



International Institute for
Applied Systems Analysis
Schlossplatz 1
A-2361 Laxenburg, Austria

Tel: +43 2236 807 342
Fax: +43 2236 71313
E-mail: publications@iiasa.ac.at
Web: www.iiasa.ac.at

Interim Report

IR-09-090

Sympathy and similarity: The evolutionary dynamics of cooperation

Karl Sigmund (ksigmund@iiasa.ac.at)

Approved by

Ulf Dieckmann
Leader, Evolution and Ecology Program

June 2010

Interim Reports on work of the International Institute for Applied Systems Analysis receive only limited review. Views or opinions expressed herein do not necessarily represent those of the Institute, its National Member Organizations, or other organizations supporting the work.

Sympathy and Similarity: The evolutionary dynamics of cooperation

The advantage of mutual help is threatened by defectors, who exploit the benefits provided by others without providing benefits in return. Cooperation can only be sustained if it is preferentially channelled towards cooperators and away from defectors. But how? A deceptively simple idea is to distinguish cooperators from defectors by tagging them. It clearly is in the interest of cooperators to use some distinctive cue to assort with their like. Such an assortment, however, conflicts with the interests of the cheaters, who have every incentive to also acquire that tag. This makes for an inherently unstable situation. The history of evolutionary thinking on this issue is long. A recent paper by Antal et al. (1) opens new ground by providing an in-depth analysis of a selection-mutation model.

The first to investigate a tag for altruism was W.D. Hamilton (2). He conceived what he called a super-gene, able to produce (a) a distinctive phenotypic trait, (b) the faculty to recognize the trait in others, and (c) the propensity to direct benefits towards bearers of that trait, even though this entails a fitness cost. Soon afterwards, Richard Dawkins described Hamilton's thought experiment by using as phenotypic trait the fanciful example of a green beard. The super-gene was now termed 'green beard gene', in part to acknowledge its inherent unlikelihood. 'Too good to be true', were Dawkins' words (3): for the gene would have to be able to programme for three effects, namely the feature, its recognition and the altruistic propensity.

The green beard concept relates to both major approaches to cooperation in evolutionary biology, namely kin selection (2) and reciprocal altruism (4). It helps in promoting assortment between cooperators; as a result, cooperators can get more than they give, so that altruism becomes a thriving business. Since wearers of green beards both confer and receive benefits, the tag works as a kind of promise that the altruistic action will be returned -- not necessarily by the recipient, but by another member of the green bearded guild. In this sense, the green beard mediates an indirect form of reciprocation, through third parties. In the usual models of indirect reciprocity, 'good guys' are recognised by their reputation, which is based on their past deeds (5). Here, however, recognition is ensured by a phenotypic trait, which is a less sophisticated (and possibly less reliable) signal.

Mostly, the green beard is studied in the context of kin selection. If you carry a green beard, your relatives are likely to carry one too. Directing benefits at green bearded individuals confers the benefits preferentially to your kin, and raises your indirect fitness (since your kin shares your genes with a higher-than-average probability). In many cases, kin are living close by. But the viscosity of the population (to use another term due to Hamilton) is not enough to guarantee a local increase in cooperation, since it is counterbalanced by a local increase in competition. Limited dispersal alone is therefore not enough. A gene for kin recognition can help to direct positive rather than negative effects towards relatives. But it is important to realise that the green beard can promote altruism beyond the realm of the family.

Some ten years ago, it was found that green beards are not as implausible as their name suggests. In particular, Haig suggested that genes for homophilic cell adhesion could perform all three tasks required from a green beard gene – trait, recognition, and action – by coding for a surface protein that allows them to stick to copies of themselves on other cells (6). A few years later, it was found that *csA* genes in *Dictyostelium discoideum* fit the bill (7). In hard times, these amoeba literally stick together to form stalks for dispersing their spores. A similar

gene has also been discovered in flocculating yeast cells (8). Other candidates for more sophisticated green beard effects have been found in ants and lizards.

