

COMMENTARY

Facing the facts

O. LEIMAR* & P. HAMMERSTEIN†

Department of Zoology, Stockholm University, Stockholm, Sweden†Institute for Theoretical Biology, Humboldt University Berlin, Berlin, Germany*

Counting from the first papers by Hamilton (1963, 1964) on the principles of kin selection, it is now more than four decades that arguments about the evolution of cooperation and altruism have played an important role in evolutionary theorizing. Much of the work is characterized by ingenuity and brilliance, possibly because the evolution of cooperation is crucial for the understanding of biological organization and, in particular, for the understanding of our own species, making it an interesting topic for some of the strongest theoreticians. After this extended period of high-quality conceptual inquiry, it seems reasonable to ask whether theory in the field is becoming more successful at interpreting facts and inspiring decisive empirical observation, which is after all a main purpose of theory in the natural sciences.

From such a perspective, an interesting classification of models ought to make distinctions that encourage critical empirical investigation. The target review by Lehmann & Keller (2006, L & K) achieves this aim in at least one important instance, by insisting on distinguishing whether or not interactions occur between related individuals. However, in other regards their classification seems less helpful in promoting contact between theory and observation.

The dangers of constructing an ivory tower

The total volume of published papers on the evolution of cooperation and altruism is quite staggering, appearing both in the mainstream and in the most prestigious journals. The rate of publication shows no sign of slowing down; if anything, it seems to have increased in recent years. A great proportion of the work consists of computer simulations of highly stylized situations, where a limited number of strategies evolve in some setting. Quite frequently, the work seems to be motivated neither by a particular set of observations nor by the ambition to delineate more precisely which kind of observations would be most important to make. Instead, the field of simulation studies on the evolution of cooperation appears to have taken a life of its own, where the phenomena to be explained are those occurring in the very same world of stylized situations. In a way it is quite

Correspondence: O. Leimar, Department of Zoology, Stockholm University, Stockholm, Sweden.

Tel.: + 46 8 164056; fax: + 46 8 167715;

e-mail: olof.leimar@zoologi.su.se

natural for this to happen, as there is such a wealth of interesting phenomena to explore in that world, but too much focus on the activity can come at the cost of losing track of the relevance of the work to the 'real world' of existing biological organisms. There is also a danger that works actually attempting to provide contact between theory and observation are drowned in an avalanche of computer simulations.

The approach of evolutionary biology

When studying any kind of traits, including those playing a role in cooperation and altruism, the general approach in evolutionary biology is to clarify under what circumstances and in which groups of organisms the traits have evolved (Futuyma, 1998). This is valuable when attempting to interpret traits as adaptations, because it has the potential to tell us something about the selective background. Good theoretical modelling, at least when reaching a somewhat mature state, ought to support the general approach. It can do this, for instance, by delivering robust predictions that qualitatively distinguish the evolutionary consequences of different sets of circumstances (like interactions between relatives or nonrelatives, between mobile or sessile individuals, between cognitively more or less complex individuals, etc.).

A revolution in our understanding of human cooperation

The recent decade has seen a dramatic change in what are generally thought to be promising ideas on the evolution of human cooperation. Innovative observation has been crucial for the change, together with a fruitful interaction between theory and observation. The work by Nowak & Sigmund (1998) on indirect reciprocity through image scoring was influential in catalysing the developments. Nowak & Sigmund proposed that, in a setting of fairly small groups of interacting individuals, possibly characteristic of human evolutionary history, nonkin helping regulated by image scoring could evolve without any directly reciprocating dyads being present in a group, and their suggestion gained support from experiments on modern humans (Wedekind & Milinski, 2000). There was, at the same time, theoretical criticism (Leimar & Hammerstein, 2001) to the effect that the presence of gene flow between groups of interacting individuals, of a magnitude thought to be realistic for historical human populations, essentially would prevent the evolution of image scoring, making it unlikely to have evolved in humans.

In a sort of second wave, a series of groundbreaking experiments by Fehr and coworkers (Fehr & Gächter, 2002; Fehr & Fischbacher, 2003) increasingly challenged the notion that human altruism needs to be based on either direct or indirect reciprocity. Instead, a con-

cept of strong reciprocity (Gintis, 2000), involving a combination of altruism and altruistic punishment of noncooperators, seemed to describe important aspects of human social behaviour. Based on current knowledge, it appears likely that cultural evolution played an important role in the emergence of strong reciprocity (Fehr & Fischbacher, 2003), for instance because this is consistent with results from cross-cultural experimentation (Henrich *et al.*, 2004). With support from a simulation model, Boyd *et al.* (2003) suggested that a form of cultural group selection could promote strong reciprocity.

