Biology and Philosophy **17:** 477–517, 2002.

Target article

Individualist and Multi-level Perspectives on Selection in Structured Populations

BENJAMIN KERR*
Department of Biological Sciences
Stanford University
Stanford, CA 94305
USA
E-mail: bkerr@leland.stanford.edu

PETER GODFREY-SMITH
Philosophy Department
Stanford University
Stanford, CA 94305-2155
USA
E-mail: pgsmith@csli.stanford.edu

Abstract. Recent years have seen a renewed debate over the importance of group selection, especially as it relates to the evolution of altruism. One feature of this debate has been disagreement over which kinds of processes should be described in terms of selection at multiple levels, within and between groups. Adapting some earlier discussions, we present a mathematical framework that can be used to explore the exact relationships between evolutionary models that do, and those that do not, explicitly recognize biological groups as fitness-bearing entities. We show a fundamental set of mathematical equivalences between these two kinds of models, one of which applies a form of multi-level selection theory and the other being a form of "individualism." However, we also argue that each type of model can have heuristic advantages over the other. Indeed, it can be positively useful to engage in a kind of back-and-forth switching between two different perspectives on the evolutionary role of groups. So the position we defend is a "gestalt-switching pluralism."

Key words: altruism, game theory, group selection, individual selection, multi-level selection, pluralism, population genetics, trait groups

1. Introduction

Few concepts in biology possess a richer and more troubled history than the concept of "group selection." Williams' book *Adaptation and Natural Selec-*

---

* Current address: Ecology, Evolution and Behavior Department, University of Minnesota, 100 Ecology, 1987 Buford Circle, St. Paul, MN 55108, USA

tion (1966) served to banish the concept from the biological community in the late 1960's. However, a number of workers continued exploring the idea (Maynard Smith 1976; Wade 1978; Alexander and Borgia 1978; Uyenoyama and Feldman 1980; Wilson 1980, 1983; Nunney 1985). Largely due to D.S. Wilson's assiduous defense, group selection has made a dramatic comeback and is once again the subject of intense biological and philosophical exploration.

If group selection was an engine for controversy over the past 35 years, the subject of altruism was the fuel. The evolution of unselfish behavior was the topic of several early population genetic models that explicitly incorporated group structure (Wright 1945; Williams and Williams 1957; Maynard Smith 1964; Wilson 1975, 1977). Several comprehensive mathematical treatments of altruism followed (Eshel 1972; Cohen and Eshel 1976; Matessi and Jayakar 1976; Uyenoyama and Feldman 1980; Karlin and Matessi 1983; Matessi and Karlin 1984). Most recently, the evolution of altruism has been cast as one aspect of the larger topic of the evolution of higher levels of organization in nature, a topic encompassing the evolution of gene network formation, multicellularity, symbiosis, and cooperating societies (Buss 1987; Maynard Smith and Szathmáry 1995; Michod 1999).

Sober and Wilson's recent book *Unto Others* (1998) puts together a detailed and systematic position on these topics. *Unto Others* (hereafter abbreviated *UO*) defends a group-selectionist approach to the evolution of altruism, and also embeds this defense within a more general conceptual framework for understanding evolution. That framework is *multi-level selection theory*, or more loosely, the *multi-level perspective*. Multi-level selection theory recognizes a plurality of levels of organization in nature, and recognizes selective forces operating at different levels. The selective forces found at different levels can be opposed to each other. Indeed, for multi-level selection theory, the key to understanding the evolution of altruism lies in recognizing a clash between *within*-group selection for selfish behavior and *between*-group selection for altruism.

Others have argued that the key models of the evolution of altruism can also be understood in a second, more individualist way. Rather than recognizing the group as a unit on which selection acts, we can treat the group in which an individual finds itself as part of the individual's *environment*. We can understand altruism while only treating individual organisms as the bearers of fitness-like properties, so long as we pay proper attention to the social context in which individual behavior is produced. Dugatkin and Reeve (1994) and Sterelny (1996) call this perspective *broad-sense individualism*.

Dugatkin, Reeve and Sterelny do not argue that the evolution of altruism *must* be modeled in this individualist way. Rather, they argue for a form of pluralism. They hold that both multi-level selection theory and the individualist approach are acceptable.

Sober and Wilson agree that it is possible to devise an evolutionary model of altruism that gets the mathematical results right, while only assigning fitnesses to individuals. However, Sober and Wilson insist that only a multi-level description of these systems can yield real understanding of the evolution of altruism, because only a multi-level description is faithful to the causal structure involved. So although Sober and Wilson also call their view "pluralist," this is not the same kind of pluralism seen in Dugatkin, Reeve and Sterelny. The individualist perspective, for Sober and Wilson, is only predictively, and not explanatorily, adequate.[1]

In the present paper, we outline a mathematical framework that can be used for a more detailed investigation of this set of problems involving multi-level selection theory, individualism, pluralism and altruism. This framework is based on several earlier discussions (Price 1970,1972; Wilson 1975; Cohen and Eshel 1976; Matessi and Jayakar 1976; Uyenoyama and Feldman 1980; Matessi and Karlin 1984). We will use this framework to explore the exact relationships between models that do, and models that do not, explicitly recognize biological groups as the bearers of fitness-like properties. We will show that a properly-formulated version of the individualist perspective is mathematically equivalent to the multi-level perspective. We will also argue that the individualist perspective facilitates understanding of important conditions bearing on the evolution of altruism. We conclude by developing a number of examples, including some two-player games, the population genetics of meiotic drive, and Sober and Wilson's fascinating case of the "brain worm" parasite.

The position we will argue for, using this mathematical discussion, is a form of pluralism that is fairly close to the positions of Dugatkin, Reeve and Sterelny. But the sense in which we advocate a plurality is slightly different from theirs. They argue that while the two perspectives are interchangeable, each view has its own set of heuristic advantages. So, in a particular case, a researcher will often have good reason to choose one perspective over the other. We accept this, but add that even when thinking about a particular case, there are advantages that accrue from using *both* perspectives. More generally, we argue that a good understanding of altruism, and the evolution of higher levels of organization in nature, is fostered by the ability to make repeated "gestalt-switches" between the different perspectives. That is, we should cultivate the ability to switch between seeing groups as fitness-bearing units in their own right and seeing groups as part of the milieu that determines the fitnesses of individuals.[2] Each perspective makes some facts vivid and obscures others.

Given the intense history of the debates over group selection and altruism, we pause here to spell out some things that are not part of our argument. First, our aim is not to banish or discredit group selection or the multi-level perspective. We agree with Sober and Wilson that attempts to remove group selection from serious consideration in evolutionary biology were mistaken (and we refer the reader to *UO* for an enthralling account of the history of this debate). In much of our paper we do stress the heuristic uses of the individualist perspective, but this is not meant to deny the value of the multi-level perspective. Second, in this paper we do not attempt to assess the idea that group selection is a *causal force*, as Sober and Wilson have it. We approach the problem from a slightly different point of view. Rather than asking "What is group selection and when does it happen?" we ask "What is the relation between models that explicitly assign fitness-like parameters to groups, and models that only assign fitness-like parameters to individuals?" What is gained and lost by taking one perspective over another?

In addition, we do not argue for the general superiority of lower-level descriptions of complex phenomena. Lastly, although our overall view in this area is a kind of pluralism, we are not committed to any view about the general superiority of pluralist outlooks. The value of pluralism clearly depends on the particular plurality involved! We cautiously suggest that gestalt-switching of the kind we describe might be generally useful in a range of different fields, but in this paper we only argue for gestalt-switching in the context of a particular area of evolutionary biology.

2. Two Perspectives

Consider an infinite population of individual organisms, which are of two types: ***A*** and ***B***. Each generation, the individuals gather to form an infinite number of social groups, all of size n. Group formation may be random or non-random. Individuals reproduce asexually, by a form of simple copying, and an individual's fitness (the number of copies it can expect) depends both on its own type and on the composition of the group in which it finds itself. Groups then dissolve, the parent individuals expire, and the process begins again with the offspring individuals forming the next set of groups.[3]

If we wish to model a situation like this, we have two ways of treating the role of groups. As sketched in the previous section, we can assign fitnesses only to individuals, and treat groups as part of the context that affects individual fitness, or we can collect the output of all the individuals within a group to define a group-level productivity, or group-level fitness. If we take the latter approach, we also have to deal with the fact that different types within a group may differ in their share of the group's output. Thus we will need parameters to describe both the group-level productivity as a whole, and the sharing of that productivity across the individual types within the group.

These two approaches to modeling the system correspond to broad-sense individualism and multi-level selection theory, respectively. Although we have introduced these perspectives by talking about individual organisms and social groups, the distinction between these views is more general. Instead of thinking of the lower-level entity as an individual organism, we might think of it as a *gene*, with ***A*** and ***B*** as two alleles at the relevant locus. For diploid organisms, a "gene group" of size $n = 2$ is formed every time gametes fuse. Here again, we can choose to model evolution in this system in two ways. We can assign fitnesses to the lower-level only, by describing genotype-dependent fitnesses of alleles. Or, following a more standard practice within population genetics, we can assign fitnesses to the groups themselves (i.e., the genotypes) and track how this group fitness is "split" between the gene pair (i.e., as segregation occurs).

