MINI REVIEW

Social semantics: altruism, cooperation, mutualism, strong reciprocity and group selection

S. A. WEST,* A. S. GRIFFIN* & A. GARDNER*†

*Institute of Evolutionary Biology, School of Biological Sciences, University of Edinburgh, King's Buildings, Edinburgh, UK
†Departments of Biology and Mathematics & Statistics, Queen's University, Kingston, ON, Canada

Abstract

From an evolutionary perspective, social behaviours are those which have fitness consequences for both the individual that performs the behaviour, and another individual. Over the last 43 years, a huge theoretical and empirical literature has developed on this topic. However, progress is often hindered by poor communication between scientists, with different people using the same term to mean different things, or different terms to mean the same thing. This can obscure what is biologically important, and what is not. The potential for such semantic confusion is greatest with interdisciplinary research. Our aim here is to address issues of semantic confusion that have arisen with research on the problem of cooperation. In particular, we: (i) discuss confusion over the terms kin selection, mutualism, mutual benefit, cooperation, altruism, reciprocal altruism, weak altruism, altruistic punishment, strong reciprocity, group selection and direct fitness; (ii) emphasize the need to distinguish between proximate (mechanism) and ultimate (survival value) explanations of behaviours. We draw examples from all areas, but especially recent work on humans and microbes.

The difference of approach is purely didactic – there is no disagreement about matters of substance Grafen (1984, p. 82)

Introduction

Scientific progress depends upon precise, reliable communication between scientists (Brown, 1983). This can be hindered if people use the same term to mean different things, or different terms to mean the same thing. In the extreme, this can lead to debates or disputes when in fact there is no disagreement, or the illusion of agreement when there is disagreement. Here, we are concerned with issues of semantic confusion that have arisen with research on the problem of cooperation. The problem is why should an individual carry out a cooperative behaviour that appears costly to perform, but benefits other individuals (Hamilton, 1963, 1964)? A

Correspondence: Institute of Evolutionary Biology, School of Biological Sciences, University of Edinburgh, King's Buildings, Edinburgh, EH9 3JT, UK.
Tel.: 0131 6505496; fax: 0131 6506564; e-mail: stu.west@ed.ac.uk

large body of theory and a rich experimental literature have arisen on this problem in the wake of Hamilton's pioneering papers.

To a large extent, semantic confusion should not be a problem in this field. Several evolutionary theoreticians have developed comprehensive overviews of the area, and there is relatively general agreement between them (Grafen, 1985; Queller, 1985; Taylor, 1996; Frank, 1998, 2003; Rousset, 2004; Sachs *et al.*, 2004; Foster & Wenseleers, 2006; Gardner *et al.*, 2007; Grafen, 2006; Lehmann & Keller, 2006). Furthermore, several possible points of confusion have long been solved (Grafen, 1984; Frank, 1998; Reeve & Keller, 1999; Rousset, 2004). However, despite this, semantic problems are still managing to creep in from a number of directions.

In particular, several factors have led to us writing this paper. First, extensions of social evolution theory to new taxa seem to allow old semantic problems to reoccur or new variants to arise. Two areas where this is currently important, that we shall discuss, are work on cooperation in humans and microbes. Other recent examples include

Box. 1 Glossary.

Actor: focal individual who performs a behaviour.

Altruism: a behaviour which is costly to the actor and beneficial to the recipient; in this case and below, cost and benefit are defined on the basis of the lifetime direct fitness consequences of a behaviour.

Cheaters: individuals who do not cooperate (or cooperate less than their fair share), but are potentially able to gain the benefit of others cooperating.

Cooperation: a behaviour which provides a benefit to another individual (recipient), and which is selected for because of its beneficial effect on the recipient.

Direct fitness: the component of fitness gained through the impact of an individual's behaviour on the production of offspring.

Inclusive fitness: 'the effect of one individual's actions on everybody's numbers of offspring … weighted by the relatedness' (Grafen, 1984); the sum of direct and indirect fitness; the quantity maximized by Darwinian individuals.

Indirect fitness: the component of fitness gained from aiding the reproduction of related individuals.

Kin selection: process by which traits are favoured because of their beneficial effects on the fitness of relatives.

Local group: a subset of the population who are interacting; the local group may vary from the perspective of different behaviours or traits.

Mutual benefit: a behaviour which is beneficial to both the actor and the recipient.

Mutualism: cooperation between species.

Neighbour-modulated fitness: total personal fitness, including the effects of one's own behaviour and the behaviours of social partners.

Recipient: an individual who is affected by the behaviour of the focal individual.

Relatedness: a measure of genetic similarity.

Selfishness: a behaviour which is beneficial to the actor and costly to the recipient.

Spite: a behaviour which is costly to both the actor and the recipient.

---

agronomy (reviewed by Denison *et al.*, 2003) and parasitology (reviewed by Read, 1994; West *et al.*, 2001, 2003). Secondly, serious semantic misunderstandings can still occur in the evolutionary literature (e.g. Zahavi, 2003, 2005; Wilson & Hölldobler, 2005; Wilson, 2005). Thirdly, even the leading social evolution theoreticians do not use fundamental terminology consistently, as illustrated by a recent target review on the topic (Lehmann & Keller, 2006) and its associated commentaries, or a recent edited volume (Hammerstein, 2003).

Our aim here was to discuss some major terms that are often misused or misunderstood (Box 1). Specifically: (i) kin selection; (ii) mutualism; (iii) cooperation; (iv) altruism; (v) group selection; (vi) direct fitness (vii) distinguishing between ultimate and proximate explanations of behaviour. We then end with the caveat that: (viii) classifying behaviours will not always be the easiest or most useful thing to do.

Social evolution theory

Before discussing possible points of confusion, it is useful to provide a basic summary of relevant theory.

As stated above, the problem of cooperation is why should an individual carry out a cooperative behaviour that appears costly to perform, but benefits other individuals (Hamilton, 1963, 1964)? Theoretical explanations for the evolution of cooperation (or any behaviour) can be broadly classified into two categories: direct fitness benefits or indirect fitness benefits (Fig. 1; Hamilton, 1964; Brown & Brown, 1981; Grafen, 1984; Taylor, 1996; Lehmann & Keller, 2006; West *et al.*, 2006b). This follows from Hamilton's insight that individuals gain inclusive fitness through their impact on the reproduction of related individuals (indirect fitness effects) as well as directly through their impact on their own reproduction (direct fitness effects) (Hamilton, 1964; Grafen, 1984). The terms direct and indirect fitness were introduced by Brown & Brown (1981), although Fisher (1930, chapter 2) discussed indirect effects in a similar context.

The first class of explanations for cooperation is that it may provide a direct fitness benefit to the individual that performs the behaviour, which outweighs the cost of performing the behaviour (Sachs *et al.*, 2004). One possibility is that individuals have a shared interest in cooperation. For example, in many cooperative breeding species, larger group size may provide a benefit to all the members of the group through factors such as greater survival or higher foraging success – in this case, individuals can be selected to help rear offspring that are not their own, in order to increase group size (Kokko *et al.*, 2001). Another possibility is that there is some mechanism for enforcing cooperation, by rewarding cooperators or punishing cheaters (Trivers, 1971; Frank, 2003). This could happen in a variety of ways, which have been termed punishment, policing, sanctions, reciprocal altruism, indirect (reputation based) reciprocity and strong reciprocity (see below).

The second class of explanations for cooperation is that it provides an indirect benefit because it is directed towards other individuals who carry the cooperative gene (Hamilton, 1964, 1970, 1975). The easiest and most common way in which this could occur is if genes are identical by descent – by helping a close relative reproduce, an individual is still passing on its own genes to the next generation, albeit indirectly. Hamilton (1964) pointed out that this could occur via two mechanisms: (i) kin discrimination, when cooperation is preferentially directed towards relatives; (ii) limited dispersal (population viscosity) keeping relatives together, allowing cooperation to be directed indiscriminately towards all neighbours (this will be favoured as those neighbours tend to be relatives). The second way to obtain an indirect fitness benefit is if cooperation is directed towards nonrelatives who share the same cooperative gene. This assortment or 'greenbeard' mechanism requires a single gene (or a number of tightly linked genes) that both causes the cooperative behaviour and can be recognized by other individuals
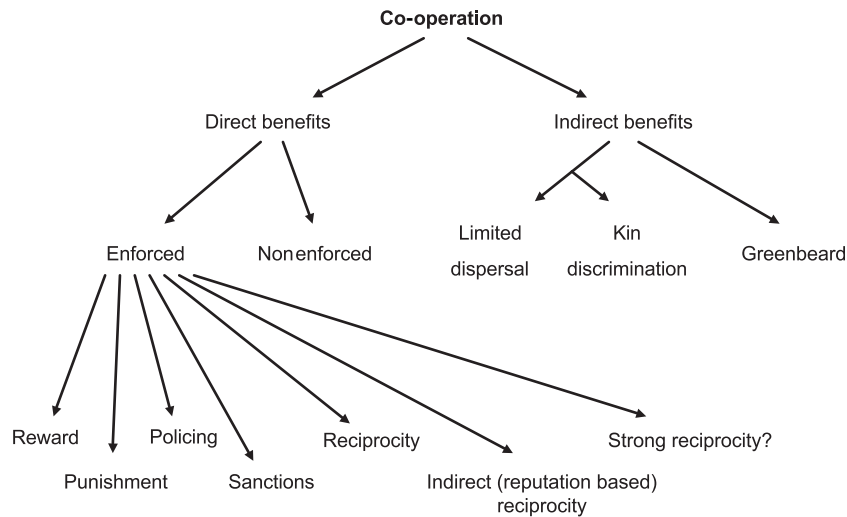
Fig. 1 A classification of the explanations for cooperation. Direct benefits explain mutually beneficial cooperation, whereas indirect benefits explain altruistic cooperation. Within these two fundamental categories, the different mechanisms can be classified in various ways – here we follow West *et al.* (2006; see also Sachs *et al.*, 2004; Lehmann & Keller, 2006). These possibilities are not mutually exclusive, for example a single act of cooperation could have both direct and indirect fitness benefits. We have listed some of the many different terms that have been used to describe the mechanisms for enforcing cooperation to emphasize that reciprocity (reciprocal altruism) is only one of many ways to obtain direct fitness benefits through cooperation. These enforcement mechanisms can also alter the indirect benefits of a behaviour (Lehmann & Keller, 2006), and determining the relationships between these terms remains an important task. Kin selection has been used to refer to (i) just those indirect benefits involving coancestry (i.e. limited dispersal and kin discrimination), or (ii) all indirect benefits (i.e. also including greenbeard effects).

due to a distinctive phenotypic marker, such as a green beard (Hamilton, 1964, 1975; Dawkins, 1976; Jansen & van Baalen, 2006). An alternative is to conceptualize greenbeards as a form of kin discrimination.

Kin selection

Maynard Smith (1964) coined the term 'kin selection' to describe how indirect fitness benefits arise from helping relatives reproduce. Since then, the phrase kin selection has been used in two different ways. The narrower use of kin selection works upon interactions between individuals who are genetically related due to common ancestry – i.e. indirect benefits due to limited dispersal or kin discrimination. The broader use of kin selection works upon interactions between individuals who share the gene of interest, regardless of whether this is due to coancestry or some other mechanism – i.e. also includes greenbeard effects. The difference between these usages is therefore whether kinship and relatedness are defined on the basis of average genetic similarity over most of the genome (narrow definition), or at the particular locus of the behaviour being examined (broad definition). Hamilton (1975) favoured the narrower definition of kin selection, arguing that inclusive fitness (and Hamilton's rule) was more general and should be distinguished from kinship effects. However, most people have used the broader definition,

using kin selection wherever genetic relatedness between social partners occurs, irrespective of the causes of relatedness. Here, we shall follow the broader definition, whilst also noting that this distinction is usually not important because kinship is by far the most common reason for indirect fitness benefits.

