What is altruism?

Benjamin Kerr[1], Peter Godfrey-Smith[2] and Marcus W. Feldman[3]

[1]Department of Ecology, Evolution and Behavior, University of Minnesota, 100 Ecology, 1987 Upper Buford Circle, St Paul, MN 55108, USA
[2]Philosophy Program RSSS, Australian National University, Canberra, ACT 0200, Australia
[3]Department of Biological Sciences, Stanford University, Stanford, CA 94305, USA

Altruism is generally understood to be behavior that benefits others at a personal cost to the behaving individual. However, within evolutionary biology, different authors have interpreted the concept of altruism differently, leading to dissimilar predictions about the evolution of altruistic behavior. Generally, different interpretations diverge on which party receives the benefit from altruism and on how the cost of altruism is assessed. Using a simple trait-group framework, we delineate the assumptions underlying different interpretations and show how they relate to one another. We feel that a thorough examination of the connections between interpretations not only reveals why different authors have arrived at disparate conclusions about altruism, but also illuminates the conditions that are likely to favor the evolution of altruism.

Sometimes, the pervasiveness of a scientific term gives one the impression that its meaning is unequivocal. Ironically, such ubiquity can go hand-in-hand with ambiguity, particularly when the term is borrowed from common parlance and 'lacks the precision, uniformity, and neutrality that scientific terms are supposed to have' [1]. Naturally, confusion arises when differences in the meaning of a key term are overlooked, often resulting in needless dispute and a failure to recognize underlying progress and consensus.

Within evolutionary biology, the term 'altruism' is an example of such ambiguity. A common definition describes altruism as behavior that simultaneously entails fitness costs to the behaving individual and fitness benefits to individuals on the receiving end of the behavior [2,3]. Several authors have provided insight into how ambiguity creeps into such a definition [1,4−7]. Here, we build on this work to delineate the primary differences among different interpretations of 'altruism' (Box 1). We then show how these different interpretations relate to one another. We do not advocate the use of one particular interpretation. Rather, we examine the evolutionary processes that are associated with each one.

To illustrate the potential for confusion, consider the following statements about altruism and fitness:

- An altruist can have a higher fitness than the non-altruists within its group.

- The fitness of an individual can decrease as the frequency of altruists in its group increases.
- Group productivity can decrease with increasing frequency of altruists.
- A non-altruist that 'switches' to altruism can reap a net fitness gain.

Each of these statements is true for at least one of the interpretations of altruism that we discuss here.

In this article, we use the trait-group framework to discuss the semantic issues surrounding altruism. In this framework, individuals form groups (the 'trait-groups') in which fitness-affecting interactions occur; after selection (which is a consequence of these interactions), the groups 'mix' (before the next generation of groups forms). The rationale for using this framework is its use in many of the models for the evolution of altruism [8−13]. Also, if genes are considered as 'individuals' and diploid organisms are taken to be 'groups of two genes', then classic population genetic theory becomes trait-group theory [7,14,15]. Many game theoretic examples also fit nicely into this framework [14,15]. The life cycle of a simple version of this framework is as follows:

(i) An infinite population of individuals of two types, A and B, (randomly or nonrandomly) forms an infinite number of groups of size n; (ii) Selection occurs, where the fitness of any focal individual (its probability of survival or its expected number of copies) is a function of the composition of its size n group. Here, we let α_i and β_i be the fitness of A and B, respectively, in a group with i A types and $n-i$ B types; (iii) The groups dissolve.

Reproduction in this model is asexual, and 'selection' in (ii) can be of several different kinds. For instance, there can be fertility selection if one type produces more copies of itself, or viability selection if fitness corresponds to the probability of survival (followed by nondifferential

Corresponding author: Benjamin Kerr (bkerr@stanford.edu).

reproduction). Kin selection is also a possibility if reproduction occurs in groups before selection does.