So the distinction between broad-sense individualism and multi-level selection theory is one instance of a more general distinction between two ways of dealing with group structure in a population. Speaking more abstractly, we have a system with two tiers: lower-level "particles" and higher-level collections of these particles.[4]

One approach assigns productivities or fitnesses only to the lower-level particles, and treats the groups as part of the selective environment experienced by the particles. We call this approach the *contextual* approach, as individual particles are assigned fitnesses in a way sensitive to group context. A second approach involves assigning productivities or fitnesses to the collection itself, while retaining information about change in particle composition within the collection. We call this approach the *collective* approach, because it recognizes collectives of particles as fitness-bearing entities in their own right. The collective approach addresses the selective process at two levels – through consideration of productivities *of* groups, and the change *within* groups. The distinction between contextual and collective approaches is a generalization of the distinction between broad-sense individualism[5] and multi-level selection theory.[6]

Above we described a very simple population and life cycle. But this set-up is a useful one for exploring a range of questions about different perspectives. Let us now see how this system will be *parameterized* according to the contextual and collective approaches. For the remainder of this section, we focus on the case where particles are individual organisms and the higher-level collections are social groups. The contextual approach assigns fitness-like properties only to individuals. Such an approach might use the following parameterization:

$$\alpha_i = \begin{cases} \text{the expected number of copies of an } \boldsymbol{A} \text{ type} \\ \text{in a group with } i\,\boldsymbol{A} \text{ types and } (n-i)\,\boldsymbol{B} \text{ types} \end{cases},$$

$$\beta_i = \begin{cases} \text{the expected number of copies of a } \boldsymbol{B} \text{ type} \\ \text{in a group with } i\,\boldsymbol{A} \text{ types and } (n-i)\,\boldsymbol{B} \text{ types} \end{cases}.$$

Note that α_0 and β_n are undefined. As fitnesses are here assigned only to individuals, but are done so in a way sensitive to group context, this paramaterization expresses broad-sense individualism of the kind described by Dugatkin, Reeve and Sterelny.[7]

Let us now look at how this system might be parameterized according to a collective approach to group structure. Groups are recognized as fitness-bearing entities in their own right. Groups differ in productivity, but in addition, the group's overall productivity may not be shared equitably among group members. Consider the following parameters:

$$\pi_i = \text{total number of copies from a group with } i\,\boldsymbol{A} \text{ types},$$

$$\phi_i = \frac{\text{number of } \boldsymbol{A} \text{ copies in a group with } i\,\boldsymbol{A} \text{ types}}{\text{total number of copies in a group with } i\,\boldsymbol{A} \text{ types}}.$$

Thus, π_i represents group productivity,[8] and ϕ_i the fraction of this productivity claimed by ***A*** types, in a group with *i* ***A*** types.[9] Differences between groups are assessed by comparing π values. Within group change is assessed by the difference between ϕ_i and *i/n* (i.e., ***A***'s frequency in the group after and before the selective episode). So this choice of parameters yields a *multi-level* description of selection, of the kind envisaged by Sober and Wilson in *UO*.

These two parameterizations seem to represent quite different approaches to the role of group structure in the population, but the two parameterizations are in fact mathematically equivalent. Each set of parameters can be defined in terms of the other. For instance, we can express the parameters π_i and ϕ_i in terms of α_i and β_i:

$$\pi_i = i\alpha_i + (n-i)\beta_i, \tag{1}$$

$$\phi_i = \frac{i\alpha_i}{i\alpha_i + (n-i)\beta_i}. \qquad (2)^{10}$$

Solving (1) and (2) for α_i and β_i gives

$$\alpha_i = \frac{\phi_i \pi_i}{i}, \tag{3}$$

$$\beta_i = \frac{(1-\phi_i)\pi_i}{n-i}. \tag{4}$$

Thus, given a collection of α's and β's, it is a triviality to produce a collection of π's and ϕ's and *vice versa*. Thus, there is a mathematical interchangeability between the individualist framework and the multi-level framework. Equations (1)–(4) are the key to translating between contextual and collective languages – a sort of Rosetta stone for fitness structure parameterizations.

Up to this point, we have focused on the fitness structure, which describes differences in productivity *after groups have formed*. This description does not yet yield evolutionary outcomes, however. In order to make evolutionary predictions, we must also know the frequencies of different types of groups. Consider the population at generation t, where $\overline{p}(t)$ and $\overline{q}(t)$ are the population-wide frequencies of $\boldsymbol{A}$ and $\boldsymbol{B}$, respectively. Let $f(t) \equiv \{f_0(t), f_1(t), f_2(t), f_3(t), \ldots f_n(t)\}$ be the group frequency distribution for generation t, where $f_i(t)$ is the frequency of groups with $i\,\boldsymbol{A}$ types. Given $f(t)$ and the fitness structure we can deduce the frequency of types in the next generation.

For instance, within the individualist perspective, the frequencies of types in the next generation are given by:

$$\overline{w}\,\overline{p}(t+1) = \sum_{i=1}^{n} f_i(t)\frac{i}{n}\alpha_i, \tag{5a}$$

$$\overline{w}\,\overline{q}(t+1) = \sum_{i=0}^{n-1} f_i(t)\frac{(n-i)}{n}\beta_i, \tag{6a}$$

with mean individual fitness given by

$$\overline{w} = \sum_{i=1}^{n} f_i(t)\frac{i}{n}\alpha_i + \sum_{i=0}^{n-1} f_i(t)\frac{(n-i)}{n}\beta_i.$$

Using relations (3) and (4), equations (5a) and (6a) can be rewritten in terms of π_i and ϕ_i

$$\overline{\pi}\,\overline{p}(t+1) = \sum_{i=1}^{n} f_i(t)\phi_i\pi_i, \tag{5b}$$[11]

$$\overline{\pi}\,\overline{q}(t+1) = \sum_{i=0}^{n-1} f_i(t)(1-\phi_i)\pi_i, \tag{6b}$$

with

$$\overline{\pi} = \sum_{i=0}^{n} f_i(t)\pi_i.$$

where $\overline{\pi}$ is identical to $n\overline{w}$ from equations (5a) and (6a). Note that if we denote the mean individual fitness within a group with i **A** types as w_i, then we have $\pi_i = n\left\{\frac{i}{n}\alpha_i + \frac{(n-i)}{n}\beta_i\right\} = nw_i$. There is an example in Dugatkin and Reeve (1994) using precisely this multi-level framework (see also Wade 1985).

In this paper, we are assuming the α's and β's (or π's and ϕ's) are *constant* values. Thus, given the fitness structure and the current group frequency distribution ($f(t)$) we can calculate **A**'s frequency after selection (i.e., $\overline{p}(t+1)$) using equations (5) and (6). However, this is not enough information to determine **A**'s frequency after a further round of selection (i.e., $\overline{p}(t+2)$). In order to compute $\overline{p}(t+2)$, we must have $f(t+1)$ as well. Simply, we must always know how groups are formed in order to calculate the frequencies of types after selection. Thus, equations (5) and (6) alone do not yield a dynamically sufficient model of the system (Lewontin 1969; Godfrey-Smith and Lewontin 1993). Dynamic sufficiency is achieved if we can express the distribution of groups, $f(t)$, as a function of type frequency, $\overline{p}(t)$. For instance, if groups are formed randomly, then the frequencies of group types follow a binomial distribution:

$$f_i(t) = \binom{n}{i}(\overline{p}(t))^i(1-\overline{p}(t))^{n-i}. \tag{7}$$

More generally, consider equations $f_i(t) = F_i(\overline{p}(t))$, such that $\sum_{i=0}^{n} F_i(\overline{p}(t)) = 1$. The F functions specify the way in which groups assemble given the current frequency of **A** types in the population. If a constant fitness structure and constant rules of group assembly (F functions) are available, then only the initial value of **A**'s frequency, $\overline{p}(0)$, is needed to determine the *entire* trajectory of the type frequencies as the population changes over time.[12] In other words, given F functions, equations (5) and (6) become recursions, where, at any time t, $\overline{p}(t+1)$ can be fully determined from $\overline{p}(t)$.

To summarize, we now specify the two essential components in understanding selection in our simple structured system.[13] First, we need the fitness structure, which can be either represented by a list of α's and β's, or by a list of π's and ϕ's. These two parameterizations of the fitness structure in the system contain exactly the same information, parceled up differently. Second, we need the distribution of types into groups (the $f(t)$ distibution). When we can represent $f(t)$ in terms of type frequencies, dynamic sufficiency is obtained.[14]

3. Fitness Structures and Altruism

Much of the dialogue that has taken place over selection in structured populations has focused on altruism. In this section we will argue that a proper understanding of the relations between individualist and multi-level perspectives clarifies a number of important issues concerning the evolution of altruism. We will do this by first discussing a few definitions of altruism, and their relations to the two perspectives. We then introduce a distinction between fitness structures that cross-cuts most definitions of altruism, and which can be used to make definite evolutionary predictions in some circumstances. This distinction is most easily represented and employed within the individualist perspective. We use that fact to introduce the idea that some key features of the problem of altruism are more readily seen from an individualist perspective.

What is altruism? Perhaps surprisingly, the literature contains a number of non-equivalent definitions. In a forthcoming paper we discuss these definitions in more detail; here we focus on issues that concern the relation between the multi-level and individualist perspectives.

A common idea in many definitions of altruism is that if ***A*** is an altruist, then ***A*** must have a lower fitness than ***B*** within any mixed group. Within the individualist perspective this can be expressed as follows:

$$\alpha_i < \beta_i, \text{ for all } i \in \{1, 2, 3, \ldots n - 1\}. \tag{8a}$$

Using relations (3) and (4), this condition can be re-expressed within the multi-level perspective as follows:

$$\phi_i < i/n, \text{ for all } i \in \{1, 2, 3, \ldots n - 1\}. \tag{8b}$$

Condition (8b) states that ***A***'s frequency within any group decreases over the selective episode.

Condition (8) is not sufficient for altruism (and for some definitions of altruism it is not even necessary; e.g., Matessi and Karlin 1984). This condition gauges the "cost" of altruism, but not its "benefit." There is a variety of natural-looking ways to introduce a further condition giving the benefit, and they are not all equivalent. In Sober and Wilson's discussion in *UO*, altruism is "behavior that decreases relative fitness within groups, but increases the fitness of groups" (p. 99). The first condition is covered by condition (8). One way to represent the second condition is as follows:

$$\pi_{i+1} > \pi_i, \text{ for all } i \in \{0, 1, 2, \ldots n - 1\}. \tag{9}$$

Condition (9) states that group fitness (productivity) increases as the number of ***A*** types in the group increase. Note that (9) could be written within the individualist perspective (see Table 1), but the resulting relation is far more cumbersome. Conditions (8) and (9) are necessary and sufficient conditions to fulfill Sober and Wilson's definition of altruism. Conditions (8) and (9) are also identical to definitions of altruism given by other authors (Matessi and Jayakar (1976); Cohen and Eshel (1976)).

There is an alternative way to conceive of the "benefit" side of altruism. Instead of using the idea of group benefit, we can require that the fitness of both ***A*** and ***B*** be an increasing function of the number of ***A*** types in an individual's group:

$$\alpha_i < \alpha_{i+1}, \text{for all } i \in \{1, 2, 3, \ldots n-1\}, \tag{10}$$

$$\beta_i < \beta_{i+1}, \text{for all } i \in \{0, 1, 2, \ldots n-2\}. \tag{11}$$

No matter what type you are, you do better if there are more type ***A*** individuals around you. This is a simple way to think about altruism in terms of individual fitness. Note that we could, of course, express (10) and (11) from the multi-level perspective (see Table 1), but the resulting relations are again more cumbersome. Other authors use a definition of altruism given by conditions (8), (10) and (11), including Uyenoyama and Feldman (1980, 1992).