It is also useful to make the distinction between process and maximand. Natural selection is the process by which fitness is maximized. Inclusive fitness is a form of analysis of social traits, and a generalization of Darwinian fitness that takes account of social interactions – natural selection leads to organisms acting as if they are maximizing their inclusive fitness (Grafen, 2006). Hamilton (1964) did not provide a specific term for the process by which inclusive fitness was maximized – his aim was to show that natural selection would maximize inclusive fitness. Consequently, the term kin selection appears to have been adopted for this role, of what provides a generalized description of natural selection on social interactions. The key point here is that kin selection is the process by which inclusive fitness is maximized only if we are using kin selection in its broadest sense. In some of his later papers, after the term kin selection had been adopted elsewhere, Hamilton (1971, 1972) seems to refer to social selection as the process by which inclusive fitness is maximized, and distinguishes between classical theory and social theory (see also Frank, 2006).

Mutualism and mutual benefit

From an evolutionary point of view, a behaviour is social if it has fitness consequences for both the individual that performs that behaviour (the actor) and another individual (the recipient). Hamilton (1964) classified social behaviours according to whether the consequences they entail for the actor and recipient are beneficial (increase direct fitness) or costly (decrease direct fitness) (Table 1). Whether a behaviour is beneficial or costly is defined on the basis of: (i) the lifetime consequences of the behaviour (i.e. not just the short-term consequences) and (ii) the absolute fitness effect – for example, does it increase or decrease the actor's number of offspring surviving to adulthood (i.e. not just relative to the individuals or social group with which the actor directly interacts). In his original papers, Hamilton (1964) provided terms for two of the four possibilities. He termed a behaviour which is beneficial to the actor and costly to the recipient (+/−) as selfishness, and a behaviour which is costly to the actor and beneficial to the recipient (−/+) as altruism. Later, he termed a behaviour which is costly to both the actor and the recipient (−/−) as spite (Hamilton, 1970). We do not discuss the semantics of the term spite in this paper, because we have done so in detail elsewhere (A. Gardner, I.C.W. Hardy, P.D. Taylor, S.A. West, unpublished data). In this section, we are concerned with defining a behaviour which is beneficial to both the actor and the recipient (+/+)?

One possible term for +/+ behaviours is 'cooperation' (Trivers, 1985; Bourke & Franks, 1995; Rousset, 2004; Lehmann & Keller, 2006). However, this use of cooperation implies that cooperation is always explained by direct fitness benefits. This can lead to confusion because others use cooperation more generally for behaviours that are beneficial to the recipient, but can be either beneficial (+/+) or costly (−/+) to the actor (Hamilton, 1972; Axelrod & Hamilton, 1981; Frank, 1995b; Maynard Smith & Szathmary, 1995; Sachs *et al.*, 2004; Foster *et al.*, 2006; West *et al.*, 2006b). Furthermore, to use cooperation for only +/+ behaviours, and not −/+ behaviours (indirect benefits) contradicts the popular use of cooperation in the empirical literature. For example, referring to cooperative breeding in vertebrates or insects does not necessarily mean that direct fitness benefits are the explanation (Rousset,

2004). Indeed, a major question in studies of cooperative breeding species is to determine the extent to which cooperation is explained by direct (+/+) or indirect (−/+) fitness benefits (Clutton-Brock, 2002; Griffin & West, 2002). 'Helping' is another possible term for +/+ behaviours, but it suffers from the same problems as cooperation (Rousset, 2004). We return to the specific definition of cooperation below.

Another possible term for +/+ behaviours is 'mutualism' (Krebs & Davies, 1993; Emlen, 1997; Foster *et al.*, 2001; Clutton-Brock, 2002; Gardner & West, 2004a; Ratnieks, 2006). There are also examples of people using cooperation and mutualism interchangeably (Brown, 1983; Alcock, 1989). However, this use of mutualism can be confusing because many also use mutualism to refer to the more specific case of cooperation between species (Wilson, 1975b; Herre *et al.*, 1999; Yu, 2001; West *et al.*, 2002b; Kiers *et al.*, 2003; Ridley, 2004; Futuyma, 2005; Foster & Wenseleers, 2006). This is also the common use of mutualism in the ecological literature. The two ideas here are quite distinct, as mutually beneficial social behaviour is a description of the effect of a single behaviour on the actor and the recipient, whereas interspecific mutualism describes the impact that each party has on the other. Using the same word to indicate both may cause confusion, for example it is easy to see how a mutually beneficial behaviour can evolve, but this does not mean that interspecific mutualisms are so easily explained. We think that the term mutualism is best reserved for its ecological usage.

In order to solve this problem we suggest the term 'mutual benefit' to describe +/+ behaviours. This term states as simply as possible, and without extra connotations, that a behaviour is beneficial to both the actor and the recipient. We emphasize that we think redefining terms is usually counter-productive, and can add more confusion than it solves. However, we think it is useful in this case, because we already have the situation where multiple terms are being used to define the same thing, and those terms also have other uses. We have found two similar uses of mutual benefit. First, Maynard Smith (1982, p. 167) appears to use mutual benefit in the same way as us, commenting that cooperation that can be explained by either interactions between kin or mutual benefits to cooperating individuals. Secondly, Maynard Smith & Szathmary (1995, p. 261) use it in a similar, but not identical way. They use mutual benefit to describe 'synergistic' benefits to cooperation, and distinguish it from when individuals force others to cooperate, which they term enforcement. This classification can be confusing because: (i) it is not clear how they classify scenarios when cooperation evolves through punishment or policing, which is enforced, but leads a mutual benefit in the long term; (ii) they seem to classify some cases which rely on kin selection into their mutual benefit category (e.g. Nowak & May, 1992).

Table 1 Social Behaviours. A Hamiltonian classification scheme for behaviours that have been selected for by natural selection.

	Effect on recipient	
	+	−
Effect on actor		
+	Mutual Benefit	Selfishness
−	Altruism	Spite

Cooperation

We now consider the formal definition for 'cooperation'. It is extremely useful to have a term that describes 'cooperative behaviours' that provide a benefit to the recipient, but could be beneficial (+/+, favoured by direct fitness benefit) and/or costly (−/+, favoured by indirect fitness benefit) to the actor. Cooperation and helping are both frequently used in this context. For example, as discussed above, a topic of debate is the extent to which cooperative breeding, or helping behaviours in cooperative breeders, can be explained by direct or indirect fitness benefits. Consequently, one possibility is to define 'cooperation' as a behaviour that provides a benefit to the recipient – this therefore includes both +/+ and −/+ behaviours (Sachs *et al.*, 2004).

However, this definition of cooperation may be overly inclusive. For example, when an elephant produces dung, this is beneficial to the elephant (emptying waste), but also beneficial to a dung beetle that comes along and uses that dung. It does not seem useful to term behaviours such as this, which provide a one-way byproduct benefit, as cooperation. Consequently, we prefer that a behaviour is only classed as cooperation if that behaviour is selected for because of its beneficial effect on the recipient. We do not wish to imply that the behaviour is selected for purely because of its beneficial effect on the recipient, just that it has at least partially done so. This is easily illustrated with an example. Suppose that two bacterial species (A and B) are interacting, and that each feeds upon a waste product of the other. This would be a mutually beneficial (+/+) behaviour, but we would not class it as cooperation. However, now suppose that a higher production of species A's waste product evolved because this benefited species B, and hence led to a higher level of waste production by species B, which was beneficial to species A. This kind of interaction has been termed byproduct reciprocity or invested benefits (Conner, 1995a; Sachs *et al.*, 2004), and we would class this as cooperation. Our definition of cooperation therefore includes all altruistic (−/+) and some mutually beneficial (+/+) behaviours.

This distinction over whether a trait is favoured for that purpose relates to the standard text book definition of adaptation (Rose & Lauder, 1996). A common definition for an adaptation is a 'trait that enhances fitness and that arose historically as a result of natural selection for its current role' (Rose & Lauder, 1996). A difference here is that our definition of cooperation considers the selective forces maintaining the trait, and not just those that led to its initial evolution – this is sometimes termed 'aptation', as opposed to adaptation (Rose & Lauder, 1996). Although we suspect this distinction will be unimportant for most real cases, there are fields in which it is important, such as distinguishing between the evolution and maintenance of sex and recombination (West *et al.*, 1999; Burt, 2000). That

historical factors can be important for social evolution is demonstrated by the example that it can be hard for eusociality to evolve in species with multiple mating, due to reduced relatedness, but that multiple mating can evolve in species which are already eusocial (Hamilton, 1964; Boomsma & Ratnieks, 1996). We admit that determining whether certain +/+ behaviours are cooperation may be hard, but this emphasizes that it is a key question.

The use of distinguishing social behaviours on the basis of the selective forces maintaining them has been demonstrated by the analogous distinction made in the communication literature between a 'signal' and a 'cue' (Maynard Smith & Harper, 2003). A signal is 'any act or structure which alters the behaviour of other organisms, which evolved because of that effect, and which is effective because the receiver's response has also evolved', whereas a cue is 'a feature of the world, animate or inanimate, that can be used by an animal as a guide to future action' (Maynard Smith & Harper, 2003). This distinction has been extremely useful, allowing a clear and general conceptual overview to be developed. Although, we note that this definition of a signal does not exclude actions that operate because of their substantive effects rather than their information content, so for example, it could include reciprocity, where cooperation is conditional upon the cooperative behaviour of others (A. Grafen, Personal Communication; see also Grafen, 1990).

Altruism

The above sections have emphasized how terms such as altruism have very specific meanings, which can convey useful information. Consequently, when these terms are misused, or redefined, it can lead to confusion. In this section we discuss three cases in which this has occurred with the term altruism. A general point here is that altruism is defined: (i) with respect to the lifetime consequences of a behaviour; (ii) on absolute fitness effects (i.e. does it increase or decrease the actor's fitness, and not relative to just some subset of the population). For example, if a cooperative behaviour was costly in the short term, but provided some long-term (future) benefit, which outweighed that, it would be mutually beneficial and not altruistic. This does not mean that such a behaviour is somehow less interesting – determining whether and how a cooperative behaviour provides short or long term direct fitness benefits remains a major problem (Clutton-Brock, 2002; Griffin & West, 2002). Instead, our aim is to distinguish the fundamental difference between direct and indirect fitness benefits.

Reciprocal altruism

Trivers (1971) suggested that cooperation could be favoured between nonrelatives, in reciprocal interactions.