An obvious way to cheat is to grow a green beard but skip the altruism. For homophilic cell adhesion, this seems barely feasible. In other examples, cheating may be prevented by genetic constraints. But in principle, one would expect that a tight link between a gene for altruistic behaviour and a gene for tag recognition will ultimately be broken, and cooperation be destroyed. Surprisingly, it turned out that if the link is not too tight (but not too loose either), a dynamic regime of cooperation can emerge, based on tag diversity. Whenever some tag becomes too frequent, it can be faked by defectors, but cooperative behaviour subsists nevertheless, by allying itself with another tag. This phenomenon has been termed 'beard chromodynamics', to suggest that green beards can over time be replaced by red, or blue, or yellow beards as rallying signals for cooperators (9,10).

The underlying principle is that of a shibboleth, or secret handshake. But such a specially contrived trait, evolved for the purpose of signalling cooperation, is not always necessary. Tag-based cooperation can also rely on self-similarity. All that is needed is some general means to recognise what is like yourself and what is not, i.e., to distinguish 'us' from 'them'. With familiars, you need no badge, or password. This has been called 'armpit effect' (by Dawkins again, see (3)). While an obvious variation of the green beard principle could mediate, in principle, symbiosis between two different species, the armpit effect is self-referential. You need not sprout a special recognition device but simply check whether the other looks, or smells, or sounds like you.

Mechanisms based on self-similarity are commonly used among cells of an organism, or among members of a species. Kin-recognition seems widespread: it is useful, not only for promoting nepotism, but also for avoiding incest (11). Bats or birds recognize their offspring on crowded cave-roofs and cliff-faces through vocalizations, hamsters and wasps pick up the odour of their nest or colony, etc. Interestingly, these faculties seem always acquired through imprinting, rather than genetically encoded. Thus they indicate in-group rather than kin. This use of associative learning is well supported by theory (12).

An armpit effect has been recently found in hamsters (13). Self-similarity appears to work in humans too: we like our like. Neat economic experiments show that players preferentially trust similar-looking co-players (14). (The players are provided with pictures of their ostensible partners, and these photos are manipulated to look to a greater or lesser degree like themselves). Clearly, such cues for self-similarity can be enhanced by cultural means. Many groups provide their members with characteristic uniforms, badges, tattoos, ties, hair-cuts, hang-outs, accents, musical tastes, or slang idioms.

In most tag-based models, the tags are discrete – you either look like me or you don't. In general, defectors can be overcome only for a restricted range of recombination between tag and behaviour, cf. (15,16,17). However, similarity is likely to be a question of degree – you can look more or less like me. In the case of continuous graduation, it is likely that cooperative behaviour is addressed towards all those who are tolerably similar. Such models show intriguing patterns: cliques of similar cooperators grow, are beset and undermined by defectors, and re-group around other phenotypes (18,19). Extending tolerance to a larger range of tag-values enlarges the basis of collaboration, whereas restricting tolerance shields from exploiters: this leads to endlessly fluctuating 'tides of tolerance' (20).

In the model of Antal et al (1), members of a well-mixed population of constant size N are distinguished by a tag which can take infinitely many values and is coded by integers. Defectors help nobody, and cooperators provide help exclusively to members of their own tag-group. From time to time, individuals produce offspring in numbers proportional to their fitness. Some N of these offspring are randomly chosen to form the next generation. Offspring inherit from their parent both their behaviour (cooperator or defector) and their tag, up to mutation. Each configuration of the population is specified by the number of defectors and cooperators for each tag. The expected payoff values for defectors and cooperators can easily be computed in terms of conditional probabilities (e.g., for defectors to interact with cooperators, etc.). This specifies the configurations for which cooperators are sufficiently assorted with other cooperators to earn more than defectors do. But the configurations move and cluster in a very fluid manner through the range of possible tags. It needs considerable mathematical dexterity to average the payoffs over all configurations, in the stationary state. This yields, under the limiting assumption of weak selection, a condition for cooperators to be more frequent than defectors in the long-term, requiring that the benefit-to-cost ratio exceeds a specific threshold. Under the most favourable conditions, i.e., when mutations between tags are frequent and mutations in the behaviour rare, that threshold is slightly larger than 2. In contrast to previous models (9,15,16), no additional requirements on spatial population distribution are used. The analysis of several limiting cases shows that the results depend significantly on mutation structure, about which empirical data is lacking at present. The elusive nature of the game of hide-and-seek between cooperators and defectors, an age-long spur for biological and cultural evolution, continues to challenge experimentalists and theoreticians alike.

