of being helped in their turn. Discriminators persist only if the population is challenged sufficiently often by defectors—if not, then society loses its immune response, so to speak.

Hence the stability of indirect reciprocity is based on a rather subtle equilibrium between different strategies. This may be why Wedekind and Milinski did not find significant correlations between payoffs and scores.

Keeping track of scores is important for both direct and indirect reciprocation. Actually, some evolutionary psychologists believe that selection has provided us with a special knack for doing exactly that (*11*). Other "mental modules," like our language instinct or our faculty for recognizing faces, work to the same purpose. The information flow within the social group is all-important; we feel cheated when our good deeds go unnoticed, and refrain from bad deeds lest they become known. The very symbol of moral pressure is the ever-watchful eye in heaven, and conscience acts as an internalization of our standing with others.

More than a hundred years ago, a Viennese playwright identified a root of social injustice in the unfortunate fact that rich people tend to invite for dinner other rich people, rather than the poor. In the eyes of evolutionary biologists, this is direct reciprocation in action, based on the tacit expectation of a return invitation. But capitalists have also always been keen on philanthropy. This entails indirect reciprocity, as shown by the fact that donations are usually well advertised (see the figure), despite the biblical injunction to keep them secret—"The left hand should not know what the right hand is doing."

The fiction of a rational "homo economicus" relentlessly optimizing material utility is giving way to "boundedly rational" decision-makers governed by instincts and emotions (*12*). Economists and biologists are increasingly drawn to the natural history of economic life. Now we have arrived at a stage where formal models can be tested by experiment. Students of animal behavior (*13*), psychologists (*14*), and experimental economists (*15*) are approaching this task from different directions, each with the bias of a long tradition. We must hope that in this convergence of three fields, the left hand will know what the right hand is doing.

References

1. R. D. Alexander, *The Biology of Moral Systems* (De Gruyter, New York, 1987).
2. C. Wedekind and M. Milinski, *Science* **288**, 850 (2000).
3. D. Haig *et al.*, *Science* **287**, 2438 (2000).
4. W. D. Hamilton, *The Narrow Roads of Gene Land* (Oxford University Press, New York, 1996), vol. 1.
5. R. Trivers, *Social Evolution* (Cummings, Menlo Park, CA, 1985).
6. R. Axelrod, *The Evolution of Cooperation* (Basic Books, New York, 1984).
7. E. Fehr and S. Gaechter, *Eur. Econ. Rev.* **42**, 232 (1998); R. Boyd and P. Richerson, *Culture and the Evolutionary Process* (Univ. of Chicago Press, Chicago, IL, 1985).
8. A. Zahavi, A. Zahavi, A. Balaban, *The Handicap Principle* (Oxford Univ. Press, London, 1996).
9. M. Ridley, *The Origin of Virtue* (Penguin, London, 1997).
10. M. A. Nowak and K. Sigmund, *Nature* **393**, 573 (1998).
11. J. H. Barkow, L. Cosmides, J. Tooby, The Adapted Mind (Oxford Univ. Press, London, 1992); see also S. Pinker, *How the Mind Works* (Penguin, London, 1997).
12. R. H. Frank, *Passions Within Reason* (Norton, New York, 1988).
13. L. A. Dugatkin, *Cooperation Among Animals: An Evolutionary Perspective* (Oxford Univ. Press, London, 1997).
14. A. Colman, *Game Theory and Its Applications* (Butterworth, Oxford, 1995).
15. J. H. Kagel and A. E. Roth, Eds., *Handbook of Experimental Economics* (Princeton Univ. Press, Princeton, NJ, 1995).

PERSPECTIVES: PLANETARY SCIENCE

The Shape of Kleopatra

William K. Hartmann

The study of asteroid 216 Kleopatra by Ostro *et al.* on page 836 of this issue (*1*) serves as a reminder of the astonishing results that can come from ground-based observations of asteroids, even in an era when the NEAR spacecraft is orbiting asteroid Eros. Ostro *et al.*'s results are provocative in several regards, from fundamental insights into the composition and structure of Kleopatra to prospects of asteroid mining.

Ostro *et al.* use a unique and powerful technique for learning about interplanetary objects by bouncing radar waves off them. This not only allows a crude imaging of asteroid shapes but also sheds light on asteroid surface properties. When Ostro pioneered asteroid radar imaging in the 1980s and 1990s, the technique was so unusual that he had the field virtually to himself, because few were equipped to follow the path he was exploring (*2*). The popular media likes to give coverage to NASA space missions to asteroids (or, arguably with even greater enthusiasm, to NASA failures). But few nonspecialists will know that Ostro has produced images of oddly shaped asteroids with craters clearly visible on their surfaces by ground-based radar.

An important aspect of the new work reported in this issue involves the understanding of metal asteroids. Since the 1970s, astronomers have grouped asteroids into various "taxonomic classes" that seem to correspond to the types of meteorites that fall on Earth, such as primitive, never-melted rocks, lavalike basaltic rocks, olivine rocks, chunks of nearly pure nickel-iron metal, and mixed stony-iron rocks. But more specific matches between asteroid classes and meteorite types will remain uncertain until samples are returned from asteroids or in situ analyses with landers are performed. So-called M-class asteroids, like Kleopatra, are a case in point. They are thought to match metal-rich meteorites, but metals have few diagnostic spectral absorption features. Ostro *et al.* use a two-step argument to claim that Kleopatra is a giant mass of metal and metal fragments. First, the radar constrains the surface bulk density to be 3.5 g/cm^3, which could match either solid rock or ground-up metal powder with a porosity of less than 60%. Second, the radar reflections show that the surface is not rough, like broken rock, but smooth, like powder lying in repose. The authors thus argue that the surface is covered by metallic powder with a porosity matching that of the powdery regolith on the moon, produced by eons of meteorite sand-

Early conceptual models for the highly elongated asteroid 624 Hektor. Hektor has a similar shape to Kleopatra and may have a similar structure and origin. Hartmann and Cruikshank (*9*) discussed the possibilities that this asteroid formed a single, elongated fragment (**top**), or a compound binary consisting of two strong, spheroidal bodies (**middle**). Weidenschilling (*10*) described how a weak, spinning contact binary could deform into two elongated lobes (**bottom**). The work of Ostro *et al.* suggests this general shape with a narrow neck possibly affected by impact erosion and redistribution of debris.

The author is at the Planetary Science Institute, 620 North Sixth Avenue, Tucson, AZ 85705–8331. E-mail: hartmann@psi.edu