The idea here is that individuals can take turns in helping each other, for example by preferentially aiding others who have helped them in the past. Trivers termed this 'reciprocal altruism'. This work was highly influential in showing that cooperation could be favoured between nonrelatives, and stimulated a huge amount of theoretical and empirical research. Current theoretical overviews place reciprocal altruism as just one of many mechanisms by which cooperation between unrelated individuals can be favoured (Krebs & Davies, 1993; Griffin & West, 2002; Frank, 2003; Sachs *et al.*, 2004; Lehmann & Keller, 2006; West *et al.*, 2006b). Furthermore, it has been suggested that the specific mechanism of reciprocity is unlikely to be of general importance outside of humans, because the conditions required can be extremely restrictive (Conner, 1995b; Dugatkin, 1997; Clutton-Brock, 2002; Hammerstein, 2003; Stevens & Hauser, 2004; Stevens *et al.*, 2005).

However, reciprocal altruism is not altruistic – it provides a direct fitness advantage to cooperating. If an individual does not pay the cost of cooperation in the short term then it will not gain the benefit of cooperation in the long term (although things could get more complicated if reciprocity was between relatives). Consequently, following Hamilton's original scheme, it is a mutually beneficial (+/+) behaviour and not an altruistic behaviour (−/+). It is presumably for this reason that Hamilton (1996, p. 263) thought that reciprocal altruism was misnamed, and that he and others have used alternative terms such as 'reciprocity' (Alexander, 1974), or 'reciprocal cooperation' (Axelrod & Hamilton, 1981). Unfortunately, the term reciprocal altruism has been in use so long that we do not expect its use be changed, although it would be preferable to use reciprocity or reciprocal cooperation. It is not clear how much confusion this different use of altruism has led to – although we suspect it is at least partially responsible for: (i) the frequent and incorrect assumption that kin selection and reciprocal altruism are the two leading explanations for cooperation or altruism and (ii) the confusing use of altruism in the human literature (see below).

Trivers (1971) originally redefined altruism in a different way to Hamilton. Specifically, he defined it with respect to inclusive fitness, apparently in the short term, rather than direct fitness in the long term: 'Altruistic behavior can be defined as a behavior that benefits another organism, not closely related, while being apparently detrimental to the organism performing the behavior, benefit and detriment being defined in terms of contribution to inclusive fitness' (Trivers, 1971). It is clear that this form of 'altruism' could not be selected for, unless benefit and cost are only measured in the short term, and that there is some longer-term benefit (i.e. it is mutually beneficial and not altruistic). In his later book, Trivers (1985) returns to Hamilton's definitions associated with Table 1 ('the altruism becomes a

form of cooperation' – having previously defined cooperation as a +/+ behaviour and altruism as a −/+ behaviour), and at times uses phrases which do not invoke altruism, such as 'return effects', 'return-benefit' and 'reciprocity' (p. 47). However, altruism is still used in other ways at other times, in a manner that is based on short term rather than long term cost and benefits. For example: 'In effect, two individuals trade altruistic acts. This can be called reciprocity or reciprocal altruism' (p. 48), and by referring to kinship and reciprocity as two ways to explain altruism (p. 49).

Weak altruism

The different use of the word altruism in the 'new' group selection literature has also led to confusion. Wilson and colleagues (Wilson, 1975a, 1977; Colwell, 1981) redefined altruism to refer to the fitness of an individual relative to the individuals that it interacts with, in its group. They define a behaviour as 'weakly altruistic' if it leads to a decrease in the fitness of the focal individual, relative to the other members of its group. This means that behaviours which provide a benefit to everyone within the local group, including the actor, such as the production of a public good, can be defined as altruistic. These are sometimes termed 'whole-group' (Pepper, 2000) or 'group beneficial' traits (Dugatkin *et al.*, 2003, 2005) – as opposed to 'other-only' traits which do not provide a benefit to the actor (Pepper, 2000). Whole-group behaviours can have both a direct and indirect benefit, and as discussed above, this means that whether they are altruistic or mutually beneficial will depend upon the relative cost and benefits of the behaviour, as well as population structure (Pepper, 2000; Rousset, 2004). This leads to the confusing situation where a trait could be favoured because it selfishly increases an individual's direct fitness, but will be weakly altruistic by Wilson's definition (Dawkins, 1979; Grafen, 1984; Harvey *et al.*, 1985)! This also emphasizes that there is a fundamental difference between altruism (−/+) and weak altruism (which can be +/+ or −/+), in contrast to the suggestion of Fletcher & Doebeli (2006).

We illustrate this point in Table 2, with the simplest possible case, where groups consist of only two individuals. We compare the fitness of cooperators (C) who perform some cooperative behaviour, and defectors (D) who do not. The cooperative behaviour is assumed to be costly to the individual who performs it (cost = x), but provide a benefit to all the members in the group. We assume that the benefit to cost ratio is three, and so each cooperative behaviour brings a benefit of $3x$ to the group, and hence $3x/2$ to each individual. This cooperative behaviour would be classed as 'weakly altruistic'. Table 2 illustrates that from the selfish perspective of an individual, C always leads to a higher fitness, irrespective of whether it is in a group with a C or a D. This shows that weakly altruistic behaviours can be selected for because

Table 2 The fitness of cooperative individuals who perform a 'weakly altruistic' trait (C), and defectors who do not (D).

Group	Two cooperators		One cooperator		No coopera-tors	
Type of Individual	C	C	C	D	D	D
Baseline Fitness	1	1	1	1	1	1
Individual cost of cooperating	x	x	x	0	0	0
Benefit of cooperation (shared within group)	$\frac{6x}{2} = 3x$	$3x$	$\frac{3x}{2}$	$\frac{3x}{2}$	0	0
Benefit – cost	$3x - x = 2x$	$2x$	$\frac{3x}{2} - x = \frac{x}{2}$	$\frac{3x}{2}$	0	0
Fitness	$1 + 2x$	$1 + 2x$	$1 + \frac{x}{2}$	$1 + \frac{3x}{2}$	1	1

The calculation assumes that the cooperators (C) invest x resources in cooperation, the benefit to cost ratio is three, and that benefits are shared amongst all group members. From the selfish perspective of an individual, C always leads to a higher fitness, irrespective of whether it is in a group with a C or a D. This shows that a behaviour which would be classed as weakly altruistic can be selected for because it increases an individua'ls direct fitness. This table was inspired by an analogous one, in the sex ratio literature (Harvey *et al.*, 1985).

they increase an individual's direct fitness, and are hence mutually beneficial. More generally, the fundamental point is that the spread of a gene is determined by its fitness 'relative to others in the breeding population, and not to others with which it happens to interact' (Grafen, 1984, 2002, 2006; Harvey *et al.*, 1985). A more complicated case illustrating the same point is given in Table 3.

Altruism in humans and strong reciprocity

Altruism has been redefined in a number of ways in the literature on cooperation in humans. One approach in the empirical side of the literature has been to describe a costly behaviour as altruistic if it benefits another individual, but then to seemingly only measure the cost to the actor over the short term or relative to who they interact with (Fehr & Fischbacher, 2003). This clearly includes the possibility for either mutually beneficial (+/+) or altruistic (−/+) behaviours, and is analogous to the definition of cooperation (or helping). In contrast, the approach taken in the relevant theoretical literature, is to use altruism for behaviours which are costly but provide a benefit to all individuals in the group (Gintis, 2000;

Boyd *et al.*, 2003) – this is the same as weak altruism described above. In both cases we are left with the possibility that cooperation can be labelled as altruistic, even when it provides a direct benefit that can outweigh its cost, and hence can be selected for from selfish individual interests (i.e. mutually beneficial cooperation, with a +/+ payoff). We focus on the relevant theoretical literature in this section, and will return to the empirical literature later.

The usage of altruism in the human cooperation theoretical literature is best illustrated by examining a specific model. Gintis (2000) compared the relative fitness of two different strategies: 'self-interested agents' who do not punish or cooperate and altruistic 'strong reciprocators' who cooperate and punish noncooperators. He labels strong reciprocators as altruistic because they 'increase the fitness of unrelated individuals at a cost to themselves'. However, in this and related models, cooperation is individually costly within the social group, but provides a benefit to all the members of the group (including the cooperative individual), through mechanisms such as increased productivity or reducing the rate of group extinction (Gintis, 2000; Henrich & Boyd, 2001;

Table 3 The relative fitness of an individuals who cooperates (C) and defectors who do not (D), when more cooperative groups are more likely to survive and reproduce.

Group	One cooperator, and N-1 defectors		N defectors
Type of individual	C	D	D
Baseline resources	1	1	1
Public good contribution	1	0	0
Benefit to group	m	0	0
Individual resources after public goods game	$\frac{m}{N}$	$1 + \frac{m}{N}$	1
Total Group Resources	$m + N - 1$		N
Proportion of Group Resources	$\frac{m/N}{m+N-1}$	$\frac{1+m/N}{m+N-1}$	$\frac{1}{N}$
Relative reproductive success of groups	1		P
Fitness of Individuals	$\frac{m/N}{m+N-1}$	$\frac{1+m/N}{m+N-1}$	$\frac{P}{N}$

We compare the fitness of a cooperator in a group with $N - 1$ defectors, with the fitness of a defector in a group with N other defectors. See main text for details. Under certain conditions, from the selfish perspective of an individual, C leads to a higher fitness. This shows that a behaviour which some would class as altruistic and not in the individuals self interest, can be selected for because it increases an individual's direct fitness, and hence is actually mutually beneficial.

Table 4 The relationship between some of the terms that we have discussed. The different categories can be grouped according to whether they have beneficial (+) or costly (−) fitness consequences for the actor and recipient.

Term	Consequences for actor	Consequences for recipient
Altruism	−	+
Reciprocal altruism	+	+
Weak altruism	− or +	+
Strong reciprocity	− or +	− or +
Cooperation	− or +	+
Mutual benefit	+	+

Bowles *et al.*, 2003; Boyd *et al.*, 2003; Gintis, 2003; Bowles & Gintis, 2004). Consequently, cooperation can provide a direct fitness benefit, as well as the potential for indirect benefits due to individuals who share the cooperative gene.

The problem here is that the definition of altruism is relative to the local group, and not the population as a whole (as with 'weak altruism'). As discussed above, natural selection selects for a gene if it causes a behaviour that leads to that gene increasing in frequency in the population, not some other arbitrarily defined scale such as social partners (Grafen, 1984) (see Table 2). Another way of looking at this problem is that by examining altruism relative to the group this means that any benefits of the behaviour which are equally spread throughout the group are ignored – traits which provide a benefit at the group level will therefore seem altruistic because the benefits are ignored. Consequently, the model of Gintis (2000) leads to the confusing situation where cooperation can be favoured because it provides a direct benefit to the cooperator, because it increases the chance they and the rest of their group survive, but that this is defined as altruistic and not in their self interest. The same issue reoccurs in a number of related models (Bowles *et al.*, 2003; Boyd *et al.*, 2003; Gintis, 2003; Bowles & Gintis, 2004). The potential direct fitness benefit of cooperation is illustrated clearly by the model of Boyd *et al.* (2003), where groups compete for territories in pairs. In their model, the territory is won by the group with the most cooperators, and so it is clear that a single individual could potentially gain a huge direct fitness advantage by cooperating, and hence making its group much more successful. The role of direct fitness benefits in this model is further emphasized by the fact that it assumes that social behaviours are culturally transmitted, by imitation on the basis that they benefit the individual.