As several authors have pointed out, there are two pieces of information that are crucial for predicting the evolutionary trajectory of such a population [3,7,15,16]. First, we must know how groups form (i.e. we must know the frequency distribution of group types). This is step (i). Second, we must know the relationships between fitnesses of individuals (i.e. we must know the fitness structure). This becomes important in step (ii). The three interpretations of altruism that we explore here focus exclusively on the fitness structure rather than on the manner in which groups form.

The 'focal-complement' interpretation of altruism

Matessi and Karlin [10] define type A individuals as altruists when the following conditions hold (Eqn 1,2)

Condition 1 : $\alpha_{i+1} < \beta_i$, for all $i \in \{0, 1, 2, \ldots n - 1\}$ [Eqn 1]

Condition 2 : $(\alpha_{i+1} - \alpha_i)i + (\beta_{i+1} - \beta_i)(n - i - 1) > 0$,

for all $i \in \{0, 1, 2, \ldots n - 1\}$ [Eqn 2]

For condition 2, we define $\alpha_0 = \beta_n = 0$.

The idea behind conditions 1 and 2 is simple. Let us consider a non-altruist (type B) in a group with i altruists. Say this focal individual switches its type ($B \rightarrow A$). This 'convert' is now in a group with $i + 1$ altruists (the extra altruist being the convert). The change in fitness experienced by the convert is $\alpha_{i+1} - \beta_i$. Condition 1 shows that altruism is costly for such a convert; that is, $\alpha_{i+1} - \beta_i < 0$.

Now, let us consider what happens with the other $n - 1$ members of the group after the convert switches. These 'static' individuals form the complement of the focal convert. The i static altruists each experience a change of fitness of $\alpha_{i+1} - \alpha_i$, and the $n - i - 1$ static non-altruists each experience a change of fitness of $\beta_{i+1} - \beta_i$. Condition 2 says that the static members, taken as a block, experience a net benefit from the altruism of the convert; that is, the net change in static fitness is positive or $(\alpha_{i+1} - \alpha_i)i + (\beta_{i+1} - \beta_i)(n - i - 1) > 0$.

Thus, altruism is costly to self, but is a benefit to others. In other words, the cost of altruism is described exclusively in terms of changes in fitness for a focal individual and the benefit of altruism is described exclusively in terms of the changes in fitness of the complement of the focal individual within its group. A fitness structure satisfying conditions 1 and 2 describes 'focal-complement altruism' (F-C altruism) and we refer to A as an 'F-C altruist' in this case. F-C altruism is essentially a more general version of Hamilton's conception of altruism [17] and is consistent with other authors' interpretation of altruism [4,5,18].

The 'multilevel' interpretation of altruism

Matessi and Jayakar [9] use the following conditions to define type A as an altruist (Eqn 3,4)

Condition 3 : $\alpha_i < \beta_i$, for all $i \in \{1, 2, \ldots n - 1\}$ [Eqn 3]

Condition 4 : $[(i + 1)\alpha_{i+1} + (n - i - 1)\beta_{i+1}]$

$> [i\alpha_i + (n - i)\beta_i]$, for all $i \in \{0, 1, 2, \ldots n - 1\}$ [Eqn 4]

The cost of altruism is given by condition 3, which says that, within any mixed group, the altruist always has a lower fitness than does the non-altruist. The benefit of altruism is given by condition 4, which says that group productivity increases with the number of altruists in the group. For instance, assume that the fitness of an individual is its expected number of offspring. In a group with i altruists, the number of offspring of the entire group will be $i\alpha_i + (n - i)\beta_i$. Condition 4 says that the number of offspring for a group should increase as the fraction of altruists within the group increases.

This interpretation of altruism relies on fitness comparisons both within and between groups. Thus, a fitness structure that satisfies conditions 3 and 4 describes 'multilevel altruism' (ML altruism), and we call type A individuals 'ML altruists'. This interpretation is very close to what Haldane [19] had in mind for altruism. Other authors [8,20–22] also use this interpretation. Sober and Wilson [3] describe altruism as 'behavior that decreases relative fitness within groups, but increases the fitness of groups'. These two clauses can be translated immediately into conditions 3 and 4. In his discussion of group selection, Price [23] discussed behaviors that are 'group-benefiting' but not 'individual-benefiting' within the context of his famous covariance framework. If one labels such behavior as altruism, Price's discussion relates to ML altruism.