These alternative approaches to defining altruism (conditions (8) and (9) and conditions (8), (10) and (11)) are not equivalent, and neither set implies the other. As Table 1 shows, Sober and Wilson's definition is easiest to use from within a multi-level parameterization of the fitness structure. Uyenoyama and Feldman's definition, on the other hand, is more transparent when using the individualist perspective. So although both definitions can be described within either perspective, each definition is easier to represent and think about from within one of the two perspectives. Almost all of the numerical examples of altruism given later in this paper qualify as cases of altruism by both definitions.[15] But for the sake of simplicity, when the term "altruism" appears in the text to follow, it refers to conditions (8) and (9).

Let us now turn to a different distinction between fitness structures. As we will see in the discussion to follow, this distinction has direct relevance to the prediction of evolutionary trajectories. Consider the following two conditions:

$$\textbf{Class I}: \quad \alpha_{i+1} < \beta_i, \text{for all } i \in \{0, 1, 2, \ldots n-1\}, \tag{12}$$

$$\textbf{Class II}: \quad \alpha_{i+1} > \beta_i, \text{for all } i \in \{0, 1, 2, \ldots n-1\}. \tag{13}$$

As Table 1 shows, relations (12) and (13) can again be rewritten in the multi-level perspective, but the resulting inequalities are less tidy.

Table 1.

	Individualist Perspective	Multi-Level Perspective
Definition of Fitness Structure	α_i = number of copies for an ***A*** type in group with i ***A*** types β_i = number of copies for a ***B*** type in group with i ***A*** types	π_i = total number of copies for a group with i ***A*** types $\phi_i = \frac{\text{number of } \boldsymbol{A} \text{ copies in a group with } i\,\boldsymbol{A}\text{'s}}{\text{total number of copies in a group with } i\,\boldsymbol{A}\text{'s}}$
Relation to Other Perspective	$\alpha_i = \frac{\phi_i \pi_i}{i}$ $\beta_i = \frac{(1-\phi_i)\pi_i}{n-i}$	$\pi_i = i\alpha_i + (n-i)\beta_i$ $\phi_i = \frac{i\alpha_i}{i\alpha_i + (n-i)\beta_i}$
Sober and Wilson's Conditions for Altruism	$\alpha_i < \beta_i$ $(i+1)\alpha_{i+1} + (n-i-1)\beta_{i+1} > i\alpha_i + (n-i)\beta_i$	$\phi_i < i/n$ $\pi_{i+1} > \pi_i$
Uyenoyama and Feldman's Conditions for Altruism	$\alpha_i < \beta_i$ $\alpha_i < \alpha_{i+1}$ $\beta_i < \beta_{i+1}$	$\phi_i < i/n$ $\frac{\phi_i \pi_i}{i} < \frac{\phi_{i+1}\pi_{i+1}}{i+1}$ $\frac{(1-\phi_i)\pi_i}{n-i} < \frac{(1-\phi_{i+1})\pi_{i+1}}{n-i-1}$
Class I Fitness Structure Condition	$\alpha_{i+1} < \beta_i$	$\frac{\phi_{i+1}\pi_{i+1}}{i+1} < \frac{(1-\phi_i)\pi_i}{n-i}$
Class II Fitness Structure Condition	$\alpha_{i+1} > \beta_i$	$\frac{\phi_{i+1}\pi_{i+1}}{i+1} > \frac{(1-\phi_i)\pi_i}{n-i}$
Dynamics	$\overline{w}\,\overline{p}(t+1) = \sum_{i=1}^{n} f_i(t)\frac{i}{n}\alpha_i$ $\overline{w}\,\overline{q}(t+1) = \sum_{i=0}^{n-1} f_i(t)\frac{(n-i)}{n}\beta_i$	$\overline{\pi}\,\overline{p}(t+1) = \sum_{i=1}^{n} f_i(t)\phi_i\pi_i$ $\overline{\pi}\,\overline{q}(t+1) = \sum_{i=0}^{n-1} f_i(t)(1-\phi_i)\pi_i$
Marginal Condition for Type ***A*** Increase	$E[\alpha_{X+1}] > E[\beta_Y]$	$E\left[\frac{\phi_{X+1}\pi_{X+1}}{X+1}\right] > E\left[\frac{(1-\phi_Y)\pi_Y}{n-Y}\right]$
Price's Condition for Type ***A*** Increase	$cov(w, p) > -E[w(\Delta p)]$	$cov(\pi, p) > E[\pi(p-\phi)]$

See text for the relevant values of *i* in the above table. The last two rows of this table are explained in our companion paper.

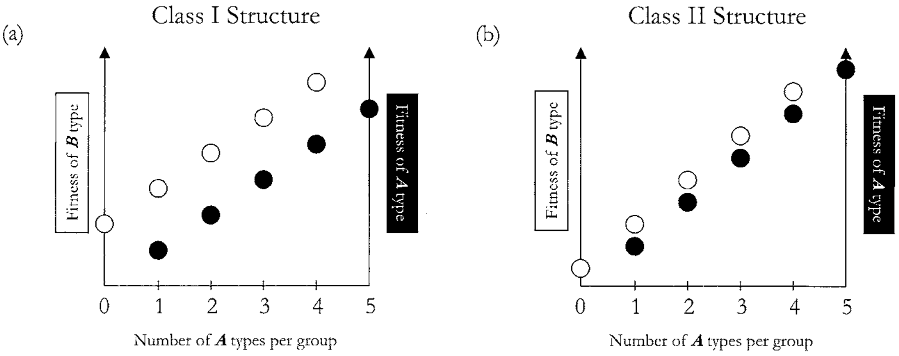


Figure 1. (a) An example of a Class I fitness structure that also satisfies Sober and Wilson's definition of altruism. Any filled point (x_A,y_A) gives the fitness, y_A, of an altruist in a group with x_A altruists (since the groups have size 5, $x_A \in \{1,2,3,4,5\}$). Similarly, any unfilled point (x_B,y_B) gives the fitness, y_B, of a selfish type in a group with x_B altruists ($x_B \in \{0,1,2,3,4\}$). If the fitness for an altruist in a group with x+1 altruists is compared to the fitness of a selfish type in a group with x altruists (where $x \in \{0,1,2,3,4\}$) we see that the selfish type always has a higher fitness. (b) Here if the same comparison between fitnesses is made, the altruist has the higher fitness. This is an example of a Class II fitness structure, which also satisfies Sober and Wilson's definition of altruism.

A useful way to visualize the individualist version of the fitness structure is shown in Figure 1 for the case of $n = 5$. On the x-axis is the number of A types in the group (0, 1, 2, 3, 4, or 5). On the y-axis in a measure of fitness (e.g., number of offspring). The fitness of A (filled points) and B (empty points) in any group context can thus be represented on a single graph.[16] Figure 1a shows a Class I structure and Figure 1b illustrates a Class II structure. To distinguish the structures, simply pick any A fitness (filled point) and compare its y-value to that of the B fitness (empty point) 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.

The distinction between Class I and Class II structures is important, because of the following result:

> **The Matessi-Jayakar Proposition**
>
> If individuals form groups at random each generation, type ***B*** will invade, fix and remain stable to invasion by type ***A*** when the fitness structure is Class I; whereas, type ***A*** will invade, fix and remain stable to invasion by type ***B*** when the fitness structure is Class II.

This proposition (hereafter abbreviated the "M-J proposition") is proved in the Appendix (see Matessi and Jayakar (1976) for another approach to demonstrating this proposition). The invasion and stability criteria have been investigated in other models (see Cohen and Eshel 1976 and Uyenoyama and

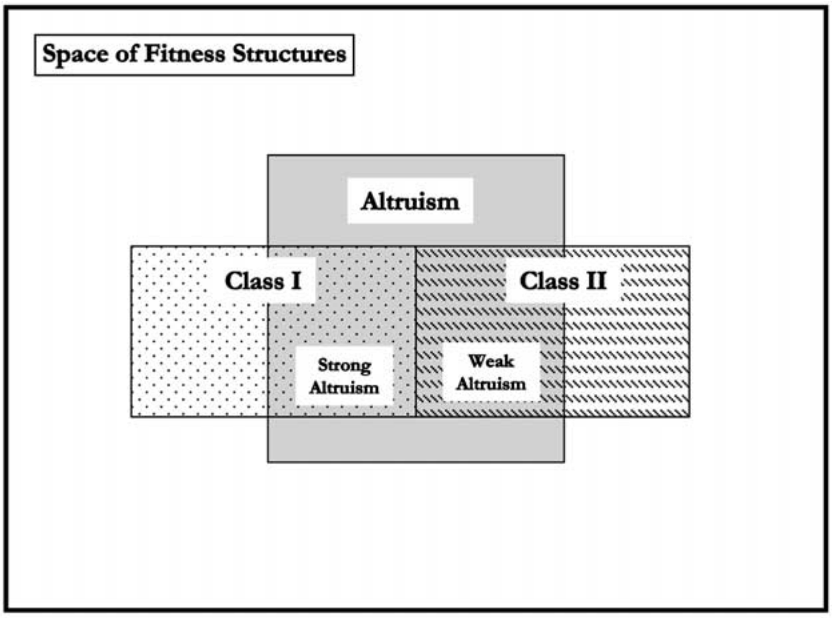


Figure 2. The fitness structures that satisfy Sober and Wilson's definition of altruism (represented by the gray square) are a proper subset within the space of all possible fitness structures (represented by the large white rectangle). Also, the set of Class I and Class II structures form two mutually exclusive proper subsets (represented by the "dotted" and "scratched" rectangles, respectively). Class I fitness structures that also satisfy the condition for altruism (the dotted gray region) are structures with "strong altruism" (Wilson 1980, 1990). Class II fitness structures that also satisfy the conditions for altruism (the scratched gray region) are structures with "weak altruism." Note that there are Class I and Class II structures that do not satisfy the conditions for altruism (the white dotted and white scratched regions, respectively). Also, there are cases of altruism that do not possess either Class I or Class II structure (the gray region that is neither dotted nor scratched).

Feldman 1980). Note that the M-J proposition requires neither inequality (8) nor (9) and thus also applies to systems which would not qualify as "altruistic" under the definition of Sober and Wilson. Even if conditions (8) and (9) are satisfied, condition (12) *or* condition (13) (or neither) may hold (see Figure 2). That is, Sober and Wilson's altruists can possess a Class I or Class II fitness structure (for instance, see Figure 1). Consequently, identifying a behavior as altruistic by Sober and Wilson's definition does not lead to a direct prediction on evolutionary dynamics when groups are formed randomly. On the other hand, if either condition (12) or condition (13) holds, a clear prediction follows from the M-J proposition.