We illustrate how cooperation can provide a direct fitness benefit in such models in Table 3. We assume that groups of *N* individuals play a public goods game within groups, and that the more productive groups are more likely to survive or reproduce. We compare the fitness of a cooperator in a group with *N* − 1 defectors, with that of

a defector in a group of *N* defectors. Individuals have one unit of resource to contribute to the public goods game – cooperators contribute this whole unit, defectors contribute nothing. For each unit invested in the group, the whole group obtains a return of *m* units of resources to share equally amongst the group (cooperation provides a group benefit: *m* > 1). Consequently, as long as *m* < *N*, the cooperator ends up with less resources than they started with (compared with Table 2, where *m* > *N*). We then assume that the group with greater resources is more likely to survive or reproduce, such that the group with greater resources has a relative reproductive success of 1.0, and the other group has a relative reproductive success of *P* (*P* < 1). If the relative fitness of individuals is their share of the group resources multiplied by the group productivity, then the fitness of the cooperator will be greater than that of a defector in a group of *N* defectors if $m > \frac{P - NP}{P - 1}$. This demonstrates that cooperation can be mutually beneficial, and hence favoured by selfish interests (a direct fitness benefit), even in cases when playing the public goods game leads to the cooperator having less resources (*m* < *N*), and a lower fitness than the other members of the group ($\frac{m/N}{m + (N-1)} < \frac{1 + m/N}{m + (N-1)}$). This is because the benefit at the level of the group, of having just one individual cooperate, can outweigh this cost, and allow a cooperator to invade a population of defectors.

Punishment can be selfish or altruistic (like cooperation) or even spiteful, and so without detailed analysis of particular situations, the word punishment should not be given a prefix such as 'altruistic'. Punishment may provide a direct benefit because it can cause higher levels of cooperation within the punisher's group (Gardner & West, 2004b; Lehmann & Keller, 2006). If this is the case in the models discussed here, then 'selfish punishment' would have been the appropriate term, rather than 'altruistic punishment'. The alternative possibility is that punishment is favoured because it leads to an indirect fitness benefit, by: (i) making the punished individual more likely to cooperate with relatives of the punisher – altruistic punishment (Gardner & West, 2004b; Lehmann & Keller, 2006); or (ii) reducing the fitness of individuals who are competing with relatives – spiteful punishment (Gardner & West, 2004b). In the human models discussed above, the simulation approaches used mean that the relative importance of direct and indirect selection in favouring punishment are not clear. A detailed analysis of this problem would be extremely useful (Gardner & West, 2004b; Lehmann & Keller, 2006).

More altruism?

We have not exhausted the redefinitions of altruism. For example, altruism has also been to used to describe behaviours that benefit other individuals (e.g. Kaushik & Nanjundiah, 2003; Zahavi, 2003, 2005) – this could

clearly include behaviours which are altruistic (−/+) or mutually beneficial (+/+). However, we think that the examples we have given above are sufficient to illustrate the general points. In particular, that redefining altruism can obscure, and lead to confusion about the selective forces at work. We stress that by pointing out when behaviours are not altruistic, we do not hope to imply they are less interesting or easier to explain. Indeed, demonstrating direct long-term fitness benefits of a cooperative behaviour can be harder than demonstrating kin selected benefits (Griffin & West, 2002, 2003). Explaining cooperation remains one of the greatest challenges for evolutionary biology, irrespective of whether it is altruistic or mutually beneficial.

Group selection

The group selection literature has generated a huge amount of semantic confusion (reviewed by Dawkins, 1979; Maynard Smith, 1983; Grafen, 1984; Trivers, 1998; Foster *et al.*, 2006). Although this debate was solved decisively during the 1960s to 1980s, by evolutionary biologists, it seems to reoccur and lead to confusion as new fields embrace the relevant aspects of social evolution theory (Reeve & Keller, 1999). We have already discussed the phrase 'weak altruism'. In this section we briefly discuss some of the other points of confusion that have arisen. An interested reader is directed elsewhere for general reviews (Grafen, 1984; Dugatkin & Reeve, 1994), or more technical summaries (Wade, 1985; Frank, 1986a; Queller, 1992; Gardner *et al.*, 2007). Readers familiar with this work could easily skip this section.

The old and the new

Before discussing the semantic issues, it is useful to distinguish between two different types of group selection, and explain their relation to social evolution theory more generally (Grafen, 1984; Trivers, 1998). During the 1960s, Wynne-Edwards (1962) argued for the importance of group selection in its original or 'old' form. He considered relatively cooperative behaviours such as reproductive constraint, as follows. In groups consisting of selfish individuals (who reproduce at the maximum rate), resources would be over exploited, and the group would go extinct. In contrast, groups consisting of cooperative individuals who restricted their birth rate would not over exploit their resources, and not go extinct. Hence, by a process of differential survival of groups, behaviour evolved that was for the good of the group.

During the 1960s and 1970s a large body of theoretical and empirical work was piled up against this idea. Theory showed that this type of group selection would only work under extremely restrictive conditions, and so its importance would be rare or nonexistent (Maynard Smith, 1964, 1976; Williams, 1966; Leigh, 1983). For example,

Maynard Smith (1976) showed that group selection would not work if the number of individuals who disperse and reproduce elsewhere (successful migrants) is greater then one per group. Empirical work showed that individuals were reproducing at the rate that maximized their lifetime reproductive success, and were not practising reproductive restraint (Lack, 1966; Krebs & Davies, 1993). It is this form of group selection that leads people to the false conclusion that individuals behave for the good of the population or species or ecosystem.

In the 1970s and 1980s a new form of group selection was developed, based on a different conception of the group (Wilson, 1975a, 1977; Colwell, 1981; Wilson & Colwell, 1981). The idea here was that at a certain stages of an organism's life cycle, interactions take place between only a small number of individuals. It can be shown that under these conditions, cooperative behaviour can be favoured. This 'new group selection' is sometimes referred to as 'trait-group selection' or 'demic selection' or 'intrademic selection'. One way of conceptualizing the difference between the old and new group selection models is that the new group selection models rely on within-population (intrademic) group selection, whereas old group selection theory worked on between-population (interdemic) group selection (Fig. 2; Reeve & Keller, 1999). Another difference is that the old group selection approach argued that selection at the group level was the driving force of natural selection, whereas the new group selection emphasizes that there are multiple levels of selection, and these can vary in their importance. Another way of looking at this is that the new group selection approach looks at the evolution of individual characters in a group structured population, whereas the old group selection approach looks at the evolution of group characters (Fig. 2; Okasha, 2005).

It has since been shown that kin selection and new group selection are just different ways of conceptualizing the same evolutionary process. They are mathematically identical, and hence are both valid (Hamilton, 1975; Grafen, 1984; Wade, 1985; Frank, 1986a, 1998; Taylor, 1990; Queller, 1992; Bourke & Franks, 1995; Gardner *et al.*, 2007). New group selection models show that cooperation is favoured when the response to between-group selection outweighs the response to within-group selection, but it is straightforward to recover Hamilton's rule from this. Both approaches tell us that increasing the group benefits and reducing the individual cost favours cooperation. Similarly, group selection tells us that cooperation is favoured if we increase the proportion of genetic variance that is between-group as opposed to within-group, but that is exactly equivalent to saying that the kin selection coefficient of relatedness is increased (Frank, 1995a). In all cases where both methods have been used to look at the same problem, they give identical results (Frank, 1986a; Bourke & Franks, 1995; Wenseleers *et al.*, 2004; Gardner *et al.*, 2007).
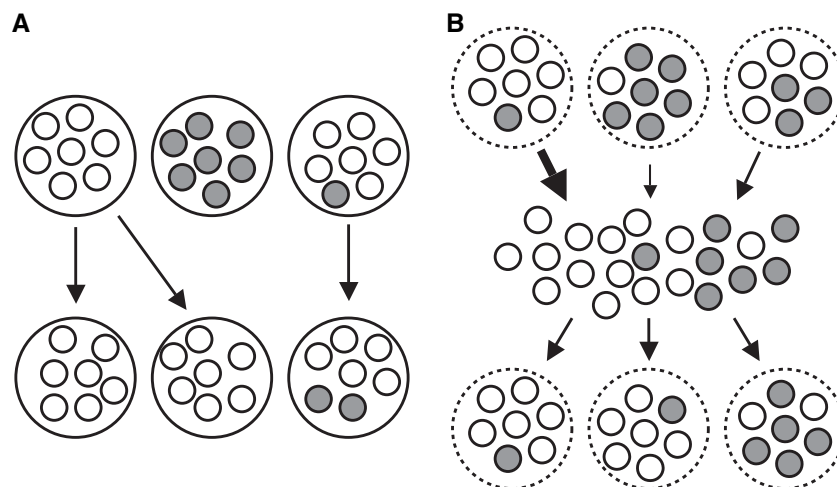
Fig. 2 The difference between old and new group selection. Panel A shows the old group selection, with well-defined groups with little gene flow between them (solid outline). The white circles represent cooperators, whereas the grey circles represent selfish individuals who do not cooperate. Competition and reproduction is between groups. The groups with more cooperators do better, but selfish individuals can spread within groups. Panel B shows the new group selection, with arbitrarily defined groups (dashed lines), and the potential for more geneflow between them. The different groups make different contributions to the same reproductive pool (although there is also the possibility of factors such as limited dispersal leading to more structuring), from which new groups are formed.

Avoiding confusion

In addition to the problems associated with the term 'weak altruism', the group selection literature has produced three other sources of semantic confusion. First, the different types of group selection can be mixed up (Grafen, 1984; Trivers, 1998; Okasha, 2005). This can give the impression that the validity of the new group selection justifies the application of the old group selection (Trivers, 1998). This can be a particular problem with nonspecialists or nontheoreticians, and we believe that it plays a major role in explaining why old group selection lingers in some fields, such as areas of microbiology (e.g. Shapiro & Dworkin, 1997; Shapiro, 1998; Bassler, 2002; Henke & Bassler, 2004), parasitology (reviewed by West *et al.*, 2001, 2003), and agronomy (reviewed by Denison *et al.*, 2003). Numerous examples of this problem are also provided by Sober & Wilson (1998), who switch confusingly between old and new group selection (Trivers, 1998).

The potential for confusion is increased when the impression is gained that evolutionary biologists do not use the new group selection methodology because they think it is unimportant (as they do with the old). In reality, the kin selection (or inclusive fitness) approach is generally preferred over new group selection because it is usually easier to construct models, interpret the predictions, and then apply these to real biological cases. For example: (i) recent methodological advances mean that kin selection and inclusive fitness models can be constructed and analysed much more simply, and for much more general cases (Taylor & Frank, 1996; Frank,

1997, 1998; Taylor *et al.*, 2007); (ii) in some of the most successful areas of social evolution, such as split sex ratios in social insects or extensions of Hamilton's (1967) basic local mate competition theory, predictions arise elegantly from kin-selection models, whereas the corresponding group selection models would be either unfeasible or so complex that they have not been developed (Frank, 1986b, 1998; Boomsma & Grafen, 1991; Queller, 2004; Shuker *et al.*, 2005); (iii) kin selection methodologies can usually be linked more clearly to empirical research, both empirically (Queller & Goodnight, 1989) and conceptually – 'knowing that $r = 0.22$ gives many biologists an understanding of the genetic closeness described; the knowledge that $n = 10$ and $v/v_b = 2.98$ is (at least for the present) less illuminating' (Grafen, 1984); (iv) the group selection approach tends to hide the distinction between direct and indirect benefits of behaviours, which many find extremely useful; (v) the group selection approach has been less useful for identifying and quantifying issues of reproductive conflict, which have provided some of the most useful areas for empirical testing of theory (Trivers, 1974; Ratnieks *et al.*, 2006); (vi) inclusive fitness theory leads to the recovery of a maximizing principle for social settings (individuals should behave as if maximizing their inclusive fitness), which is a useful reasoning tool (it is easier to think of individuals optimizing something rather than the evolutionary dynamics), provides formal justification for the use of intentional language (selfishness, altruism, conflict), and legitimizes discussion of 'function' and 'design' of social behaviours (Hamilton, 1964; Grafen, 1999, 2006) – whereas the group selection view does not lead so easily to a

maximizing principle (instead it emphasizes the tension between individuals and groups), and indeed, people looking for one could fall into the trap of old group selection, which incorrectly views Darwinian individuals as acting to maximize group or species fitness.