Condition 4 is related to condition 2 of F-C altruism. Specifically, we can rewrite condition 4 as follows (Eqn 5):

Condition 5 : $(\alpha_{i+1} - \alpha_i)i + (\beta_{i+1} - \beta_i)(n - i - 1)$

$+ (\alpha_{i+1} - \beta_i) > 0$, for all $i \in \{0, 1, 2, \ldots n - 1\}$.   [Eqn 5]

The left side of condition 5 is the same as the left side of condition 2, except that it has an extra term, which is the change in fitness to the 'convert' (i.e. $\alpha_{i+1} - \beta_i$). Thus, condition 5 requires that the altruism of a convert causes a positive net change in fitness for the whole group, including the convert. Thus, whereas F-C altruists must help others (condition 2) at a personal cost (condition 1), ML altruists can be helping themselves ($\alpha_{i+1} - \beta_i > 0$) whilst helping their group. Therefore, the benefit of altruistic behavior from a focal F-C altruist is shared in some way by its complement, whereas the benefit of ML altruism is shared in some way by members of the entire group, including the focal altruist.

The 'individual-centered' interpretation of altruism

Uyenoyama and Feldman [11] define altruism by the following three conditions (Eqn 6–8):

Condition 6 : $\alpha_i < \beta_i$, for all $i \in \{1, 2, 3, \ldots n - 1\}$ [Eqn 6]

Condition 7 : $\alpha_i < \alpha_{i+1}$, for all $i \in \{1, 2, 3, \ldots n - 1\}$ [Eqn 7]

Condition 8 : $\beta_i < \beta_{i+1}$, for all $i \in \{0, 1, 2, \ldots n - 2\}$ [Eqn 8]

Condition 6 is identical to condition 3. Thus, the cost of altruism is depicted as a lower fitness of the altruist within its group (as was the case for ML altruists). Conditions 7 and 8 state that the individual fitnesses of both A and B individuals increase with the number of A types in the

group. The benefit of altruism here is measured by the increase in the fitness of every (nonconverting) individual, because of the addition of more altruists to its group.

The focus in this interpretation is on individuals. Thus, we label this type of fitness structure as 'individual-centered altruism' (I-C altruism) and, given this structure, we call type A an 'I-C altruist'. There is no condition based on the fitness boost to the complement of a focal altruist as in condition 2 of F-C altruism, and there is no condition based on group productivity as in condition 4 of ML altruism.

The fundamental differences

If the conditions of one definition of altruism are sufficient to satisfy all the conditions of another definition of altruism, then the latter definition includes the former. As Table 1 shows, none of the three definitions of altruism includes, or is included by any other. The differences between the interpretations concern the manner in which the benefit and cost are assessed. Within all interpretations, the benefit of altruism is measured through some comparison of fitness in which the number of altruists in the group varies (conditions 2, 4, 7 and 8 all involve quantities with subscripts i and $i + 1$, where i denotes the number of altruists in the group). However, the benefiting party might be the individual (I-C altruism), the group complement of the altruist considered as a block (F-C altruism), or the whole group (ML altruism). The cost of altruism in the case of F-C altruism involves a comparison of individual fitness across groups. However, the cost of altruism in both ML and I-C altruism involves a comparison of individual fitness within groups.

Returning to the idea of a $B \rightarrow A$ convert, F-C altruism requires a decrease in fitness for the convert. However, because the cost to the altruist is given by a lower fitness within its group for ML or I-C altruists, in these cases, fitness need not decrease for a convert [1,20,21]; indeed, a ML or I-C convert to altruism can increase its fitness. This difference is crucial, because, under certain assumptions about how groups form, the evolutionary behavior of the system will be fully determined by the sign of the change in convert fitness.