An intuitive way to understand the M-J proposition involves "last member analysis." As in Figure 1, assume that groups are of size 5. Suppose you are the last member to join a group (that is, the group already contains 4 individuals). For instance, assume that all of the other members are selfish. Now, if someone gives you a choice between being altruistic or selfish, and if your goal is to maximize your fitness, which option should you pick? The choice here is between receiving a fitness value of α_1 (if you decide to be altruistic) and receiving a fitness of β_0 (if you go the selfish route). For the Class I structure, $\alpha_1 < \beta_0$ (see condition (12) and Figure 1a); therefore, you should pick selfishness. For the Class II structure, $\alpha_1 > \beta_0$ (see condition (13) and Figure 1b); therefore, you should pick altruism. In general, if there are i altruists already in the group and you are given a choice between altruism and selfishness, then your choice will depend on the relationship between α_{i+1} and β_i. Thus, given a Class I structure, you should always pick selfishness (since $\alpha_{i+1} < \beta_i$), while for a Class II structure, you should always pick altruism (since $\alpha_{i+1} > \beta_i$). As shown rigorously in our companion paper (in this issue), last member analysis is always reliable when there is random formation of groups.[17]

While Sober and Wilson's definition of altruism in *UO* is well defined, it does miss the distinction between Class I and Class II fitness structures. The things it lumps together can be distinguished in a vivid way via a particular kind of individualist representation of the system. We should note that in some of his other work, Wilson gives a more fine-grained set of definitions. If an altruist receives a net benefit from its behavior, but other group members receive a larger benefit in the process, we have a case of "weak altruism." Alternatively, if the altruist incurs a net cost by virtue of its action, while other group members benefit, we have a case of "strong altruism" (Wilson 1980, 1990). Thus, a system with "strong altruism" possesses a Class I fitness structure while a system with "weak altruism" possesses a Class II fitness structure (see Wilson 1980 and Figure 2). Some authors have made the Class I structure a necessary condition for their definition of altruism (Hamilton 1975; Matessi and Karlin 1984; Grafen 1984; Nunney 1985). These authors would not consider Class II structures to be "altruism" at all, not even "weak altruism".[18] We will not address these semantic issues in this paper and instead will focus on the evolutionary roles of these two fitness structures (see entries in Keller and Lloyd 1992 for discussion of different definitions of altruism).

Table 1 summarizes the results from this and the previous section by laying out the perspectives side by side. This format emphasizes the interchangeability of the two perspectives. While the connection between the approaches is transparent (see the second row of Table 1), certain results will be clearer from one perspective over the other. Thus, we see that Sober and Wilson's definition for altruism is more compact from the multi-level perspective,

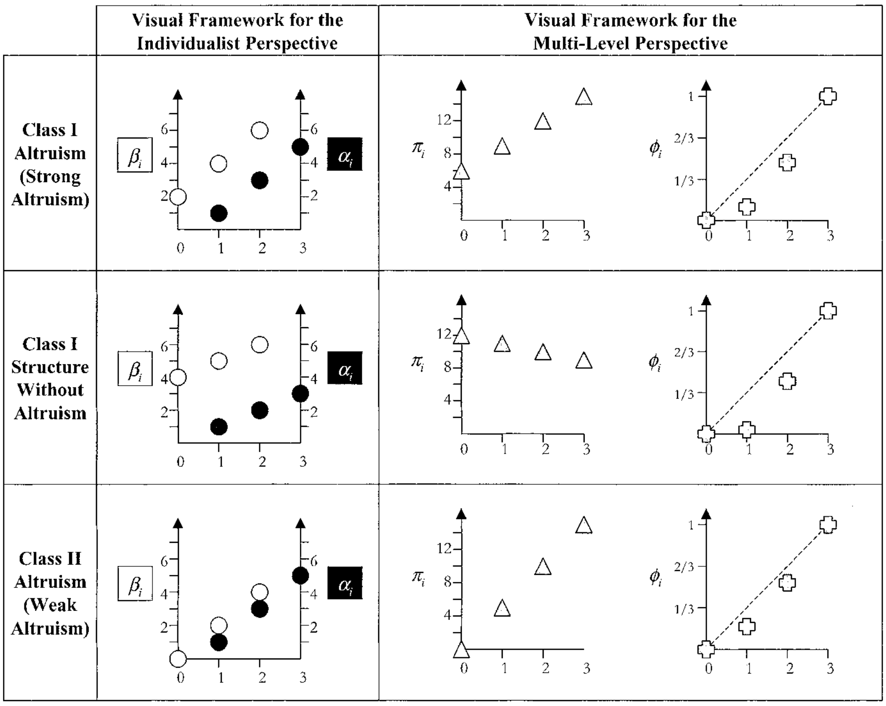


Figure 3. The visual representations of the individualist and multi-level frameworks. Each row corresponds to a certain fitness structure. For the individualist perspective, both α and β values (the black and white points, respectively) can be represented on the same graph with the number of *A* types (i) on the x-axis. For the multi-level perspective, two graphs are needed, one showing group productivity (π values given by the plotted triangles) and the other showing the fraction of this productivity claimed by *A* type (ϕ values given by the plotted crosses). As described in the text, identification of altruism is easier using the visual representation of the multi-level perspective, while identification of Class I or Class II structure is facilitated by the visual representation of the individualist perspective.

while the distinction between Class I and Class II structure is simpler from the individualist perspective. As we describe in our companion paper, certain statistical approaches to evolutionary dynamics will be more natural from either the individualist perspective (marginal approach) or the multi-level perspective (Price's approach).

Figure 3 shows the visual representations associated with each perspective, where each row of the figure corresponds to a specific fitness structure. Identifying "altruism" within a system is facilitated by the visual representation of the multi-level perspective (the second column of Figure 3).[19] To satisfy Sober and Wilson's definition, the following must hold: on a graph of π values, π_i must monotonically increase with i; and on a graph of ϕ values, ϕ_i must remain below the line $\phi_i = i/n$ (the dashed line) for all

$i \in \{1,2,3, \ldots n-1\}$ (see Wilson 1990). Thus, it is immediately clear that the systems in rows 1 and 3 of the figure describe "altruism," whereas the system from the second row does not. The visual framework of the individualist perspective is less helpful in identifying altruism within systems. This is because one condition for altruism involves combining α and β values in ways that are difficult to process visually (see Table 1).[20]

On the other hand, the individualist graphs do facilitate identification of Class I or Class II structure. Thus, it is immediately clear that systems in rows 1 and 2 are Class I, while the system in row 3 is Class II. The visual representation of the multi-level perspective is less helpful in distinguishing Class I versus Class II structure (i.e., rows 1 and 3 look qualitatively similar within the multi-level framework). This is because the condition for Class I or Class II structure involves combining π and ϕ values in ways that are difficult to process visually (see Table 1). When a fitness structure is Class I or Class II, evolutionary information on the system is given by the M-J proposition (given random group formation). Since the evolutionary behavior of systems is often of primary interest, the visual representation of the individualist framework can be particularly useful. We will see this in the next section.

4. Biological and Game-Theoretic Examples

In this section, we illustrate the heuristic power of the individualist perspective and, in particular, its associated visual representation. However, we will also begin to make a case for the value of gestalt-switching. To achieve these aims, we consider several examples. We start with some symmetric two-player games. Next, we consider some classic one-locus diallelic population genetic models, focusing on meiotic drive. We conclude with the amazing life cycle of the brain worm.

Games

In this section, we discuss four symmetric two-player games. Three are well known (the Prisoner's Dilemma, the Iterated Prisoner's Dilemma, and the Hawk-Dove game), while the last game (the Tandem Bicycle Contest) is new. In size 2 groups, there are *only* four altruistic fitness structures possessing unique evolutionary characteristics. Maynard Smith (1987) gives examples of three of the four structures from an individualist perspective and Wilson (1990) gives examples of (a different set of) three of the four structures from a multi-level perspective. Each game exemplifies one of these four structures. As we shall see (and as Maynard Smith (1987) and Wilson (1990) point out)

Table 2. Payoffs to a Focal Individual Playing a Prisoner's Dilemma

	Partner Cooperates	Partner Defects
Focal Individual Cooperates	$R = \alpha_2 = 3$	$S = \alpha_1 = 0$
Focal Individual Defects	$T = \beta_1 = 5$	$P = \beta_0 = 1$

In the table, R is the reward for mutual cooperation, P is the punishment for mutual defection, S is the "sucker's payoff" (cooperating when your partner defects), and T is the temptation to defect (a high payoff for defecting when your partner cooperates). The numbers and parameter labels are taken from Axelrod (1984).

in the next section these structures resurface in meiotic drive systems within a population genetic context.

The Prisoner's Dilemma:

We begin with the most famous of all cases from game theory, the Prisoner's Dilemma. This game has played a central role in both models and discussions about the evolution of altruism. Here, we consider two types of individuals, cooperators (***A***) and defectors (***B***). These individuals form pairs and their fitness depends on their partner. Using the individualist perspective, α_i gives the fitness of a cooperator in a pair with i cooperators (where $i \in \{1,2\}$) and β_j gives the fitness of a defector in a pair with j cooperators (where $j \in \{0,1\}$). If $\alpha_1 < \beta_0 < \alpha_2 < \beta_1$ and $(\alpha_1 + \beta_1)/2 < \alpha_2$, then individuals are playing a Prisoner's Dilemma (see Uyenoyama and Feldman 1992).

In Figure 4a, we plot fitnesses for the one-shot Prisoner's Dilemma used in Axelrod's famous computer tournament (Axelrod 1984). In Table 2, we give the standard "payoff matrix" for the Prisoner's Dilemma for comparison to Figure 4a. From the figure, we see that the fitness structure is Class I and therefore defectors will sweep to fixation if pairs are formed at random. In general, the relationship $\alpha_{i+1} < \beta_i$, which defines a Class I structure, is a necessary condition for the Prisoner's Dilemma. Therefore, by the M-J proposition, the defectors must always fix if partners are picked randomly.[21]

As is widely known, there must be non-random partner selection if cooperators are to stand a chance. Specifically, a *clumped* distribution of groups is necessary, where cooperators pair preferentially with cooperators and defectors are left to pair with defectors (i.e., the tails of $f(t)$ must be given more weight than the tails of the binomial distribution). Let us see how to approach this issue within our framework. Consider an extreme clumping case in which only pure groups form, and the mixed pair fitnesses are never realized. We see that the fitness structure has become completely "Class II like" (see Figure 4b) and cooperation fixes.