A second source of confusion is the incorrect idea that inclusive fitness theory or kin selection are distinct from, or just special cases of, new group selection. This idea lingers on, with group selection being invoked in cases where kin selection is suggested to be unimportant (reviewed by Gardner & West, 2004b; Foster *et al.*, 2006; Helantera & Bargum, 2007). Recent examples can be found in the human (e.g. Sober & Wilson, 1998), social insect (e.g. Wilson, 2005; Wilson & Hölldobler, 2005) and microbial (e.g. Kreft, 2005) literature. Whilst coancestry is without doubt the most important cause of relatedness between individuals, and the underlying factor in the examples discussed above, the inclusive fitness approach applies more generally when genetic correlations between social partners occur for any reason. Although this point was realized by Hamilton (1964, 1975), and has been developed substantially (Frank, 1998), it is often not appreciated. Inclusive fitness is a generalization of Darwinian fitness, and inclusive fitness theory (kin selection in its broadest form) is a generalized description of natural selection, and these are not simply special cases that are appropriate only for when individuals interact with their relatives (Grafen, 2006). Perhaps most importantly, there is no biological model or empirical example that can be explained with the new group selection approach, that cannot also be understood in terms of kin selection and inclusive fitness.

A third source of confusion is that the new group selection approach has involved the use of several fundamental terms in ways that were different from their established (valuable and clear) meanings (Dawkins, 1979; Grafen, 1984). Specifically it has identified within-group selection as 'individual selection' and between-group selection as 'group selection'. This can be confusing as in some situations a social behaviour can be selected for by a selfish direct benefit, but be classed as being selected against by 'individual selection' and favoured by 'group selection' (Grafen, 1984; Tables 2 and 3). Even a nonsocial trait can be ascribed a group selection component simply because groups containing fitter individuals are themselves favoured by selection (Hamilton, 1975; Okasha, 2004)! Group selection was originally defined as the differential survival or extinction of whole groups (Maynard Smith, 1976). The new group selection models do not rely on the maintenance of whole groups, and so the terms trait-group or demic selection are perhaps more appropriate. An alternative is to state as simply as possible what they are – models of nonrandom assortment of altruistic genes (e.g. because of relatedness, or altruists choosing to interact with each other; Maynard Smith, 1976). When this is done, the links with inclusive fitness theory become transparently

clear. The potential meanings for 'group selection' are so numerous because the partitioning of selection into within-group and between-group components can be done for any arbitrarily defined group (Wade, 1985). Indeed, group selection has also been used to describe species level sorting (Williams, 1992), and new variants are being introduced such as 'cultural group selection' (see below). These points emphasize the use and power of Hamilton's original terminology, and the gains to be made from the minimal use of jargon.

Direct fitness

Direct fitness is usually used to describe the component of fitness gained through the behaviour of an individual influencing their reproductive success. However, in recent years, the term direct fitness has also been used to describe a powerful method for constructing kin selection models, which also allows for how an individuals fitness is influenced by their social partners (Taylor & Frank, 1996; Frank, 1997, 1998; Rousset, 2004; Taylor *et al.*, 2007). This approach involves: (i) writing down how the personal fitness of an individual depends upon their behaviour, and the behaviour of the individuals with which they interact; (ii) differentiating this equation in such a way that the appropriate relatedness term appears, leading to an expression that can be conceptualized alternatively as a direct fitness effect or as an inclusive fitness effect (these two views are mathematically equivalent). The development of this methodology has revolutionized social evolution theory, providing a simpler method that produces more general models, compared with the inclusive fitness approach. Happily, because of the mathematical equivalence, the results of direct fitness analyses are readily interpreted in terms of inclusive fitness, which can be a more natural way of understanding kin selection (Taylor *et al.*, 2007).

However, this new use of the term direct fitness can cause confusion. It is clear why this method for constructing kin selection models could be called the direct fitness approach – the first step involves writing the direct (or personal) fitness of an individual. The confusion is that by terming this the direct fitness approach, it can lead to the impression that it ignores the indirect fitness consequences of the behaviour and so does not take account of inclusive fitness. However, it does. The direct fitness approach has also been termed the 'neighbour-modulated' approach (Hamilton, 1964; Maynard Smith, 1983; Grafen, 2006).

Proximate and ultimate explanations

The Nobel Prize winner Niko Tinbergen famously clarified the different approaches to studying animal behaviour, in the most influential paper of his career (Tinbergen, 1963; Kruuk, 2003). His key insight was to show that the different methodologies are

complementary and not alternatives. Of particular relevance here is his distinction between: (i) proximate explanations which are concerned with the mechanisms underlying a behaviour (causation; how questions) and (ii) ultimate explanations which examine the fitness consequences or survival value of a behaviour (why questions) (Mayr, 1961).

One of Tinbergen's classic studies to illustrate this distinction was on the removal of eggshells from their nests by black-headed gulls. The mechanistic (proximate) explanation for this is that individuals are more likely to remove objects from their nest if they are white or egg-coloured, have frilly edges, and if they are feather-light. The evolutionary (ultimate) explanation for this is that it makes aerial predators such as herring gulls less likely to find their brood. These explanations are clearly not competing (each answer cannot provide a solution to the other problem), and a fuller understanding is gained by considering both. In this section we use the literature on cooperation in humans to illustrate the problems that can arise through blurring these approaches. In particular, we show how distinguishing between these hypotheses: (i) clarifies the evolutionary forces at work and (ii) allows human behaviour to be placed in a wider context.

Strong reciprocity in humans

In the human cooperation literature, the definition and discussion of terms related to altruism has often mixed proximate and ultimate factors. An example of this is provided by 'strong reciprocity', which is defined proximately, but then given as a solution to an ultimate problem. A strong reciprocator has been defined as a combination of 'a predisposition to reward others for cooperative, norm-abiding behaviours' and 'a propensity to impose sanctions on others for norm violations' (Fehr & Fischbacher, 2003). This is a description of a proximate mechanism. However, it is then given as a solution to an ultimate problem: 'Strong reciprocity thus constitutes a powerful incentive for cooperation even in non-repeated interactions when reputation gains are absent' (Fehr & Fischbacher, 2003).

This approach mixes up two different questions (how and why?). The proximate question is how is cooperation maintained? The answer to this is through punishment and reward – i.e. what has been termed strong reciprocity. The ultimate question is why is cooperation maintained, or more specifically, why is cooperation and punishment (strong reciprocity) maintained? The possible answers to this are because it provides either a direct or an indirect fitness benefit. The related theory, which we have discussed above, suggests that the answer is a direct-fitness benefit, because cooperation increases group productivity and/or survival and hence provides a direct benefit to individuals who cooperate.

The difference between proximate and ultimate questions can be illustrated with a discussion of the impor-

tance of the role of imitation (cultural transmission) or learning (Boyd *et al.*, 2003). Learning and imitation are a solution to the proximate not the ultimate problem. They provide a mechanism to learn how to punish and cooperate, and at a rate that maximizes fitness. For example, the imitation of successful strategies (Boyd *et al.*, 2003). The ultimate question is why is cooperation and punishment favoured in the first place? Put another way, what makes punishers and cooperators so successful that they are worth imitating? Another example is provided by neurobiological work. Experiments have shown that punishment of individuals who do not cooperate leads to stimulation in the dorsal striatum, an area which has been implicated in reward-related brain circuits (Quervain *et al.*, 2004). Consequently, at a proximate level, individuals punish others who do not cooperate because it gives them 'satisfaction'. The ultimate question is why has the brain circuitry evolved so that punishment provides satisfaction. The answer to this must be that punishment has provided some fitness advantage, either direct or indirect.

The potential confusion that can be caused by these semantic issues is illustrated by the suggestion that standard evolutionary theory cannot explain cooperation between humans, and that alternatives such as 'cultural group selection' or 'gene-culture coevolution' are needed (Bowles *et al.*, 2003; Fehr & Fischbacher, 2003). Whilst these alternative models work on standard social evolution principles, the terminology involved gives the misleading impression they do not. For example, Gintis (2000) describes how strong reciprocity 'cannot be justified in terms of self-interest', but then as we have explained in a previous section, his model to explain it appears to rely on cooperation providing a direct fitness benefit. To give another example, Fehr & Fischbacher (2003) suggest that 'cultural group selection' or 'gene-culture coevolution' provide an alternative to selfish or kin selected explanations of cooperation. However, the models that they refer to (Henrich & Boyd, 2001; Bowles *et al.*, 2003; Boyd *et al.*, 2003; Gintis, 2003) appear to rely on selfish benefits to cooperation (although also possibly some indirect benefits as we have discussed earlier). Social learning answers how questions, not why questions – whilst both of these are important, they are different (although there can be some interesting interactions, which we discuss below).

Interactions and special humans

A lack of a clear distinction between ultimate and proximate factors can obscure biological differences and similarities. This is illustrated by the discussion of whether cooperation in humans is special. Bowles & Gintis (2004) suggest that a problem is 'why similar behaviours are seldom observed in other animals'. We would suggest that at the ultimate level this statement is incorrect. Punishment has been shown or argued to

stabilize cooperation in a number of situations, not just between animals, but also with plants and bacteria (reviewed by Trivers, 1985; Clutton-Brock & Parker, 1995; Frank, 2003; Sachs *et al.*, 2004; Foster & Wenseleers, 2006; Kiers & van der Heijden, 2006; Ratnieks *et al.*, 2006; West *et al.*, 2006b). For example: between legume plants and their rhizobia bacteria (West *et al.*, 2002a; Kiers *et al.*, 2003, 2006), within the social hymenoptera (ants, bees and wasps; Ratnieks & Visscher, 1989; Ratnieks *et al.*, 2006), between yucca plants and their pollinator moths (Pellmyr & Huth, 1994), between cleaner fish (Bshary & Grutter, 2002, 2005), and within colonies of naked mole-rats (Reeve, 1992). Furthermore, the direct benefits of cooperation, via increasing group survival, and hence individual survival, have been suggested to be important in a range of species, especially cooperative breeding vertebrates, but also social insects (Wiley & Rabenold, 1984; Queller *et al.*, 2000; Kokko *et al.*, 2001; Clutton-Brock, 2002; Griffin & West, 2002).