Although much remains to be explored in terms of modelling and experimental tests of competing explanations, the dual inheritance theory of Boyd & Richerson (1985) appears to be the basic framework needed to understand human cooperation. If so, the framework and classification by L & K seems too limited to encompass the study of human cooperation and altruism, because L & K appear not to take cultural evolution into account. It is of course interesting to compare predictions of models based on some version of dual inheritance theory with those based only on genetical evolution, if only distinguishing predictions were available. Gardner & West (2004) suggested that strong reciprocity in humans could be maintained without an influence of human cultural practices, but they did not offer any critical predictions. Judging from recent cross-cultural comparisons (Henrich *et al.*, 2004), the effect of culture on human behaviour in experimental games is quite noticeable, which is at least suggestive of cultural evolution being important.

Repeated interactions

The papers by Trivers (1971) and Axelrod & Hamilton (1981) on reciprocal altruism were met with great enthusiasm and with a hope that many features of the interactions between organisms would find an explanation in the give and take of direct reciprocation. On the whole, the hopes have gone unfulfilled and it seems important to clarify the reasons for this failure. Although the proposed strategy of Tit-for-Tat can be criticized as perhaps being an unlikely end product of evolution, there are many qualitatively similar cooperative strategies that are more robust, so the problem cannot lie in the very particular implementation of reciprocity investigated by Axelrod & Hamilton (1981). What remains is then either that the structure of real interactions differs in important ways from the original theoretical assumptions, or that the proposed strategies are unlikely to be reached by evolutionary change in real organisms because they correspond to unlikely behavioural mechanisms. Both these possibilities were brought up as important neglected problems by Hammerstein (2003), with an expression of hope that scientists in the field would apply themselves towards a resolution of the seeming discrepancy. This kind of re-evaluation seems to be taking place now (Stevens *et al.*, 2005).

There is of course a fair amount of understanding of the many instances where organisms possess traits that appear designed for providing benefits to others, and where these organisms are relatively independent and interactions are repeated or extend over time. That kind of traits and behaviours are most prevalent in interspecific interactions and the concept of pseudoreciprocity, or investment in by-product mutualism, seems to be the most broadly applicable idea about the phenomenon (Connor, 1995; Leimar & Connor, 2003). Cases where organisms provide benefits mainly for themselves, but where the benefits also can be utilized by others, or where the benefits can be enhanced through behavioural coordination, fall under the heading of by-product mutualism and are relatively well understood. There is also the very large category where organisms become bound to each other in ways that make them succeed or fail together, resulting in common interest and, in some cases, permanent symbiosis (Maynard Smith & Szathmáry, 1995). In the L & K classification all these forms of mutualism would appear as direct benefits, but there seems to be good reason to use several separate categories. If reciprocal altruism, as traditionally defined, is a rare phenomenon outside humans, the L & K classification puts the vast majority of cases of cooperation between unrelated individuals into the direct benefit category, making it too coarse a classification to be of any real use.

Kin selection

Among the strongest points made by L & K is their claim that the importance of kin selection should be recognized for all models where relatives interact. Given the crucial role kin selection has played in evolutionary theory of altruism, this might at first seem like an obvious requirement. Nevertheless, the much studied modelling category of spatial games, where individuals are placed and interact with neighbours on some grid-like structure, and also reproduce locally, was originally presented without a strong focus on kin selection (Nowak & May, 1992; Nowak *et al.*, 1994) and the tradition continues to this day. An explanation for the variation in basic explanatory frameworks could be that it has taken long to properly understand the balance between helping relatives and competing with relatives for reproductive opportunities (e.g. Wilson *et al.*, 1992; West *et al.*, 2002; Doebeli & Hauert, 2005). The work by Rousset (2004) could play a role in establishing a more unified theoretical perspective on the interactions between relatives in spatial settings.

There are many cases in nature where both interaction and reproduction are quite localized in space, but there are also many examples where more-or-less sessile individuals interact with unrelated neighbours, for instance as a result of seed dispersal in plants or planktonic larval dispersal in marine organisms. To

deliver critical predictions, modelling should examine the possibility of a separation of the spatial scales of within-generation interactions and between-generation reproduction. The literature on spatial games can be criticized for underemphasizing this, thus missing an opportunity for greater contact with observation. Even without a deeper understanding of the principles of kin selection, it is possible to vary the range of dispersal in grid- or graph-based models, in order to examine what the role of relatedness might be.