- (1) Antal T et al. (2009) Evolution of cooperation by phenotypic similarity, Proc. Natl. Acad. Sci. USA, this issue.
- (2) Hamilton WD (1964) The genetical basis of social behaviour I, J. Theor. Biol. 7, 1-16.
- (3) Dawkins R (1982) The extended phenotype, Oxford University Press.
- (4) Trivers R (1971) The evolution of reciprocal altruism, Quarterly Rev. Biol. 46, 35-57.
- (5) Nowak MA, Sigmund K (2005) Evolution of indirect reciprocity, Nature 437, 1292-1298.
- (6) Haig D (1996) Gestational drive and the green-beard placenta, PNAS 93, 6547-51.
- (7) Queller D C et al. (2003) Single gene green beard effects in the social amoeba *Dictyostelium discoideum*, Science 299, 105-106.
- (8) Smukalla S et al. (2006) Flo1 is a variable green beard gene that drives biofilm-like cooperation in budding yeast, Cell 135, 726-737.
- (9) Jansen VAA, van Baalen M (2006) Altruism through beard chromodynamics, Nature 440, 663-664.
- (10) Traulsen A, Nowak MA (2007) Chromodynamics of cooperation in finite populations, PLoS ONE 2(3) e270.

- (11) Pfennig DW (2002) Kin recognition, in M. Pagel (ed) *Encyclopedia of Evolution*, Oxford University Press.
- (12) Lehmann L, Perrin N (2002) Altruism, dispersal, and phenotype-matching kin recognition, *Amer. Naturalist* 159, 451-467.
- (13) Mateo JM, Johnston RE (2000) Kin recognition and the ‘armpit effect’: evidence of self-referent phenotype matching, *Proc. Royal Society London B*, 267, 695-700.
- (14) Krupp DB, DeBruine LM, Barclay P (2008) A cue of kinship promotes cooperation for the public good, *Evolution and Human Behaviour* 29, 49-55.
- (15) Rousset R, Roze, D (2007) Constraints on the origin and maintenance of genetic kin recognition, *Evolution*, 61, 2320-2330.
- (16) Axelrod R, Hammond RA, Grafen A (2004) Altruism via kin-selection strategies that rely on arbitrary tags with which they co-evolve, *Evolution* 58, 1833-38.
- (17) Gardner A, West SA (2007) Social evolution: the decline and fall of genetic kin recognition, *Current Biology* 17, R810-812.
- (18) Riolo RL, Cohen MD, Axelrod R (2001) Evolution of cooperation without reciprocity, *Nature* 414, 441-443.
- (19) Traulsen A, Schuster HG (2003) Minimal model for tag-based cooperation, *Phys. Rev. E* 68, 046129.
- (20) Sigmund K, Nowak MA (2001) Tides of tolerance, *Nature* 414, 403-404.

Fig: Face transitions. Players in a game theoretic experiment are provided with pictures of their partners who, through digital sorcery, are made to look like themselves, to a greater or lesser extent. Here, the face in the middle is the result of a 60:40 mix of the other two faces. Players preferentially trust co-players who look more like themselves. Thus familiarity enhances trust. With kind permission from Lisa DeBruine, see (14).

Karl Sigmund
Faculty of Mathematics, University of Vienna,
and International Institute for Applied Systems Analysis (IIASA), Laxenburg, Austria
Email: karl.sigmund@univie.ac.at