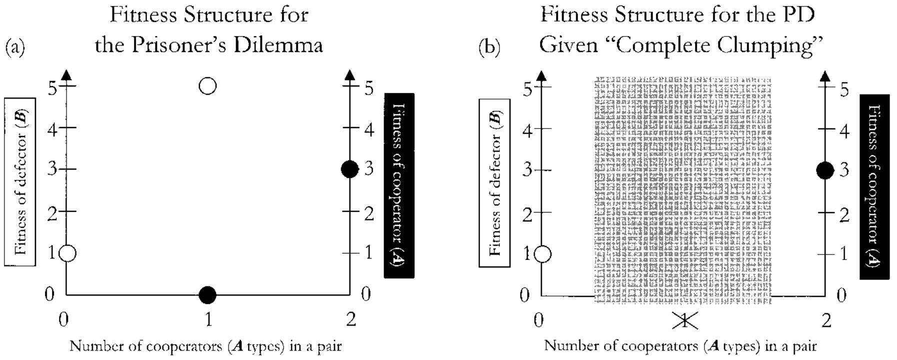


Figure 4. (a) The fitness structure for the Prisoner's Dilemma. Using the payoff values from Axelrod (1984), a defector playing a defector gets 1, a defector playing a cooperator gets 5, a cooperator playing a defector gets 0, and a cooperator playing a cooperator gets 3. This structure is Class I and one would predict that defection will fix if pairs are formed at random. (b) If cooperators only play with cooperators (and consequently defectors must play one another) the mixed pair is non-existent. The newly weighted fitness structure favors cooperators.

In cases like this, where certain types of groups are not formed at all (i.e., $f_j(t) = 0$ for some j value), there may be a "Class I like" or "Class II like" fitness relation discernable by looking at a certain α_i and some β value to the left (other than β_{i-1}). However, we must stress that the full evolutionary consequences within systems where groups do not form randomly will depend critically on the distribution of group types (i.e., $f(t)$). Thus, there are no general visual shortcuts that will yield evolutionary predictions when non-random group formation occurs. With the extreme clumping in this example, however, each type only appears in one combination (its pure group) and hence only has one fitness value. Since the cooperator's fitness is higher, cooperation will fix.[22]

Tit for Tat:

Individuals may form pairs that persist over many time steps, where in each time step, each individual will decide whether to cooperate or defect. If at each time step the payoff structure satisfies a Prisoner's Dilemma, the individuals are said to play an Iterated Prisoner's Dilemma. One strategy that has proven to be particularly successful for the Iterated Prisoner's Dilemma is Tit-for-Tat (TFT). A TFT individual will cooperate with its partner in the first time step, and then, for all subsequent time steps, copy its partner's play (cooperate or defect) from the previous time step. If we consider a population made up of TFT individuals (***A***'s) and individuals that always defect (***B***'s), then, with sufficient time steps of play, the fitness structure can take the form presented in Figure 5.

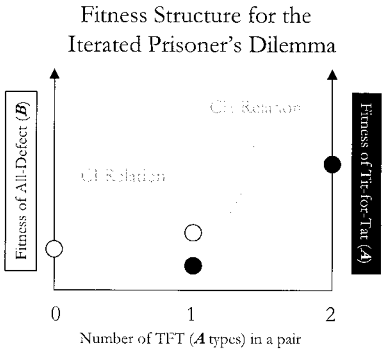
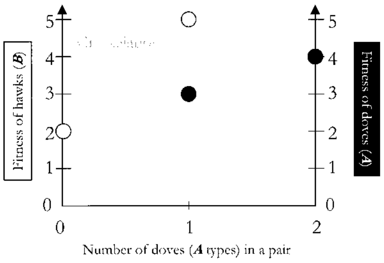


Figure 5. Given sufficient rounds of play, the fitness structure of the Iterated Prisoner's Dilemma with Tit-for-Tat (TFT) and All-Defect (ALLD). We see that when we compare the fitness of TFT in the mixed pair with the fitness of ALLD in its pure pair, ALLD has a higher fitness (i.e., a Class I relationship is found). We also see that when we compare the fitness of TFT in its pure pair with the fitness of ALLD in the mixed pair, TFT has a higher fitness (i.e., a Class II relationship). In this case, we have disruptive selection, where either TFT or ALLD can be favored if common enough.

This fitness structure is neither Class I nor Class II, but rather a hybrid. If mostly always defecting individuals (ALLD) are present and group formation is random, then most pairs will be (ALLD, ALLD) or (ALLD, TFT). This part of the fitness structure is Class I (see Figure 5). Specifically, since $\alpha_1 < \beta_0$, ALLD is stable to invasion by TFT given random pair formation (see Matessi and Jayakar 1976 and Cohen and Eshel 1976). However, if mostly TFT individuals are present and group formation is random, then most pairs will be (TFT, TFT) and (TFT, ALLD). As shown in Figure 5, this part of the fitness structure is Class II. Since $\alpha_2 > \beta_1$, TFT is stable to invasion by ALLD given random pair formation.

This is an example of disruptive selection, in which we have two attracting states of the population. If a sufficient portion of the population is TFT, the TFT strategy fixes; otherwise, ALLD will fix. Note that the number of time steps (as well as the payoff structure of the Prisoner's Dilemma) will determine the size of the basins of attraction of each strategy. In general, the longer the game is played by each pair of players, the wider the TFT basin will be. However, with random group formation, TFT cannot invade the population. Like the non-iterated Prisoner's Dilemma, the initial success of TFT in a defecting population depends on clumping the distribution of pairs. Unlike the non-iterated Prisoner's Dilemma, the continued success of TFT does not depend on clumping. Once the population contains enough TFT players, random group formation will not prevent TFT fixation.



Fitness Structure for the Hawk-Dove Game

Figure 6. The fitness structure of the Hawk-Dove game. Using the values from Maynard Smith (1982), a hawk playing a hawk gets 2, a hawk playing a dove gets 5, a dove playing a hawk gets 3, and a dove playing a dove gets 4. This structure is a hybrid of Class II and Class I parts, and evolution leads to a stable mix of hawks and doves.

Hawks and Doves:

Here we consider one version of the well-known Hawk-Dove game (Maynard Smith 1982). Imagine a population of doves (***A***'s) and hawks (***B***'s) in which pairs form to contest a resource (such as territory). If two doves meet, they peacefully split the resource in two equal parts. If two hawks meet, they fight with one another until one hawk is the sole possessor of the resource. The fighting is costly and each hawk is assumed to have equal probability of obtaining the resource. If a hawk and a dove meet, the hawk secures the resource without any fighting, as the dove calmly retreats from the contest.

If the costs of fighting are more substantial than the value of the resource, we have a fitness structure shown in Figure 6 (the numbers are taken from Maynard Smith 1982, p. 17). Again we have a hybrid fitness structure. Since $\alpha_1 > \beta_0$ (a Class II relation), a population of hawks is unstable to invasion by doves given random pair formation. However, since $\alpha_2 < \beta_1$ (a Class I relation), a population of doves is unstable to invasion by hawks given random pair formation. The result is a protected polymorphism where both strategies coexist (in fact, a stable polymorphism is realized).

A Tandem Bicycle Contest:

In this game, two riders enter a tandem bicycle contest, and the team is awarded 20 dollars if it completes the course within a certain time limit (this award is split evenly between the two riders). Riders come in two varieties: (1) those that invest a dollar in special shoes that attach to the pedals of the bicycle or (2) those that save the dollar and just wear sneakers to the contest. If neither rider of the bicycle wears the special shoes, there is no chance of winning the award. If one rider wears the shoes there is a 50% chance of

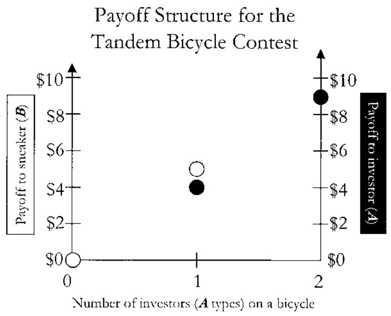


Figure 7. The payoff structure for the Tandem Bicycle Contest. A sneaker riding with a sneaker can expect \$0, a sneaker riding with an investor can expect \$5, an investor riding with a sneaker can expect \$4, and an investor riding with an investor can expect \$9. This structure is Class II. If payoffs are considered to be fitnesses and if pairs are formed at random, then investors should sweep to fixation.

winning. If both riders wear the special shoes, the team is certain to win the award.

Imagine a population of two types of riders, investors (***A***) and "sneakers" (***B***). Using the individualist perspective, α_i gives the expected payoff to an investor in a pair with *i* investors (where $i \in \{1,2\}$) and β_j gives the expected payoff to a sneaker in a pair with *j* investors (where $j \in \{0,1\}$). The expected payoffs are given in Figure 7, where we see $\beta_0 = \$0$, $\beta_1 = \$5$, $\alpha_1 = \$4$, and $\alpha_2 = \$9$. The payoff structure is Class II. If the payoffs are interpreted as fitnesses and partners are chosen at random, then the investors should drive the sneakers to extinction by the M-J proposition. Thus, even though on a mixed tandem bicycle the sneaker does better on average than its investing partner ($\alpha_1 < \beta_1$), the investor sweeps to fixation in the population.[23]

Evolutionary Stable Strategies:

Maynard Smith (1982) defines an evolutionary stable strategy (ESS) as "a strategy such that, if all the members of a population adopt it, no mutant strategy can invade" (p. 204). Suppose there are only two strategies possible, call them ***A*** and ***B***, and fitness is determined by randomly assembled two-player contests in an infinite asexual population.[24] Let W(**I**,**J**) be the fitness of strategy **I** given a partner with strategy **J** (note that possibly **I** = **J**). Strategy ***A*** is an ESS if $W(\boldsymbol{A},\boldsymbol{A}) > W(\boldsymbol{B},\boldsymbol{A})$ (or if $W(\boldsymbol{A},\boldsymbol{A}) = W(\boldsymbol{B},\boldsymbol{A})$ and $W(\boldsymbol{A},\boldsymbol{B}) > W(\boldsymbol{B},\boldsymbol{B})$) (see Maynard Smith 1982, p. 24). Using the terminology from above, if $\alpha_2 > \beta_1$, strategy ***A*** is an ESS. Alternatively, if $\alpha_1 < \beta_0$, strategy ***B*** is an ESS.