Conclusion

There is of course more to modelling than being an aid to critical observation. One might, for instance, be interested in the methodology of modelling or in the generality of some theoretical principle. Nevertheless, for the evolution of cooperation and altruism, an understanding of traits that have evolved in real organisms seems to be the major reason for the interest in the topic. With all the ingenuity and effort spent on modelling in the field, it would reflect badly on the significance of modelling to evolutionary biology if the gap to observation remains large. An important influence shaping the output of the scientific community lies in the recognition it receives. For the study of cooperation and altruism, we think that the time has come to value work more highly that brings theory and observation into closer contact, compared with work that merely adds another twist to modelling.

References

- Axelrod, R. & Hamilton, W.D. 1981. The evolution of cooperation. *Science* **211**: 1390–1396.
- Boyd, R. & Richerson, P.J. 1985. *Culture and the Evolutionary Process*. The University of Chicago Press, Chicago.
- Boyd, R., Gintis, H., Bowles, S. & Richerson, P. 2003. The evolution of altruistic punishment. *Proc. Natl. Acad. Sci. U S A* **100**: 3531–3535.
- Connor, R.C. 1995. The benefits of mutualism: a conceptual framework. *Biol. Rev.* **70**: 427–457.
- Doebeli, M. & Hauert, C. 2005. Models of cooperation based on the Prisoner's Dilemma and the Snowdrift game. *Ecol. Lett.* **8**: 748–766.
- Fehr, E. & Gächter, S. 2002. Altruistic punishments in humans. *Nature* **415**: 137–140.
- Fehr, E. & Fischbacher, U. 2003. The nature of human altruism. *Nature* **425**: 785–791.
- Futuyma, D.J. 1998. *Evolutionary Biology*, 3rd edn. Sinauer, Sunderland, MA.
- Gardner, A. & West, S.A. 2004. Cooperation and punishment, especially in humans. *Am. Nat.* **164**: 753–764.
- Gintis, H. 2000. Strong reciprocity and human sociality. *J. Theor. Biol.* **206**: 169–179.
- Hamilton, W.D. 1963. The evolution of altruistic behavior. *Am. Nat.* **97**: 354–356.
- Hamilton, W.D. 1964. The genetical evolution of social behaviour, I, II. *J. Theor. Biol.* **7**: 1–52.
- Hammerstein, P. 2003. Why is reciprocity so rare in social animals? A protestant appeal. In: *Genetical and Cultural Evolution of Cooperation* (P. Hammerstein, ed.), pp. 83–93. MIT Press, Cambridge, MA.
- Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E. & Gintis, H. 2004. *Foundations of Human Sociality*. Oxford University Press, Oxford.
- Lehmann, L. & Keller, L. 2006. The evolution of cooperation and altruism. A general framework and a classification of models. *J. Evol. Biol.* **19**: 1365–1376.
- Leimar, O. & Connor, R.C. 2003. By-product benefits, reciprocity, and pseudoreciprocity in mutualism. In: *Genetical and Cultural Evolution of Cooperation* (P. Hammerstein, ed.), pp. 203–222. MIT Press, Cambridge, MA.
- Leimar, O. & Hammerstein, P. 2001. Evolution of cooperation through indirect reciprocity. *Proc. R. Soc. Lond. B* **268**: 745–753.
- Maynard Smith, J. & Szathmáry, E. 1995. *The Major Transitions in Evolution*. Oxford University Press, Oxford.
- Nowak, M. & May, R.M. 1992. Evolutionary games and spatial chaos. *Nature* **359**: 826–829.
- Nowak, M. & Sigmund, K. 1998. Evolution of indirect reciprocity through image scoring. *Nature* **393**: 573–577.
- Nowak, M., Bonhoeffer, S. & May, R.M. 1994. Spatial games and the maintenance of cooperation. *Proc. Natl. Acad. Sci. U S A* **91**: 4877–4881.
- Rousset, F. 2004. *Genetic Structure and Selection in Subdivided Populations*. Princeton University Press, Princeton, NJ.
- Stevens, J.R., Cushman, F.A. & Hauser, M.D. 2005. Evolving the psychological mechanisms for cooperation. *Ann. Rev. Ecol. Syst.* **36**: 499–518.
- Trivers, R.L. 1971. The evolution of reciprocal altruism. *Quart. Rev. Biol.* **46**: 35–57.
- Wedekind, C. & Milinski, M. 2000. Cooperation through image scoring in humans. *Science* **288**: 850–852.
- West, S.A., Pen, I. & Griffin, A.S. 2002. Cooperation and competition between relatives. *Science* **296**: 72–75.
- Wilson, D.S., Pollock, G.B. & Dugatkin, L.A. 1992. Can altruism evolve in purely viscous populations? *Evol. Ecol.* **6**: 31–34.