In contrast, what appears to be special about cooperation in humans is the proximate factors involved. Experimental work has shown very sophisticated punishment and reward systems, which can be facultatively fine tuned in response to variation in local conditions that alter the direct benefit of cooperation (Fehr & Gächter, 2002; Wedekind & Braithwaite, 2002; Fehr & Fischbacher, 2003; Fehr & Rockenbach, 2003; Fehr & Fischbacher, 2004; Henrich *et al.*, 2005; Crespi, 2006; West *et al.*, 2006a). For example, human cognitive abilities allow individuals to alter their level of cooperation in response to whether there is the possibility for punishment (Fehr & Gächter, 2002), and whether they are competing locally or globally for resources (West *et al.*, 2006a). Furthermore, although proximate factors alone cannot supply the answer to the ultimate problem of why cooperate, there is the possibility for some interesting interactions that can alter the evolutionary dynamics, and provide a mechanism that could rapidly generate cultural differences (Boyd *et al.*, 2003; Gintis, 2003; Hagen & Hammerstein, 2006). For example, imitation of successful strategies allows individuals to fine tune their behaviour in response to local conditions. Consequently, although plants and bees have impressive strategies for enforcing cooperation with punishment, humans can be even more sophisticated.

Interactions between ultimate and proximate factors can also play a key role in constraining adaptation, and hence determining how perfect behaviours should be. Proximate mechanisms determine the possibilities that evolution can work on, and hence the level of perfection that should be expected in behaviours (West & Sheldon, 2002; Boomsma *et al.*, 2003; Shuker & West, 2004). This interplay is likely to be key to explaining the experimental result that humans will cooperate and punish in one-shot interactions, where they will never get to interact with the same person again (Fehr & Gächter, 2002; Fehr & Fischbacher, 2003; Fehr & Rockenbach,

2003; Fehr & Fischbacher, 2004). From an ultimate perspective, theory has suggested that cooperation and punishment have generally been favoured in humans because they have provided direct fitness benefits. From a proximate perspective, the way humans do this is what has been termed strong reciprocity, and can be fine tuned to local conditions with social learning (see above). A problem with this proximate mechanism is that it can lead to imperfect behaviour when the individual is subjected to some situations, such as one-shot interactions, where there are no direct fitness benefits to cooperation. However, it is well accepted that numerous factors are expected to constrain adaptation and prevent animals from behaving perfectly in certain test situations (Herre, 1987; Wehner, 1987; Parker & Maynard Smith, 1990; Partridge & Sibley, 1991; Herre *et al.*, 2001; West & Sheldon, 2002; Boomsma *et al.*, 2003; Shuker & West, 2004).

Defining behaviours will not always be easy or the most useful thing to do

Although we believe that well-defined terms are needed, we would also like to stress that there are situations where defining a specific behaviour could be extremely hard, and might not be the most useful way forward. In particular, when a cooperative behaviour can provide both direct and indirect benefits, it can be hard to determine whether it is altruistic or mutually beneficial, because this will depend upon the relative importance of the direct and indirect benefits. This could happen with whole-group behaviours that provide a benefit to everyone within the local group, including the actor. For example: (i) the production of public goods, such as iron-scavenging siderophore molecules, by many bacteria (West & Buckling, 2003); (ii) helping in cooperative-breeding vertebrates, if that leads to a larger group size, which increases the survival of everyone, including the helper (termed group augmentation; Kokko *et al.*, 2001). In cases such as this it can be useful to describe behaviours in terms of who incurs costs and benefits, and not their net sum for the actor.

A related problem is that behaviours should be classified according to their impact on total lifetime reproductive success. So, whilst a cooperative behaviour that increased the chance of attaining reproductive dominance in a group would be classified as mutually beneficial, this may not be immediately obvious. Fitting specific behaviours into this classification can be hard because of difficulties with measuring future (and hence total lifetime) fitness consequences (Griffin *et al.*, 2003; MacColl & Hatchwell, 2004). This means that in some situations there can be an advantage to describing behaviours according to immediate effects, as the future effects may not be certain, and we are often not in a position to view them (Trivers, 1985). However, this also emphasizes the pitfalls of describing behaviours with

terms such as altruism, based upon immediate fitness benefits to participants, rather than total lifetime benefits.

Useful predictions can often be made and tested, even when we do not know whether a cooperative behaviour is altruistic or mutually beneficial. We illustrate this with three examples. First, Taylor (1992a,b) showed that, contrary to expectations, limited dispersal would not necessarily favour higher levels of cooperation. However, the cooperative behaviour modelled by Taylor provides a benefit to everyone in the group, and can be altruistic or mutually beneficial depending upon exact parameter values (Rousset, 2004). Secondly, it is possible to make predictions for how selection for investment into public goods should vary with environmental conditions and population structure (West & Buckling, 2003), and to test these predictions (Griffin *et al.*, 2004), without determining whether this is altruistic or mutually beneficial. Thirdly, in cooperative breeding vertebrates, we can predict that the extent to which helpers preferentially aid relatives (kin discrimination), will vary with the benefit of helping, despite the fact that we do not know the relative importance of indirect and direct fitness benefits of helping for particular species (Griffin & West, 2003).

Another general issue is that, although we have focused on cooperative behaviours that are clearly helping another individual (+/+ or −/+), the classification in Table 1 can be applied to all forms of social behaviour. Other possibilities would include dispersal or sex ratio evolution. For example, dispersal can be altruistic (−/+) if it reduces the local competition for resources and hence provides a benefit to relatives, or mutually beneficial (+/+) if it also provides a direct benefit such as moving to a better habitat or a reduced likelihood of inbreeding (Hamilton & May, 1977). When considering specific cases it will often not be necessary or useful to think about behaviours such as the sex ratio or dispersal in terms of the classification provided in Table 1 (or Hamilton's rule). However, the insight that these behaviours can be thought about in the same ways was a huge breakthrough that has allowed very general overviews to be developed (Hamilton, 1971, 1972, 1975; Frank, 1986a, 1998; Taylor, 1990, 1996; Taylor & Frank, 1996). Indeed one reason why evolutionary biologists may have solved the levels of selection debate that still rages in other areas is that they have been able to apply theory to areas such as the sex ratio that are not so loaded with preconceptions of how individuals are expected (or ought) to behave.

Conclusions

Science depends upon the communication of facts and ideas. The misuse or redefinition of widely used terms can hinder this. In the case of altruism, the various redefinitions that we have discussed are well meaning, but lead to confusion because there are so many potential

ways to do it (Table 4). For example, with respect to population or local group, with respect to personal or inclusive fitness, with respect to short or long term etc. Consequently, unless we are careful, a term that had a single useful meaning can becomes meaningless and confusing.

We have discussed the human cooperation literature in some detail because we think this illustrates the general points. Another area where this problem is growing, is the literature on cooperation in microbes (reviewed by West *et al.*, 2006b). First, the old group selection ideas, and even species or community level selection, are still suggested (e.g. Shapiro & Dworkin, 1997; Shapiro, 1998; Bassler, 2002; Henke & Bassler, 2004). Secondly, terms have been redefined. Altruism has been defined relative to the local scale, analogous to weak altruism (e.g. Kreft, 2005), or to mean a behaviour that benefits others (e.g. Kaushik & Nanjundiah, 2003), and cooperation has been defined to refer to only the specific case of public goods (e.g. Velicer, 2003; Travisano & Velicer, 2004). Thirdly, it has been assumed that proximate answers can be given to ultimate questions. We suspect that this proximate/ultimate issue is likely to be particularly important in research on communication between bacteria (quorum sensing), where signals can be defined both proximately and ultimately (reviewed by Diggle *et al.*, 2007; Keller & Surette, 2006).

A general point here is that the potential for semantic confusion is greatest with interdisciplinary research. One reason for this is that different fields may use the same term differently (Read, 1994). For example, we have focused on the evolutionary definition of altruism, whereas the psychological definition is based on motivations (Quervain *et al.*, 2004). Another reason is that different areas may focus traditionally on different questions, such as proximate vs. ultimate. In all cases, confusion is best avoided by clear and specific statements that minimize jargon.

References

Alcock, J. 1989. *Animal Behavior*, 4th edn. Sinauer Associates, Sunderland, Massachusetts.
Alexander, R.D. 1974. The evolution of social behavior. *Annu. Rev. Ecol. Syst.* **5**: 325–383.

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

Bassler, B.L. 2002. Small talk: cell-to-cell communication in bacteria. *Cell* **109**: 421–424.

Boomsma, J.J. & Grafen, A. 1991. Colony-level sex ratio selection in the eusocial Hymenoptera. *J. Evol. Biol.* **44**: 383–407.

Boomsma, J.J. & Ratnieks, F.L.W. 1996. Paternity in eusocial Hymenoptera. *Phil. Trans. R. Soc. Lond. B* **351**: 947–975.

Boomsma, J.J., Nielsen, J., Sundstrom, L., Oldham, N.J., Tentschert, J., Petersen, H.C. & Morgan, E.D. 2003. Informational constraints on optimal sex allocation in ants. *Proc. Natl. Acad. Sci. USA* **100**: 8799–8804.

Bourke, A.F.G. & Franks, N.R. 1995. *Social Evolution in Ants*. Princeton University Press, Princeton, New Jersey.

Bowles, S. & Gintis, H. 2004. The evolution of strong reciprocity: cooperation in heterogeneous populations. *Theor. Popul. Biol.* **65**: 17–28.

Bowles, S., Choi, J.-K. & Hopfensitz, A. 2003. The co-evolution of individual behaviors and social institutions. *J. Theor. Biol.* **223**: 135–147.

Boyd, R., Gintis, H., Bowles, S. & Richerson, P.J. 2003. The evolution of altruistic punishment. *Proc. Natl. Acad. Sci. USA* **100**: 3531–3535.

Brown, J.L. 1983. Cooperation – a biologist's dilemma. *Adv. Study behav.* **13**: 1–37.

Brown, J.L. & Brown, E.R. 1981. Kin selection and individual selection in babblers. In: *Natural Selection and Social Behavior: Recent Research and New Theory* (R. D. Alexander & D. W. Tinkle, eds), pp. 244–256. Chiron Press, New York.

Bshary, R. & Grutter, A.S. 2002. Experimental evidence that partner choice is a driving forces in the payoff distribution among cooperators or mutualists: the cleaner fish case. *Ecol. Lett.* **5**: 130–136.

Bshary, R. & Grutter, A.S. 2005. Punishment and partner switching cause cooperative behaviour in a cleaning mutualism. *Biol. Lett.* **1**: 396–399.

Burt, A. 2000. Sex, recombination and the efficacy of selection – was Weismann right? *Evolution* **54**: 337–351.

Clutton-Brock, T.H. 2002. Breeding together: kin selection, reciprocity and mutualism in cooperative animal societies. *Science* **296**: 69–72.

Clutton-Brock, T.H. & Parker, G.A. 1995. Punishment in animal societies. *Nature* **373**: 209–216.

Colwell, R.K. 1981. Group selection is implicated in the evolution of female-biased sex ratios. *Nature* **290**: 401–404.

Conner, R.C. 1995a. The benefits of mutualism: a conceptual framework. *Biol. Rev.* **70**: 427–457.

Conner, R.C. 1995b. Altruism among non-relatives: alternatives to the 'Prisoner's Dilemma'. *Trends Ecol. Evol.* **10**: 84–86.

Crespi, B.J. 2006. Cooperation: close friends and common enemies. *Curr. Biol.* **16**: R414–R415.

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

Dawkins, R. 1979. Twelve misunderstandings of kin selection. *Z. Tierpsychol.* **51**: 184–200.

Denison, R.F., Kiers, E.T. & West, S.A. 2003. Darwinian agriculture: when can humans find solutions beyond the reach of natural selection. *Q. Rev. Biol.* **78**: 145–168.