In the two-player games above, investors, cooperators, TFT's, and doves are all altruists by Sober and Wilson's definition.[25] However, depending on the fitness structure there can be a single altruistic ESS (e.g., investors in the Tandem Bicycle Contest), a single selfish ESS (e.g., defectors in the Prisoner's Dilemma), an altruistic ESS *and* a selfish ESS (TFT and ALLD in the Iterated Prisoner's Dilemma), or no ESS's at all (there are no pure strategy ESS's in the Hawk-Dove game outlined above). On the other hand, if the fitness structure is known to be Class I or Class II, a definite statement concerning ESS's can be made. If the fitness structure is Class I altruism (i.e., strong altruism), there is a unique selfish ESS. If the fitness structure is Class II altruism (i.e., weak altruism), there is a unique altruistic ESS.

One Locus Population Genetics

Even the most familiar of all evolutionary models, a diploid one-locus population genetics model, can be viewed from either of our two perspectives. We will show this, and then discuss how the game-theoretic models of the previous section can be mapped onto this basic population genetic model.

Let us consider the equations describing the dynamics of two alleles at one locus. Label two alleles at a single locus ***A*** and ***B***, with frequencies in the population at generation t given by $\overline{p}(t)$ and $\overline{q}(t)$, respectively. If we assume random mating (i.e., random formation of allele groups) then the following equations apply:

$$\overline{\omega}\,\overline{p}(t+1) = (\overline{p}(t))^2\omega_{AA} + 2(\overline{p}(t))(\overline{q}(t))\left(\frac{1}{2}\right)\omega_{AB}, \tag{14}$$

$$\overline{\omega}\,\overline{q}(t+1) = 2(\overline{p}(t))(\overline{q}(t))\left(\frac{1}{2}\right)\omega_{AB} + (\overline{q}(t))^2\omega_{BB}, \tag{15}$$

with

$$\overline{\omega} = (\overline{q}(t))^2\omega_{BB} + 2(\overline{p}(t))(\overline{q}(t))\omega_{AB} + (\overline{p}(t))^2\omega_{AA}.$$

Note that equations (14) and (15) can be rewritten in the form of equations (5b) and (6b) by making the following substitutions: $\pi_0 = \omega_{BB}$, $\pi_1 = \omega_{AB}$, $\pi_2 = \omega_{AA}$, $\overline{\pi} = \overline{\omega}$, $\phi_0 = 0$, $\phi_1 = 1/2$, $\phi_2 = 1$, $f_0(t) = (\overline{q}(t))^2$, $f_1(t) = 2(\overline{p}(t))(\overline{q}(t))$, and $f_2(t) = (\overline{p}(t))^2$. Thus, these standard equations are essentially written from the multi-level perspective. Fitnesses are ascribed to collections of two alleles (i.e., the diploid organism) and division of the productivity of the mixed group (i.e., the heterozygote) is fair (i.e., Mendelian segregation implies $\phi_1 = 1/2$). In general, ω_k, with $k \in \{\boldsymbol{AA}, \boldsymbol{AB}, \boldsymbol{BB}\}$, can be thought of as the expected number of "successful" copies of alleles from an organism with genotype k (i.e., alleles that end up in offspring). This is, of course, equivalent to the number of offspring genotype k can expect.

Wilson (1990, 1997) has argued cogently that there is no reason that such a perspective, used freely in population genetics, cannot be used for individuals in groups. We agree. But there is a flip side to this point. There is no reason the individualist perspective, used freely in game theory, cannot be employed for genes in diploid organisms.[26] Standard population genetics treats group structure in a *collective* way, but these models could also be written from a *contextual* perspective. In such a case, the focus shifts to the lower-level particle (the gene) and the genotype of the individual organism in which it appears is now treated as the particle's environment. The contextual perspective yields an orthodox individualist model in the case of the game theory, but it yields an unorthodox gene's eye view model in the case of population genetics.

To see the contextual perspective, we divide both sides of equations (14) and (15) by 2:

$$\frac{\overline{\omega}}{2}\,\overline{p}(t+1) = (\overline{p}(t))^2\frac{\omega_{AA}}{2} + 2(\overline{p}(t))(\overline{q}(t))\left(\frac{1}{2}\right)\frac{\omega_{AB}}{2}, \tag{16}$$

$$\frac{\overline{\omega}}{2}\,\overline{q}(t+1) = 2(\overline{p}(t))(\overline{q}(t))\left(\frac{1}{2}\right)\frac{\omega_{AB}}{2} + (\overline{q}(t))^2\frac{\omega_{BB}}{2}. \tag{17}$$

Equations (16) and (17) can be easily translated into the form of equations (5a) and (6a) via the following relationships: $\alpha_1 = (1/2)\omega_{AB}$, $\alpha_2 = (1/2)\omega_{AA}$, $\beta_0 = (1/2)\omega_{BB}$, $\beta_1 = (1/2)\omega_{AB}$, $\overline{w} = \frac{\overline{\omega}}{2}$, $f_0(t) = (\overline{q}(t))^2$, $f_1(t) = 2(\overline{p}(t))(\overline{q}(t))$, and $f_2(t) = (\overline{p}(t))^2$. Thus, the dynamics of the population genetic model can be redescribed using the fitnesses of alleles. The fitness of, for instance, allele $\boldsymbol{A}$ in a given genotypic context is given by the α_i terms, where α_i represents the expected number of copies contributed to the next generation by an $\boldsymbol{A}$ allele in group with i $\boldsymbol{A}$'s ($i \in \{1,2\}$).[27] Note that given Mendelian segregation, both alleles in the heterozygote share the same fitness (i.e., $\alpha_1 = \beta_1$). In this way, the simple system can be formulated from either a gene's eye view (contextual) or a multi-level perspective (collective). The two perspectives become more discernable when the assumption of Mendelian inheritance is violated – for instance, when there is meiotic drive.

Wilson (1990) gives the equations for allelic frequency change when there is meiotic drive as follows:

$$\overline{\omega}\,\overline{p}(t+1) = (\overline{p}(t))^2\omega_{AA} + 2(\overline{p}(t))(\overline{q}(t))(g)\omega_{AB}, \tag{18}$$

$$\overline{\omega}\,\overline{q}(t+1) = 2(\overline{p}(t))(\overline{q}(t))(1-g)\omega_{AB} + (\overline{q}(t))^2\omega_{BB}, \tag{19}$$

with

$$\overline{\omega} = (\overline{q}(t))^2\omega_{BB} + 2(\overline{p}(t))(\overline{q}(t))\omega_{AB} + (\overline{p}(t))^2\omega_{AA}.$$

If $g < 1/2$, then segregation is being distorted in favor of ***B*** alleles in heterozygotes. Again, if we set $\pi_0 = \omega_{BB}$, $\pi_1 = \omega_{AB}$, $\pi_2 = \omega_{AA}$, $\phi_0 = 0$, $\phi_1 = g$, and $\phi_2 = 1$, equations (18) and (19) are written from the multi-level perspective.

On the other hand, by setting $\alpha_1 = (g)\omega_{AB}$, $\alpha_2 = (1/2)\omega_{AA}$, $\beta_0 = (1/2)\omega_{BB}$, and $\beta_1 = (1\text{-}g)\omega_{AB}$, we obtain the gene's eye view. Note that the α's and β's are interchangeable with the π's and ϕ's by the operations spelled out in section 2. As must be the case, both descriptions of the fitness structure have the same number of degrees of freedom (namely, 4). One question remains: Why use the unconventional contextual perspective for this system?

The answer again relates to the heuristic value of the contextual perspective. Let us assume that the segregation distorting allele is the ***B*** allele. Therefore, heterozygotes produce more ***B*** gametes than ***A*** gametes ($g<1/2$) and condition (8) holds. Let us also assume that the number of successful gametes from a diploid organism (i.e., gametes that end up in the offspring of the parent generation) increases with the frequency of ***A*** alleles in the organism (this frequency is either 0, 1/2 or 1) and thus condition (9) holds. This second assumption may be reasonable given that segregation distorters often compromise the fitness of their owners, especially in the double dose. So, we see the allele that is not the segregation distorter is an altruist by Sober and Wilson's definition. Given these assumptions, Figure 8 graphs out the four possible types of fitness structures from the contextual perspective.[28]

We can have either a pure Class I structure or pure Class II structure (Figure 8a and 8b, respectively). Note that these structures are identical in form to the one-shot Prisoner's Dilemma and the Tandem Bicycle Contest, respectively. Here we know the evolutionary outcomes from the M-J proposition and the gray arrows indicate these outcomes: an arrow to the left indicates that ***B*** is favored and an arrow to the right indicates that ***A*** is favored. However, there are also two hybrid structures to consider. The first is a replica of the Tit-for-Tat structure (Figure 8c). In this case, the left-ward pointing gray arrow indicates that ***B*** will fix when common, while the right-ward pointing gray arrow indicates that ***A*** will fix when common. Disruptive selection results. The second hybrid is a replica of the Hawk-Dove structure (Figure 8d). For this structure, the left-ward pointing gray arrow indicates that ***B*** will increase when rare and the right-ward pointing gray arrow indicates that ***A*** will increase when rare. A protected polymorphism results (as in the Hawk-Dove game, the dynamical behavior of the system leads to a *stable* polymorphism).[29] In general, if groups contain only two particles (for instance, genes in a diploid organism or individuals in a two-player game), the space of possible fitness structures is *exhausted* by the four types of fitness structures shown in Figure 8 (this breakdown is shown schematically in Figure 9).