A further distinction

F-C altruism is unique in assessing the cost of altruism by considering the change in individual fitness owing to a conversion to altruism. The fitness change of a convert can be used to delineate a further classification of fitness structures (Eqn 9,10)

Class I : $\alpha_{i+1} < \beta_i$, for all $i \in \{0, 1, 2, \ldots n - 1\}$ [Eqn 9]

Class II : $\alpha_{i+1} > \beta_i$, for all $i \in \{0, 1, 2, \ldots n - 1\}$ [Eqn 10]

Table 1. Sufficiency of conditions[a]

	C1	C2	C3	C4	C6	C7	C8
F-C altruism (C1 and C2)			No	No	No	No	No
ML altruism (C3 and C4)	No	Yes			Yes	No	No
I-C altruism (C6, C7 and C8)	No	Yes	Yes	No			

[a]C stands for 'condition'. The entries describe whether the conditions in the rows are sufficient for the condition in the column.

Thus, F-C altruism requires a Class I fitness structure, whereas both ML and I-C altruism can have a Class I structure, a Class II structure or neither (Box 2). This distinction becomes important when we consider how groups are formed [step (i) of the life-cycle]. When groups form at random, we have the following result [8,9,11,15]: if individuals form groups at random each generation, type B individuals will invade, fix and remain stable to invasion by type A individuals when the fitness structure is Class I; whereas, type A individuals will invade, fix and remain stable to invasion by type B individuals when the fitness structure is Class II. This we call the Matessi-Jayakar Proposition [9,15].

Connections between the interpretations

Here, we illustrate the relationships between the three interpretations of altruism. First, we state some general claims concerning sufficient conditions for one type of altruism to be another type (we offer a proof for the first claim; the other claims have similar proofs). Second, we give examples of fitness structures that satisfy and violate

---

Box 2. Different classes of altruism

We can represent the fitness structure (i.e. the relationships between fitnesses of individuals) within the simple trait-group framework on a single graph. In Figure I, any filled point (i, α_i) gives the fitness, α_i, of an altruist in a group with i altruists, whereas any unfilled point (i, β_i) gives the fitness, β_i, of a selfish type in a group with i altruists.

Here, we illustrate the fitness structure of the Prisoner's Dilemma (Figure Ia) with type A as a cooperator (altruist) and type B as a defector (selfish type) in trait-groups of size two. Payoff values are taken to be fitnesses [24]. Thus, a defector paired with another defector has fitness 1 [the unfilled point (0,1)]. Within a mixed pair, a cooperator has fitness 0 [the filled point (1,0)] and the defector has fitness 5 [the unfilled point (1,5)]. A cooperator paired with another cooperator has fitness 3 [the filled point (2,3)]. The Prisoner's Dilemma is an example of a Class I fitness structure. If the fitness for an altruist in a group with $i + 1$ altruists is compared with the fitness of a selfish type in a group with i altruists [where $i \in \{0,1\}$], the selfish type always has a higher fitness (i.e. $\alpha_{i+1} < \beta_i$). Figure Ib describes a different fitness structure, where if the same comparison between fitnesses is made, the altruist has the higher fitness (i.e. $\alpha_{i+1} > \beta_i$). This is an example of a Class II fitness structure.

To distinguish the structures, simply pick any A fitness and compare its y-value to that of the B fitness one x-value to the left. If the A fitness is always less than the B fitness in this comparison, then the fitness structure is Class I. If the A fitness always exceeds the B fitness in this comparison, then the fitness structure is Class II.

(a) **(b)**

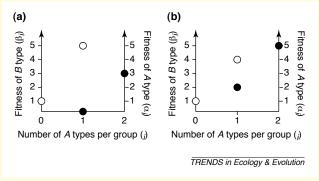*TRENDS in Ecology & Evolution*

Figure I.

different combinations of the three interpretations of altruism.

- Claim 1: ML altruism with a Class I structure is F-C altruism.