Diggle, S.P., Gardner, A., West, S.A. & Griffin, A.S. 2007. Evolutionary theory of bacterial quorum sensing: when is a signal not a signal? *Phil. Trans. R. Soc. Lond. B* in press.

Dugatkin, L.A. 1997. *Cooperation among Animals: An Evolutionary Perspective*. Oxford University Press, Oxford.

Dugatkin, L.A. & Reeve, H.K. 1994. Behavioral ecology and levels of selection: dissolving the group selection controversy. *Adv. Study Behav.* **23**: 101–133.

Dugatkin, L.A., Perlin, M. & Atlas, R. 2003. The evolution of group-beneficial traits in the absence of between-group selection. *J. Theor. Biol.* **220**: 67–74.

Dugatkin, L.A., Perlin, M., Lucas, J.S. & Atlas, R. 2005. Group-beneficial traits, frequency-dependent selection and genotypic diversity: an antibiotic resistance paradigm. *Proc. Roy. Soc. Lond. B* **272**: 79–83.

Emlen, S.T. 1997. Predicting family dynamics in social vertebrates. In: *Behavioural Ecology* (J. R. Krebs & N. B. Davies, eds), pp. 228–253. Blackwell, Oxford, UK.

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

Fehr, E. & Fischbacher, U. 2004. Third-party punishment and social norms. *Evol. Hum. Behav.* **25**: 63–87.

Fehr, E. & Gächter, S. 2002. Altruistic punishment in humans. *Nature* **415**: 137–140.

Fehr, E. & Rockenbach, B. 2003. Detrimental effects of sanctions on human altruism. *Nature* **422**: 137–140.

Fisher, R.A. 1930. *The Genetical Theory of Natural Selection*. Clarendon, Oxford.

Fletcher, J.A. & Doebeli, M. 2006. How altruism evolves: assortment and synergy. *J. Evol. Biol.* **19**: 1389–1393.

Foster, K.R. & Wenseleers, T. 2006. A general model for the evolution of mutualisms. *J. Evol. Biol.* **19**: 1283–1293.

Foster, K.R., Wenseleers, T. & Ratnieks, F.L.W. 2001. Spite: Hamilton's unproven theory. *Ann. Zool. Fennici* **38**: 229–238.

Foster, K.R., Wenseleers, T. & Ratnieks, F.L.W. 2006. Kin selection is the key to altruism. *Trends Ecol. Evol.* **21**: 57–60.

Frank, S.A. 1986a. Hierarchical selection theory and sex ratios. I. General solutions for structured populations. *Theoret. Popul. Biol.* **29**: 312–342.

Frank, S.A. 1986b. The genetic value of sons and daughters. *Heredity.* **56**: 351–354.

Frank, S.A. 1995a. George Price's contributions to evolutionary genetics. *J. Theor. Biol.* **175**: 373–388.

Frank, S.A. 1995b. Mutual policing and repression of competition in the evolution of cooperative groups. *Nature* **377**: 520–522.

Frank, S.A. 1997. Multivariate analysis of correlated selection and kin selection, with an ESS maximization method. *J. Theor. Biol.* **189**: 307–316.

Frank, S.A. 1998. *Foundations of Social Evolution*. Princeton University Press, Princeton.

Frank, S.A. 2003. Repression of competition and the evolution of cooperation. *Evolution* **57**: 693–705.

Frank, S.A. 2006. Social selection. In: *Evolutionary Genetics: Concepts and Case Studies* (C. W. Fox & J. B. Wolf, eds), pp. 350–363. Oxford University Press, Oxford.

Futuyma, D.J. 2005. *Evolution*. Sinauer Associates, Sunderland, Massachusetts.

Gardner, A. & West, S.A. 2004a. Spite and the scale of competition. *J. Evol. Biol.* **17**: 1195–1203.

Gardner, A. & West, S.A. 2004b. Cooperation and punishment, especially in humans. *Am. Nat.* **164**: 753–764.

Gardner, A., West, S.A. & Barton, N.H. 2007. The relation between multilocus population genetics and social evolution theory. *Am. Nat.*, in press.

Gintis, H. 2000. Strong reciprocity and human sociality. *J. Theor. Biol.* **206**: 169–179.

Gintis, H. 2003. The hitchhiker's guide to altruism: gene-culture coevolution, and the internalization of norms. *J. Theor. Biol.* **220**: 407–418.

Grafen, A. 1984. Natural selection, kin selection and group selection. In: *Behavioural Ecology: An Evolutionary Approach* (J. R. Krebs & N. B. Davies, eds), pp. 62–84. Blackwell Scientific Publications, Oxford, UK.

Grafen, A. 1985. A geometric view of relatedness. *Oxford Surv. Evol. Biol.* **2**: 28–89.

Grafen, A. 1990. Biological signals as handicaps. *J. Theor. Biol.* **144**: 517–546.

Grafen, A. 1999. Formal Darwinism, the individual-as-maximizing-agent analogy and bet-hedging. *Proc. Roy. Soc. Lond. B* **266**: 799–803.

Grafen, A. 2002. A first formal link between the Price equation and an optimization program. *J. Theor. Biol.* **217**: 75–91.

Grafen, A. 2006. Optimisation of inclusive fitness. *J. Theor. Biol.* **238**: 541–563.

Griffin, A.S. & West, S.A. 2002. Kin selection: fact and fiction. *Trends Ecol. Evol.* **17**: 15–21.

Griffin, A.S. & West, S.A. 2003. Kin discrimination and the benefit of helping in cooperatively breeding vertebrates. *Science* **302**: 634–636.

Griffin, A.S., Pemberton, J.M., Brotherton, P.N.M., McIlrath, G., Gaynor, D., Kansky, R., O'Riain, J. & Clutton-Brock, T.H. 2003. A genetic analysis of breeding success in the cooperative meerkat (*Suricata suricatta*). *Behav. Ecol.* **14**: 472–480.

Griffin, A.S., West, S.A. & Buckling, A. 2004. Cooperation and competition in pathogenic bacteria. *Nature* **430**: 1024–1027.

Hagen, E.H. & Hammerstein, P. 2006. Game theory and human evolution: a critique of some recent interpretations of experimental games. *Theor. Popul. Biol.* **69**: 339–348.

Hamilton, W.D. 1963. The evolution of altruistic behaviour. *Am. Nat.* **97**: 354–356.

Hamilton, W.D. 1964. The genetical evolution of social behaviour, I & II. *J. Theor. Biol.* **7**: 1–52.

Hamilton, W.D. 1967. Extraordinary sex ratios. *Science* **156**: 477–488.

Hamilton, W.D. 1970. Selfish and spiteful behaviour in an evolutionary model. *Nature* **228**: 1218–1220.

Hamilton, W.D. 1971. Selection of selfish and altruistic behaviour in some extreme models. In: *Man and Beast: Comparative Social Behavior* (J. F. Eisenberg & W. S. Dillon, eds), pp. 57–91. Smithsonian Press, Washington, DC.

Hamilton, W.D. 1972. Altruism and related phenomena, mainly in social insects. *Annu. Rev. Ecol. Syst.* **3**: 193–232.

Hamilton, W.D. 1975. Innate social aptitudes of man: an approach from evolutionary genetics. In: *Biosocial Anthropology* (R. Fox, ed.), pp. 133–155. Wiley, New York.

Hamilton, W.D. 1996. *Narrow roads of gene land: I Evolution of Social Behaviour*. W. H. Freeman, Oxford.

Hamilton, W.D. & May, R. 1977. Dispersal in stable habitats. *Nature* **269**: 578–581.

Hammerstein, P. 2003. *Genetic and Cultural Evolution of Cooperation*. MIT Press, Cambridge.

Harvey, P.H., Partridge, L. & Nunney, L. 1985. Group selection and the sex ratio. *Nature* **313**: 10–11.

Helantera, H. & Bargum, K. 2007. Pedigree relatedness, not greenbeard genes, explains eusociality. *Oikos* (in press).

Henke, J.M. & Bassler, B.L. 2004. Bacterial social engagements. *Trends Cell Biol.* **14**: 648–656.

Henrich, J. & Boyd, R. 2001. Why people punish defectors. *J. Theor. Biol.* **208**: 79–89.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., McElreath, R., Alvard, M., Barr, A., Ensminger, J., Henrich, N.S., Hill, K., Gil-White, F., Gurven, M., Marlowe, F.W., Patton, J.Q. & Tracer, D. 2005. ''Economic man'' in cross-cultural perspective: behavioral experiments in 15 small-scale societies. *Behavioral and Brain Sciences* **28**: 795–855.

Herre, E.A. 1987. Optimality, plasticity and selective regime in fig wasp sex ratios. *Nature* **329**: 627–629.

Herre, E.A., Knowlton, N., Mueller, U.G. & Rehner, S.A. 1999. The evolution of mutualisms: exploring the paths between conflict and cooperation. *Trends Ecol. Evol.* **14**: 49–53.

Herre, E.A., Machado, C.A. & West, S.A. 2001. Selective Regime and Fig Wasp Sex Ratios: Towards Sorting Rigor from Pseudo-Rigor in Tests of Adaptation. In: *Adaptionism and Optimality* (S. Orzack & E. Sober, eds), pp. 191–218. Cambridge University Press, Cambridge.

Jansen, V.A.A. & van Baalen, M. 2006. Altruism through beard chromodynamics. *Nature* **440**: 663–666.

Kaushik, S. & Nanjundiah, V. 2003. Evolutionary questions raised by cellular slime mould development. *Proc. Indian Natn. Sci. Acad.* **B69**: 825–852.

Keller, L. & Surette, M.G. 2006. Communication in bacteria: an ecological and evolutionary perspective. *Nat. Rev. Microbiol.* **4**: 249–258.

Kiers, E.T. & van der Heijden, M.G.A. 2006. Mutualistic stability in the arbuscular mycorrhizal symbiosis: exploring hypotheses of evolutionary cooperation. *Ecology* **87**: 1627–1636.

Kiers, E.T., Rousseau, R.A., West, S.A. & Denison, R.F. 2003. Host sanctions and the legume-rhizobium mutualism. *Nature* **425**: 78–81.

Kiers, E.T., Rousseau, R.A. & Denison, R.F. 2006. Measured sanctions: legume hosts detect quantitative variation in rhizobium cooperation and punish accordingly. *Evol. Ecol. Res.* **8**: 1077–1086.

Kokko, H., Johnstone, R.A. & Clutton-Brock, T.H. 2001. The evolution of cooperative breeding through group augmentation. *Proc. R. Soc. Lond. B* **268**: 187–196.

Krebs, J.R. & Davies, N.B. 1993. *An Introduction to Behavioural Ecology*, 3rd edn. Blackwell Scientific Publications, Oxford, England.

Kreft, J.-U. 2005. Conflicts of interest in biofilms. *Biofilms* **1**: 265–276.

Kruuk, H. 2003. *Niko's Nature*. Oxford University Press, Oxford.

Lack, D. 1966. *Population Studies of Birds*. Clarendon Press, Oxford.

Lehmann, L. & Keller, L. 2006. The evolution of cooperation and altruism. A general framework and classification of models. *J. Evol. Biol.* **19**: 1365–1725.

Leigh, E. 1983. When does the good of the group override the advantage of the individual? *Proc. Natl. Acad. Sci. (USA)* **80**: 2985–2989.