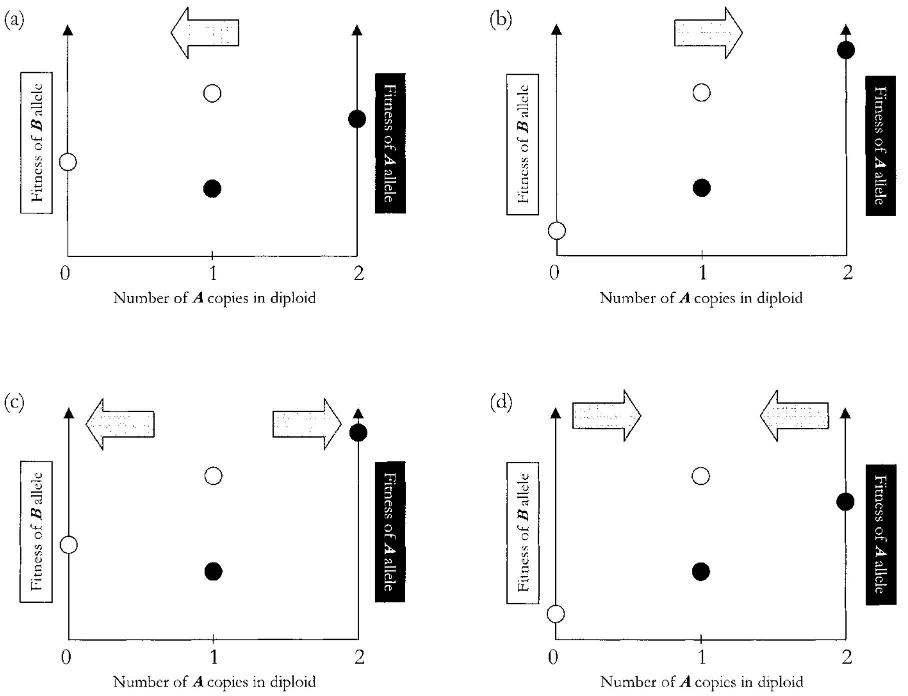


Figure 8. Fitness structures for a diallelic one locus segregation distortion system assuming that Sober and Wilson's definition of altruism applies. In all diagrams, the ***BB*** homozygote corresponds to x-value 0, the ***AB*** heterozygote corresponds to x-value 1, and the ***AA*** homozygote corresponds to x-value 2. (a) A Class I fitness structure. As the left-ward pointing gray arrow indicates and as the M-J proposition claims, the population will fix for the ***B*** allele. (b) A Class II fitness structure. As the right-ward pointing gray arrow indicates and as the M-J proposition claims, the population will fix for the ***A*** allele. (c) A Class I-Class II hybrid. As in the TFT-ALLD case of Figure 5, either allele will fix if common enough. This is represented by the two gray arrows, one of which points to the ***A*** fixation (right-ward) and the other pointing to the ***B*** fixation (left-ward). Disruptive selection results. (d) A Class II-Class I hybrid. Here, either allele will invade when rare. This is represented by two gray arrows, the left-ward pointing one represents invasion of ***B*** when rare and the right-ward pointing one represents invasion of ***A*** when rare. A (stable) polymorphism results.

If there is random mating and we observe a polymorphism maintained indefinitely in a diallelic segregation distortion system, then the fitness structure must be of the type shown in Figure 8d (unless the population starts off and stays on the unstable equilibrium separating the basins of attraction in Figure 8c – an extremely unlikely event). Note that the gray arrows drawn in Figure 8 rely on the assumption of random union of gametes (i.e., random mating). If there is non-random mating (for instance, positive assortment mating or mating in a spatially structured population) then the ***A*** allele might increase when rare in either Figure 8a or 8c.

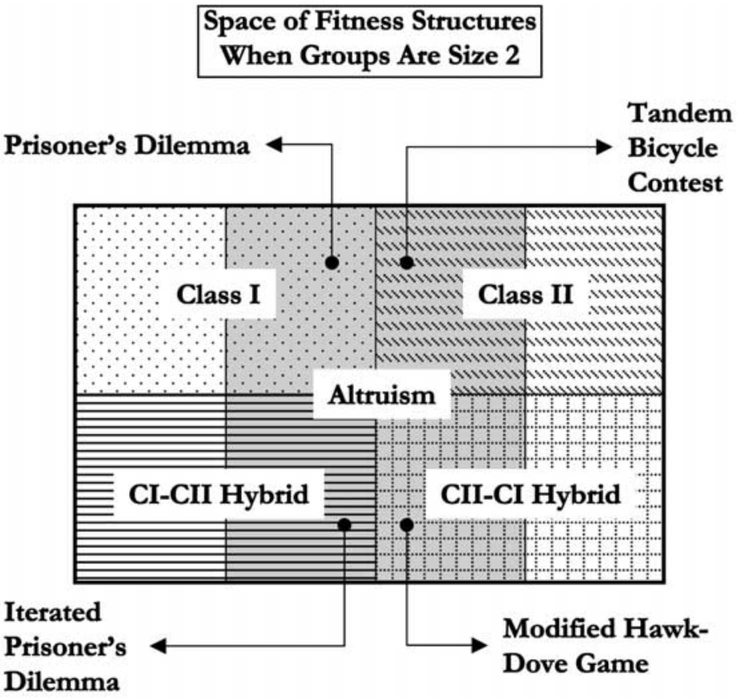


Figure 9. When the groups are of size 2, we are able to describe all fitness structures as one of four fundamental types (if we make the assumption that $\alpha_{i+1} \neq \beta_i$ for $i \in \{0,1\}$). These types include Class I structures (dotted region), Class II structures (scratched region), Class I/Class II hybrids (horizontal lined region) and Class II/Class I hybrids (cross-hatched region). Within each of these four basic types, there will be some structures that satisfy the conditions for altruism (gray region) and some that do not (white region). The four game-theoretic examples considered exemplify each fundamental type as shown. (Note that we must use a *modified* Hawk-Dove game as explained in note 25).

In this section, we have framed evolutionary outcomes at a segregation locus in terms of population structured selection and explored the heuristic value of the contextual perspective in classifying these outcomes. However, other population genetic systems can also be framed in terms of the gene's eye view (as opposed to the more standard multi-level perspective). Graphs like those from Figure 8 can be created for systems with overdominance and underdominance and the evolutionary outcome can be read off the graph. In general, selection on genes in diploid groups *is* a type of population structured selection and the contextual perspective can have heuristic value when it comes to predicting evolutionary trajectories.

The Brain Worm

Our last example will be more complicated. In *UO*, Sober and Wilson describe the fascinating life cycle of the parasite *Dicrocoelium dendriticum* (pg. 18). This tremotode spends its adult stage in the liver of cows and sheep. Eggs leave the host with the feces, where a land snail picks up some quantity. In the snail, the parasite undergoes a couple of rounds of asexual reproduction and exits the snail as cercaria (another stage in the life-cycle) enveloped in mucus. An ant then becomes infected with a population of cercariae via consumption of the snail's mucus. If there is an "altruistic" individual in the bunch (the brain worm), it will sacrifice itself by forming a cyst in the ant's brain. The presence of the brain cyst causes the ant to climb to the tip of a grass blade, where the probability of being consumed by a cow or sheep is greatly enhanced. Thus, the likelihood of any individual contributing to future generations is augmented considerably if there is a self-sacrificing brain worm in its "ant-group" (the collection of cercariae in its host ant). Of course, this good fortune is not enjoyed by the hapless worm forming the brain cyst!

We will gloss over some of the biological details here in order to make a few points and to connect this system to our graphical representation. We posit two types of parasites, the altruistic brain worm (***A***) and its selfish counterpart (***B***). For simplicity, assume that the snail always takes up three eggs at random. While in the snail, we will assume that each of the three parasites doubles for a total of six cercariae and we will assume that a single ant ingests this group in its entirety (see Figure 10a).

Assume that any selfish parasite in an ant-group with one or more brain worms will, with absolute certainty, contribute to the next adult generation. Thus, we will say this individual has fitness 1 ($\beta_i = 1$, for $i \in \{1,2,3,4,5,6\}$). We will assume that there is a very low probability that a purely selfish group will end up in a mammalian host (for instance, say $\beta_0 = 0.1$). If there is only one altruist in the group, it has fitness 0 ($\alpha_1 = 0$) since it dies when forming the brain cyst. If there is more than one altruist, we assume that the individual forming the brain cyst is chosen at random and receives fitness 0, while the other altruists get fitness 1. Now, considering average fitness of an altruist within a certain group context (note we are averaging within groups here, not over groups), we have $\alpha_2 = 1/2$, $\alpha_3 = 2/3$, $\alpha_4 = 3/4$, $\alpha_5 = 4/5$, and $\alpha_6 = 5/6$. That is, in a group with i altruists ($i \geq 2$), one individual has fitness 0 and $i - 1$ individuals have fitness 1. This gives an average altruist fitness of $(i - 1)/i$.

For all possible groups of 6 worms (that is, groups with 0, 1, 2, 3, 4, 5, or 6 altruists) the fitness structure is given by Figure 10b. Note that the fitness structure in Figure 10b is Class I. Thus, if groups of six cercariae are formed randomly, then altruism could not invade the population. However, the groups

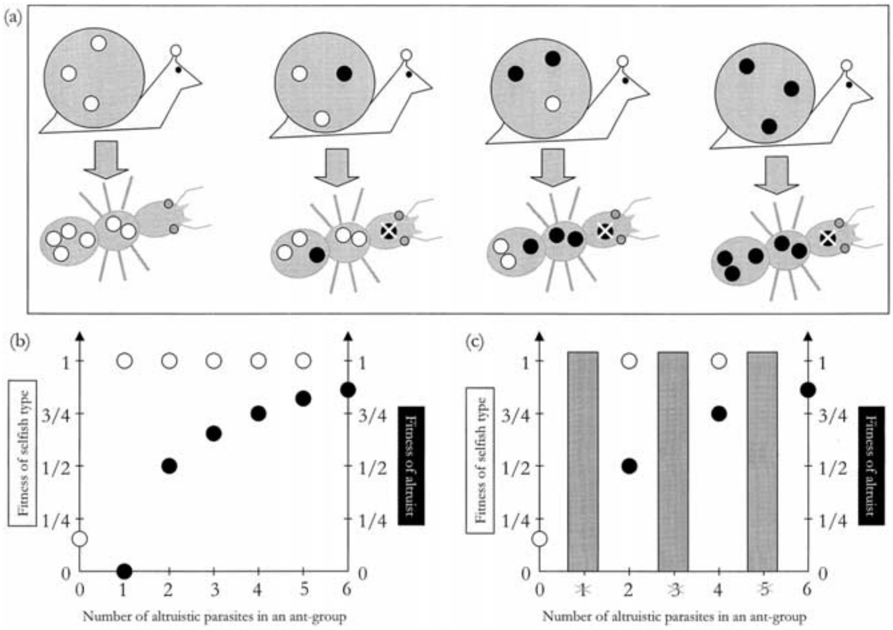


Figure 10. (a) A simplified schematic showing the life cycle of the trematode parasite. Given that there are two types of parasites (***A*** as black circles and ***B*** as white circles), there are four possible groups of size 3 which originally inhabit the land snail (called "snail-groups"). These individuals double and the group of six is taken up by the ant host. The four possible groups of six parasites within the ant (called "ant-groups") are shown. If there is one or more type ***A*** parasite (the brain worm), one of these ***A***'s will form a cyst in the ant's brain and affect its behavior. (b) The fitness structure if all seven possible groups of six parasites could be formed in the ant. See text for the derivation. Note this is a Class I fitness structure, so if groups are formed at random (*which they are not*) then the brain worm (type ***A***) would go extinct. (c) Because of the reproductive event depicted in (a), there are only four possible types of ant-groups. Thus, only parts of the fitness structure from (b) are seen by the parasites. The newly weighted fitness structure favors the initial invasion of altruism, but not its fixation.