Proof: Under Class I structure, condition 1 holds. Condition 4 can be rearranged as follows (Eqn 11):

$$(\alpha_{i+1} - \alpha_i)i + (\beta_{i+1} - \beta_i)(n - i - 1) > \beta_i - \alpha_{i+1} \quad \text{[Eqn 11]}$$

for all $i \in \{0, 1, 2, \ldots n - 1\}$. However, because we have assumed a Class I structure, we know that $\beta_i > \alpha_{i+1}$ for all $i \in \{0, 1, 2, \ldots n - 1\}$. In such a case, Equation [11] guarantees condition 2.

- Claim 2: I-C altruism with a Class I structure is F-C altruism.
- Claim 3: ML altruism with a Class II structure is I-C altruism.
- Claim 4: I-C altruism with a Class II structure is ML altruism.

In claims 1–4, we show that, if the structure of ML altruism or I-C altruism is known to be either Class I or Class II, then we have sufficient conditions to satisfy another interpretation of altruism.

Box 3 discusses the space of fitness structures and the way in which different sets of fitness structures intersect. The following is demonstrated by examples in Box 3: first, fitness structures exist that satisfy the conditions for any one of the three interpretations, but which do not satisfy the conditions for the remaining two interpretations (e.g. ML altruists exist that are neither I-C altruists nor F-C altruists). Second, there are structures that satisfy the conditions for any two of the three interpretations, but not the conditions for the third (e.g. ML altruists exist that are I-C altruists, but not F-C altruists). Third, structures exist that satisfy the conditions for all three interpretations (e.g. the Prisoner's Dilemma).

With the aid of Figure I in Box 3, we can now return to the statements introduced at the beginning of this article and add the relevant qualifiers for each.

- An F-C altruist can have a higher fitness than the non-altruists within its group (yellow fitness structure: Box 3).
- The fitness of an individual can decrease as F-C or ML altruist frequency in its group increases (yellow, orange and red fitness structures: Box 3).
- Group productivity can decrease with increasing F-C or I-C altruist frequency (green and blue fitness structures: Box 3).
- A non-altruist that 'switches' to ML or I-C altruism can reap a net fitness gain (red, blue and purple fitness structures: Box 3).

One reason that the existence of some of these puzzles has not been realized is that there has been a focus on simple additive models (see [25] for an exception). For instance, in Wilson's [7] framework, each altruist donates r

units of fitness to every member of its group and has its own fitness changed by an amount d as a consequence, in which case (Eqn 12,13):

$$\alpha_i = z + d + r(i - 1) \quad \text{[Eqn 12]}$$

$$\beta_i = z + ri \quad \text{[Eqn 13]}$$

where z represents the 'base' fitness of an individual. We assume that $r > 0$, $d < r$ and $d + (n - 1)r > 0$. If $d \neq 0$, then a fitness structure must be either Class I (which Wilson calls 'strong altruism', where $d < 0$) or Class II (which Wilson calls 'weak altruism', where $0 < d < r$). If $d < 0$, then it is not difficult to show that conditions 1–8 hold, and thus Wilson's strong altruists are also F-C, ML, and I-C altruists. If $0 < d < r$, then conditions 3–8 hold and Wilson's weak altruists are ML and I-C altruists, but not F-C altruists. The additive fitness structure given by Equations 12 and 13 restricts the space of possible fitness structures to the white oval in the center of Figure I in Box 3 (the left half of the oval represents strong altruism and the right half represents weak altruism). Outside of this white oval, matters can get more complicated.

The evolutionary consequences of different interpretations

Given random group formation, the Matessi-Jayakar proposition predicts specific evolutionary behavior for Class I and Class II structures. ML altruism and I-C altruism can have a Class I structure, a Class II structure, or neither, whereas F-C altruism must have a Class I structure. Consequently, given random group formation, both ML altruists and I-C altruists can exhibit a range of dynamical behavior (including fixation, extinction, fixation or extinction depending on initial type frequencies and protected polymorphism), whereas F-C altruists are evolutionarily doomed.