MacColl, A.D.C. & Hatchwell, B.J. 2004. Determinants of lifetime fitness in a cooperative breeder, the long tailed tit *Aegithalos caudatus*. *J. Anim. Ecol.* **73**: 1137–1148.

Maynard Smith, J. 1964. Group selection and kin selection. *Nature* **201**: 1145–1147.

Maynard Smith, J. 1976. Group selection. *Q. Rev. Biol.* **51**: 277–283.

Maynard Smith, J. 1982. *Evolution and the Theory of Games*. Cambridge University Press, Cambridge.

Maynard Smith, J. 1983. Models of evolution. *Proc. Roy. Soc. Lond. B* **219**: 315–325.

Maynard Smith, J. & Harper, D. 2003. *Animal Signals*. Oxford University Press, Oxford.

Maynard Smith, J. & Szathmary, E. 1995. *The Major Transitions in Evolution*. W.H. Freeman, Oxford.

Mayr, E. 1961. Cause and effect in biology. *Science* **134**: 1501–1506.

Nowak, M.A. & May, R.M. 1992. Evolutionary games and spatial chaos. *Nature* **359**: 826–829.

Okasha, S. 2004. Multilevel selection and the partitioning of covariance: a comparison of three approaches. *Evolution* **58**: 486–494.

Okasha, S. 2005. Maynard Smith on the levels of selection question. *Biol. Philos.* **20**: 989–1010.

Parker, G.A. & Maynard Smith, J. 1990. Optimality theory in evolutionary biology. *Nature* **348**: 27–33.

Partridge, L. & Sibley, R. 1991. Constraints in the evolution of life histories. *Phil. Trans. Roy. Soc. Lon. B* **332**: 3–13.

Pellmyr, O. & Huth, C.J. 1994. Evolutionary stability of mutualism between yuccas and yucca moths. *Nature* **372**: 257–260.

Pepper, J.W. 2000. Relatedness in trait group models of social evolution. *J. Theor. Biol.* **206**: 355–368.

Queller, D.C. 1985. Kinship, reciprocity and synergism in the evolution of social behaviour. *Nature* **318**: 366–367.

Queller, D.C. 1992. Quantitative genetics, inclusive fitness, and group selection. *Am. Nat.* **139**: 540–558.

Queller, D.C. 2004. Kinship is relative. *Nature* **430**: 975–976.

Queller, D.C. & Goodnight, K.F. 1989. Estimating relatedness using genetic markers. *Evolution* **43**: 258–275.

Queller, D.C., Zacchi, F., Cervo, R., Turillazzi, S., Henshaw, M.T., Santorelli, L.A. & Strassmann, J.E. 2000. Unrelated helpers in a social insect. *Nature* **405**: 784–787.

Quervain, J.-F.d., Fischbacher, U., Treyer, V., Schellhammer, M., Schnyder, U., Buck, A. & Fehr, E. 2004. The neural basis of altruistic punishment. *Science* **305**: 1254–1258.

Ratnieks, F.L.W. 2006. The evolution of cooperation and altruism: the basic conditions are simple and well known. *J. Evol. Biol.* **19**: 1413–1414.

Ratnieks, F.L.W. & Visscher, P.K. 1989. Worker policing in the honeybee. *Nature* **342**: 796–797.

Ratnieks, F.L.W., Foster, K.R. & Wenseleers, T. 2006. Conflict resolution in insect societies. *Annu. Rev. Entomol.* **51**: 581–608.

Read, A.F. 1994. The evolution of virulence. *Tr. Microbiol.* **2**: 73–76.

Reeve, H.K. 1992. Queen activation of lazy workers in colonies of the eusocial naked mole-rat. *Nature* **358**: 147–167.

Reeve, H.K. & Keller, L. 1999. Levels of selection: burying the units-of-selection debate and unearthing the crucial new issues. In: *Levels of Selection in Evolution* (L. Keller, ed.), pp. 3–14. Princeton University Press, Princeton.

Ridley, M. 2004. *Evolution*, 3rd edn. Blackwell Publishing, Oxford.

Rose, M.R. & Lauder, G.V., eds. 1996. *Adaptation*. Academic Press, San Diego.

Rousset, F. 2004. *Genetic Structure and Selection in Subdivided Populations*. Princeton University Press, Princeton, NJ.

Sachs, J.L., Mueller, U.G., Wilcox, T.P. & Bull, J.J. 2004. The evolution of cooperation. *Q. Rev. Biol.* **79**: 135–160.

Shapiro, J.A. 1998. Thinking about bacterial populations as multicellular organisms. *Annu. Rev. Microbiol.* **52**: 81–104.

Shapiro, J.A. & Dworkin, M. 1997. *Bacteria as Multicellular Organisms*. Oxford University Press, New York.

Shuker, D.M. & West, S.A. 2004. Information constraints and the precision of adaptation: sex ratio manipulation in wasps. *Proc. Natl. Acad. Sci. USA* **101**: 10363–10367.

Shuker, D.M., Pen, I., Duncan, A.B., Reece, S.E. & West, S.A. 2005. Sex ratios under asymmetrical local mate competition: theory and a test with parasitoid wasps. *Am. Nat.* **166**: 301–316.

Sober, E. & Wilson, D.S. 1998. *Unto Others: The Evolution and Psychology of Unselfish Behavior*. Harvard University Press, Boston, MA.

Stevens, J.R. & Hauser, M.D. 2004. Why be nice? Psychological constraints on the evolution of cooperation. *Trends in Cognitive Sciences* **8**: 60–65.

Stevens, J.R., Cushman, F.A. & Hauser, M.D. 2005. Evolving the psychological mechanisms for cooperation. *Ann. Rev. Ecol. Evol. Syst.* **36**: 499–518.

Taylor, P.D. 1990. Allele-frequency change in a class structured population. *Am. Nat.* **135**: 95–106.

Taylor, P.D. 1992a. Inclusive fitness in a homogeneous environment. *Proc. R. Soc. Lond. B* **249**: 299–302.

Taylor, P.D. 1992b. Altruism in viscous populations – an inclusive fitness model. *Evol. Ecol.* **6**: 352–356.

Taylor, P.D. 1996. Inclusive fitness arguments in genetic models of behaviour. *J. Math. Biol.* **34**: 654–674.

Taylor, P.D. & Frank, S.A. 1996. How to make a kin selection model. *J. Theor. Biol.* **180**: 27–37.

Taylor, P.D., Wild, G. & Gardner, A. 2007. Direct fitness or inclusive fitness: how shall we model kin selection. *J. Evol. Biol.* **20**: 296–304.

Tinbergen, N. 1963. On aims and methods of ethology. *Z. Tierpsychol.* **20**: 410–433.

Travisano, M. & Velicer, G.J. 2004. Strategies of microbial cheater control. *Trends. Microbiol.* **12**: 72–78.

Trivers, R.L. 1971. The evolution of reciprocal altruism. *Q. Rev. Biol.* **46**: 35–57.

Trivers, R.L. 1974. Parent-offspring conflict. *Am. Zool.* **14**: 249–264.

Trivers, R.L. 1985. *Social Evolution*. Benjamin/Cummings, Menlo Park, CA.

Trivers, R.L. 1998. As they would do to you. *Skeptic* **6**: 81–83.

Velicer, G.J. 2003. Social strife in the microbial world. *Trends. Microbiol.* **11**: 330–337.

Wade, M.J. 1985. Soft selection, hard selection, kin selection, and group selection. *Am. Nat.* **125**: 61–73.

Wedekind, C. & Braithwaite, V.A. 2002. The long-term benefits of human generosity in indirect reciprocity. *Curr. Biol.* **12**: 1–4.

Wehner, R. 1987. 'Matched filters' – neural models of the external world. *J Comp Physiol A* **161**: 511–531.

Wenseleers, T., Helantera, H., Hart, A. & Ratnieks, F.L.W. 2004. Worker reproduction and policing in insect societies: an ESS analysis. *J. Evol. Biol.* **17**: 1035–1047.

West, S.A. & Buckling, A. 2003. Cooperation, virulence and siderophore production in bacterial parasites. *Proc. Roy. Soc. Lond. B* **270**: 37–44.

West, S.A. & Sheldon, B.C. 2002. Constraints in the evolution of sex ratio adjustment. *Science* **295**: 1685–1688.

West, S.A., Lively, C.M. & Read, A.F. 1999. A pluralist approach to the evolution of sex and recombination. *J. Evol. Biol.* **12**: 1003–1012.

West, S.A., Reece, S.E. & Read, A.F. 2001. The evolution of gametocyte sex ratios in malaria and related apicomplexan (protozoan) parasites. *Trends Parasitol.* **17**: 525–531.

West, S.A., Kiers, E.T., Pen, I. & Denison, R.F. 2002a. Sanctions and mutualism stability: when should less beneficial mutualists be tolerated? *J. Evol. Biol.* **15**: 830–837.

West, S.A., Kiers, E.T., Simms, E.L. & Denison, R.F. 2002b. Sanctions and mutualism stability: why do rhizobia fix nitrogen? *Proc. Roy. Soc. Lond. B* **269**: 685–694.

West, S.A., Reece, S.E. & Read, A.F. 2003. *Toxoplasma gondii*, sex and premature rejection. *Trends Parasitol.* **19**: 155–157.

West, S.A., Gardner, A., Shuker, D.M., Reynolds, T., Burton-Chellow, M., Sykes, E.M., Guinnee, M.A. & Griffin, A.S. 2006a. Cooperation and the scale of competition in humans. *Curr. Biol.* **16**: 1103–1106.

West, S.A., Griffin, A.S., Gardner, A. & Diggle, S.P. 2006b. Social evolution theory for microbes. *Nat. Rev. Microbiol.* **4**: 597–607.

Wiley, R.H. & Rabenold, K.N. 1984. The evolution of cooperative breeding by delayed reciprocity and queuing for favorable social positions. *Evolution* **38**: 609–621.

Williams, G.C. 1966. *Adaptation and Natural Selection*. Princeton University Press, Princeton.

Williams, G.C. 1992. *Natural Selection. Domains, Levels and Challenges*. Oxford University Press, Oxford.

Wilson, D.S. 1975a. A theory of group selection. *Proc. Natl. Acad. Sci. USA* **72**: 143–146.

Wilson, E.O. 1975b. *Sociobiology*. Harvard University Press, Cambridge, Massachusetts.

Wilson, D.S. 1977. Structured demes and the evolution of group advantageous traits. *Am. Nat.* **111**: 157–185.

Wilson, E.O. 2005. Kin selection as the key to altruism: its rise and fall. *Soc. Res.* **72**: 159–168.

Wilson, D.S. & Colwell, R.K. 1981. The evolution of sex ratio in structured demes. *Evolution* **35**: 882–897.

Wilson, E.O. & Hölldobler, B. 2005. Eusociality: origin and consequences. *Proc. Natl. Acad. Sci. USA* **102**: 13367–13371.

Wynne-Edwards, V.C. 1962. *Animal Dispersion in Relation to Social Behaviour*. Oliver and Boyd, Edinburgh.

Yu, D.W. 2001. Parasites of mutualisms. *Biol. J. Linn. Soc.* **72**: 529–546.

Zahavi, A. 2003. Indirect selection and individual selection in sociobiology: my personal views on theories of social behaviour. *Anim. Behav.* **65**: 859–863.

Zahavi, A. 2005. Is group selection necessary to explain social adaptations in microorganisms? *Heredity* **94**: 143–144.