Given the extinction of F-C altruists under random group formation, it is not surprising that the authors who employ this interpretation focus on mechanisms of nonrandom group formation [5,17]. Specifically, one way to avoid random group formation within the trait-group framework is for reproduction to occur within 'parent groups' of size n before selection operates. The resulting 'offspring groups' of size N (with $N > n$) will tend to have a more clumped distribution than if these size N groups had been assembled at random (i.e. A types will tend to be with other A types and B types will tend to be with other B types). This will be the case even if parents come together at random [12,13,20]. Some members of these offspring groups will be kin, and the relatedness among all members will determine whether F-C altruism can evolve. In this case, highly clumped group distributions (e.g owing to reproduction within small 'parent groups') will correspond to high coefficients of relatedness. And, as a result, kin selection is a potential mechanism for the evolution of F-C altruism, whereas it need not be invoked at all for the evolution of some cases of ML or I-C altruism.

Although the evolution of F-C altruism is certainly possible, there is another definition of altruism that leads to a theoretical dead-end. If one considers both the fitness structure (i.e. the α and β values) and the group frequency

Box 3. The space of fitness structures

Within our simple trait-group framework, groups have size n and α_i and β_i are the fitness of type A and type B, respectively, in a group with i A types and $n-i$ B types. We discuss three varieties of altruism (where type A is the altruist) (Eqn I–III):

• Focal-Compliment (F-C) altruism :

$$\alpha_{i+1} < \beta_i \text{ and } (\alpha_{i+1} - \alpha_i)i + (\beta_{i+1} - \beta_i)(n - i - 1) > 0, \qquad \text{[Eqn I]}$$

• Multilevel (ML) altruism :

$$\alpha_i < \beta_i \text{ and } [(i+1)\alpha_{i+1} + (n - i - 1)\beta_{i+1}] > [i\alpha_i(n - i)\beta_i], \qquad \text{[Eqn II]}$$

• Individual-Centered (I-C) altruism :

$$\alpha_i < \beta_i, \ \alpha_i < \alpha_{i+1} \text{ and } \beta_i < \beta_{i+1}. \qquad \text{[Eqn III]}$$
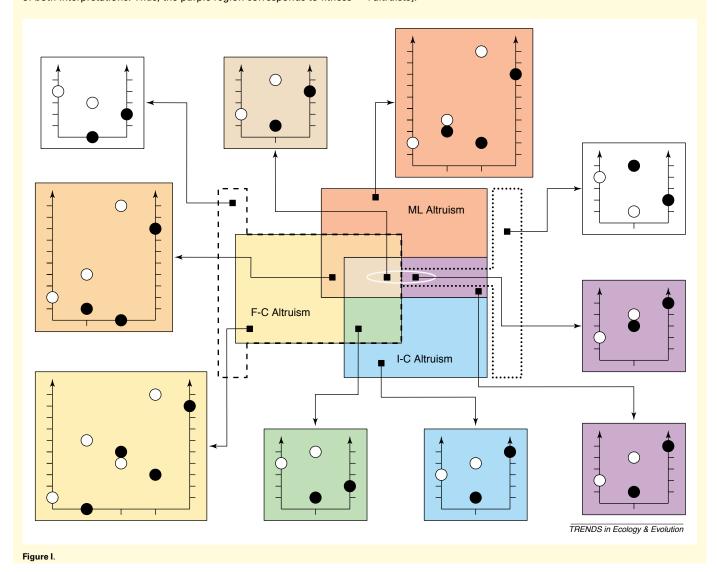
The above conditions hold over all well-defined values of i. In Figure I, the set of fitness structures that satisfy the conditions for F-C altruism, ML altruism, and I-C altruism are represented by the yellow, red, and blue rectangles, respectively. The overlap between any two rectangles corresponds to fitness structures that satisfy the conditions of both interpretations. Thus, the purple region corresponds to fitness structures showing ML altruism and I-C altruism, but not F-C altruism. The green region corresponds to fitness structures showing I-C altruism and F-C altruism, but not ML altruism. The orange region corresponds to fitness structures showing ML altruism and F-C altruism, but not I-C altruism. The brown region corresponds to fitness structures that satisfy the conditions for all three definitions of altruism (e.g. the Prisoner's Dilemma). On top of this rectangle schematic, we have represented regions satisfying the following conditions (over all well defined values of i) (Eqn IV,V)

$$\text{Class I}: \ \alpha_{i+1} < \beta_i \qquad \text{[Eqn IV]}$$

$$\text{Class II}: \ \alpha_{i+1} > \beta_i \qquad \text{[Eqn V]}$$

The region enclosed in the dashed boundary represents the structures that are Class I, whereas the region enclosed in the dotted boundary represents the structures that are Class II. (In general, all regions are drawn to emphasize overlap and their sizes are not drawn to scale). For every distinct region, we show an example of a potential fitness structure [filled points (i, α_i) and unfilled points (i, β_i) correspond to the fitness of an altruist, α_i, and a non-altruist, β_i, in a group with i altruists].

Figure I.

TRENDS in Ecology & Evolution

distribution (i.e. how different individuals come together to form groups), one can compute the marginal fitness of A and B (i.e. the fitness of a type averaged over group constitutions). If we think of an altruist as an individual who 'gives away fitness', it might initially seem reasonable to express this by saying 'an altruistic type must not have a higher marginal'. If so, altruism is an evolutionary impossibility in the trait-group model we describe. We

call this 'marginal altruism'. The requirement for marginal altruism depends not only on the fitness structure, but also on the group frequency distribution. Wilson and others [1,3,21,26] are extremely critical of this interpretation, because it entails that altruism is a kind of behavior that can never evolve.

We make one final point about marginal altruism. Imagine an F-C altruist in a population where groups form randomly. This F-C altruist would also be a marginal altruist, and would decline in frequency (the Matessi-Jayakar proposition). However, if the group frequency distribution suddenly became clumped (perhaps owing to reproduction within groups before selection), this F-C altruist might become a marginal selfish type! In the three interpretations discussed here, the identity of the altruist depends only on fitness relations. However, marginal altruism requires more information: the identity of the marginal altruist can change even if the group-dependent fitnesses of the individual types (i.e. the α and β values) remain constant. Thus, if the label 'altruist' is meant to apply to one type in relation to another only because of relationships between their group-dependent fitnesses, marginal altruism is indeed a poor choice.

Conclusions

To some extent, the different definitions of altruism discussed here correspond to different ways of thinking about the relevant 'units' in discussions of behavior and evolution. We see this especially on the benefit side of the definitions. Do whole groups function as evolutionary units to which benefits can be assigned, or are individuals the only beneficiaries? The definitions discussed here also differ in which kinds of evolutionary processes are required for altruism to evolve. Some conceptions of altruism are very demanding on this point; F-C altruism cannot survive with random group formation, for instance. Other conceptions are less demanding. For example, verbal formulations of I-C and ML altruism sometimes make the evolution of altruism look rather implausible. But both these categories include cases in which a convert to altruism increases its individual fitness. So the plausibility of evolutionary hypotheses concerning altruism will be affected by the exact definition of altruism used.

Although we embedded our discussion in the framework of evolutionary biology, similar issues arise in economics, social psychology and anthropology [27–29]. These fields also encounter subtly different kinds of cooperation and defection 'in collective goods games' and 'social dilemma' situations. In biology and other fields, we believe that careful consideration of the relations between definitions can prevent unproductive dispute and redirect focus on the true object of study – the conditions that are important for the evolution of different kinds of social behavior.

References
1 Wilson, D.S. and Dugatkin, L.A. (1992) Altruism: contemporary debates. In *Keywords in Evolutionary Biology* (Keller, E.F. and Lloyd, E.A., eds), pp. 29–33, Harvard University Press
2 Bourke, A.F.G. and Franks, N.F. (1995) *Social Evolution in Ants*, Princeton University Press
3 Sober, E. and Wilson, D.S. (1998) *Unto Others: The Evolution and Psychology of Unselfish Behavior*, Harvard University Press
4 Grafen, A. (1984) Natural selection, kin selection and group selection. In *Behavioral Ecology*, 2nd edn, (Krebs, J.R. and Davies, N.B., eds), pp. 62–91, Sinauer
5 Nunney, L. (1985) Group selection, altruism, and structured-deme models. *Am. Nat.* 126, 212–230
6 Uyenoyama, M.K. and Feldman, M.W. (1992) Altruism: some theoretical ambiguities. In *Keywords in Evolutionary Biology* (Keller, E.F. and Lloyd, E.A., eds), pp. 34–40, Harvard University Press
7 Wilson, D.S. (1990) Weak altruism, strong group selection. *Oikos* 59, 135–140
8 Cohen, D. and Eshel, I. (1976) On the founder effect and the evolution of altruistic traits. *Theor. Popul. Biol.* 10, 276–302
9 Matessi, C. and Jayakar, S.D. (1976) Conditions for the evolution of altruism under Darwinian selection. *Theor. Popul. Biol.* 9, 360–387
10 Matessi, C. and Karlin, S. (1984) On the evolution of altruism by kin selection. *Proc. Natl. Acad. Sci. U. S. A.* 81, 1754–1758
11 Uyenoyama, M. and Feldman, M.W. (1980) Theories of kin and group selection: a population genetics perspective. *Theor. Popul. Biol.* 17, 380–414
12 Wilson, D.S. (1975) A theory of group selection. *Proc. Natl. Acad. Sci. U. S. A.* 72, 143–146
13 Wilson, D.S. (1977) Structured demes and the evolution of group-advantageous traits. *Am. Nat.* 111, 157–185
14 Maynard Smith, J. (1987) How to model evolution. In *The Latest on the Best: Essays on Evolution and Optimality* (Dupre, J., ed.), pp. 119–131, MIT Press
15 Kerr, B. and Godfrey-Smith, P. (2002) Individualist and multi-level perspectives on selection in structured populations. *Biol. Philos.* 17, 477–517
16 Dugatkin, L.A. and Reeve, H.K. (1994) Behavioral ecology and levels of selection: dissolving the group selection controversy. *Adv. Study Behav.* 23, 101–133
17 Hamilton, W.D. (1975) Innate social aptitudes in man: an approach from evolutionary genetics. In *Biosocial Anthropology* (Fox, R., ed.), pp. 133–155, John Wiley & Sons
18 Karlin, S. and Matessi, C. (1983) Kin selection and altruism. *Proc. R. Soc. Lond. Ser. B* 219, 327–353
19 Haldane, J.B.S. (1932) *The Causes of Evolution*, Longsman & Green
20 Wilson, D.S. (1980) *The Natural Selection of Populations and Communities*, Benjamin/Cummings
21 Wilson, D.S. (1983) The group selection controversy: history and current status. *Annu. Rev. Ecol. Syst.* 14, 159–187
22 Maynard Smith, J. (1964) Kin selection and group selection. *Nature* 201, 1145–1147
23 Price, G.R. (1972) Extension of covariance selection mathematics. *Ann. Hum. Genet.* 35, 485–490
24 Axelrod, R. (1984) *The Evolution of Cooperation*, Basic Books
25 Cavalli-Sforza, L. and Feldman, M.W. (1978) Darwinian selection and 'altruism'. *Theor. Popul. Biol.* 14, 268–280
26 Sober, E. (1984) *The Nature of Selection: Evolutionary Theory in Philosophical Focus*, University of Chicago Press
27 Boyd, R. and Richerson, P.J. (1985) *Culture and the Evolutionary Process*, University of Chicago Press
28 Hauert, C. *et al.* (2002) Volunteering as red queen mechanism for cooperation in public goods games. *Science* 296, 1129–1132
29 Orbell, J. and Dawes, R. (1981) Social dilemmas. In *Progress in Applied Social Psychology* (Vol. 1) (Stephenson, G.M. and Davis, J.M., eds), pp. 37–65, John Wiley & Sons