in the ants are not formed randomly because a reproductive event took place in the snail. If there was a strict doubling of the original three worms in the snail, there are no groups in the ants that have an odd number of brain worms. That is, there are no groups with 1, 3 or 5 brain worms in the ants. Thus, the possible individual fitnesses are shown in Figure 10c. Note that $\beta_0 < \alpha_2$; thus, there is a "Class II relationship" between the altruist in the group with the smallest number of altruists and a selfish type in a uniformly selfish group. We can think of reproduction in the snail as a "clumping mechanism" which adjusts $f(t)$, so that altruism can invade. Altruism will not fix in this population, however. This can be understood intuitively by looking at the far right of Figure 10c – here the fitness structure is still "Class I like." Thus, we

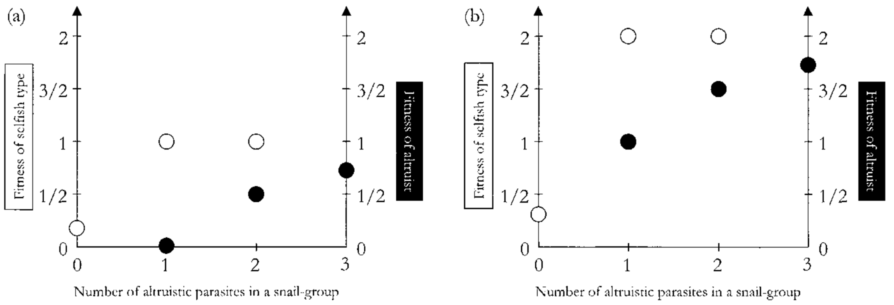


Figure 11. (a) The fitness structure for parasites in the snail-group (i.e., the parent group) if they were to abstain from reproduction. Thus, the ant-groups would also be of size 3 (i.e., they would be the same as the snail-groups) and we see that the fitness structure is Class I, which favors the selfish parasite. (b) The fitness structure for the parasites in the snail-group given reproduction. Here fitness is measured by the number of offspring that successfully enter the mammalian host. Now we see that we have a Class II-Class I hybrid fitness structure and a protected polymorphism will result.

expect a protected polymorphism of types in the population (either altruists or selfish types can invade when rare) as Sober and Wilson predict in *UO*.

There is another way to understand evolution in our simplified brain worm system from Figure 10. This is done by calculating the fitnesses of the three parasites while still in the land snail (before reproduction and before entering into the ant-group). Fitness in this case will be the number of offspring the parasite can expect to successfully inhabit a mammalian host in the next generation. Given our previous assumptions: $\alpha_1 = 1$, $\alpha_2 = 3/2$, $\alpha_3 = 5/3$, $\beta_0 = 1/5$, $\beta_1 = 2$, and $\beta_2 = 2$ as shown in Figure 11b. We see that this fitness structure is a hybrid of Class I and Class II parts of the type shown in Figure 8d. As in that case of a segregation distortion system, we expect a protected polymorphism of altruists and selfish types. And this occurs even when the groups in the snails *are formed at random*.

If the parasites in this snail-group did not undergo reproduction, the fitness structure would be $\alpha_1 = 0$, $\alpha_2 = 1/2$, $\alpha_3 = 2/3$, $\beta_0 = 1/10$, $\beta_1 = 1$, and $\beta_2 = 1$ as shown in Figure 11a. Here the fitness structure is Class I and random formation of groups would favor selfish types. Thus, we see that the effect of reproduction is to alter the fitness structure of our *snail*-groups. Thus, from the perspective of *offspring* groups, reproduction is a clumping mechanism favoring altruism where $f(t)$ is affected (compare Figure 10b to Figure 10c). Whereas, from the perspective of *parent* groups in this system, reproduction is a means of fitness structure manipulation where relationships between the values of α's and β's change (compare Figure 11a to Figure 11b).

In our simplification above, we forced the individuals to undergo a strict doubling. One might ask if the conclusions differ if we assume “random production of offspring" from the parent group; that is, parents are picked with equal probability and with replacement to copy into offspring. Wilson (1977) discusses this case. Essentially, even if both parent group formation and offspring group formation are random processes, the offspring groups will be more clumped than the binomial distribution (a scenario tending to favor altruism). This is analogous to the Walhund effect, where there is an excess of homozygotes when random mating occurs in *structured* populations (i.e., the distribution of the size 2 gene groups, otherwise known as genotypes, deviates from the binomial distribution generated using population-wide allele frequencies). Such production of clumped offspring groups offers a biologically plausible mechanism by which altruism might "swim upstream" against a Class I fitness structure. This is one incarnation of kin selection, where the coefficient of relatedness between members of the offspring group is high due to a "parent bottleneck" (Hamilton 1975; Wilson 1975; Maynard Smith 1976). The smaller the parent group, the more clumped these offspring groups will be and the resulting group frequency distribution, $f(t)$, will be more likely to work in favor of altruism.

Gestalt-Switching

In section 2 we presented two mathematically equivalent ways of modeling selection in structured populations. These are two ways of "packaging" the same information about a fitness structure. The various examples we have just described show that although the two parameterizations are mathematically interchangeable, some facts will be more vivid from one perspective than they are from the other. The way that the information is packaged makes a difference.

We have discussed game theoretic examples, population genetics examples, and the unusual case of the brain worm. In the literature, some of these cases are typically handled in a particular manner; that is, they have a standard description that treats group structure in either a collective or a contextual way. Game theoretic modeling is usually done from a contextual perspective – game theory is often considered a paradigm case of an individualist approach to the evolution of behavior (Axelrod and Hamilton 1981). Kin selection models are often thought of in the same way. Population genetics, on the other hand, is usually done using a collective perspective. In each case, one treatment of group structure is orthodox, but another is available. One of our aims here has been to introduce the option of using the alternative perspective. This can be done for each case despite the respective orthodoxy. We suggest that when the selective process is viewed from a different angle, new parts of the total picture come into focus.

In population genetics, a contextual approach to group structure is unorthodox. Applying this approach yields a version of the "gene's eye view" of evolution. When we do this, we find that we can map some standard examples from game theory onto various mixes of viability/fertility selection and meiotic drive. The role of meiotic drive is important here; we find that the fairness of meiosis will *prevent* Mendelian genes from playing some standard games like the Prisoner's Dilemma and Hawk/Dove.

In the case of game theory, a collective approach to group structure is the unorthodox option. Sober and Wilson consider this perspective for the Iterated Prisoner's Dilemma in *UO*. They argue that the success of Tit-for-Tat can be easily understood by employing a multi-level perspective. Their argument goes as follows: A pair with two Tit-for-Tat players has a higher group productivity than a mixed pair, which has a higher productivity than a pair with two defectors. That is, there are differences in productivity *between* groups. Relative to the defector, Tit-for-Tat does have a lower fitness in a mixed pair. However, unlike an unconditional altruist, the Tit-for-Tat player defects shortly after pairing with a defector, thus minimizing the differences in fitness *within* the mixed group. As a result, the TFT player improves its share of the mixed group productivity. Sober and Wilson claim that the diminution of fitness differences within the mixed group (working against Tit-for-Tat) tips the scales in favor of productivity differences between groups (working for Tit-for-Tat), which allows Tit-for-Tat to evolve by group selection. These sorts of facts about group-level productivity (π) and the sharing of productivity within groups (ϕ) are naturally represented with a collective rather than contextual perspective.

Rather than arguing that the Iterated Prisoner's Dilemma *must* be understood as a case of multi-level selection, or arguing that the gene's eye view is the *right* way of thinking about population genetics, we argue that maximum understanding is achieved by retaining the ability to look at each case in both ways. Clearly in structured populations it is often important to ask questions about group-level productivity; this is best done using a collective perspective. But as we saw in the case of the brain worm, some rather complex systems become easier to understand if we use the Class I/Class II distinction to think about fitness structures. That distinction is more easily deployed using an α/β parameterization, hence a contextual treatment of group structure. In section 3 we showed that different authors writing about the evolution of altruism use different definitions of altruism. We have shown that some definitions are more naturally represented and modeled from one perspective or the other.

The central message is that it is a mistake to argue that either a collective or contextual approach to group structure is *better in general*. And even when thinking about a particular example, it will usually be helpful to consider different ways to think about group structure. When we consider selection in structured populations each perspective has advantages for seeing some *pieces* of the entire picture.

5. Summary and Conclusions

In this paper, we focus on different ways of understanding the role of population structure in processes of natural selection. Our discussion shows the compatibility of two themes in the recent literature. One is the need to properly represent the important role of population structure in models of evolution. Sober and Wilson make this theme vivid in *UO*, but they insist that only the multi-level framework is adequate for understanding these issues. The other theme is the interchangeability of individualist and multi-level perspectives, asserted by many and developed in detail by Dugatkin and Reeve (1994) and Sterelny (1996). We agree with Dugatkin, Reeve and Sterelny that arguments about the general superiority of one view over the other should be retired.

We also hope to have cast some light on the overall structure of the selection process in structured populations. The two major ingredients in understanding population structured selection are (1) the fitness structure for the population, which can be represented in different ways, and (2) the distribution of groups in the population. In order to construct a dynamically sufficient model, we must know *how groups are formed* (this requires understanding the biology of the system). If we can obtain such information, then we may be able write *recursions* to describe the dynamics of the system under either the individualist or multi-level perspective.

Even though these perspectives are isomorphic, there can be legitimate heuristic motivation to use one perspective over the other. In this paper we have discussed how the individualist perspective offers a particularly practical and useful visual representation. Specifically, there are two classes of fitness structure that are easy to visually distinguish when this perspective is used. Under certain assumptions, evolutionary outcomes of a system can be read off a graph.

The evolution of altruistic behavior has been a reliable source of controversy. Part of the controversy is a result of different definitions of the word. How does our mathematical development of two distinct perspectives help? First of all, we see that a given definition may be much clearer from one perspective. Second, since the perspectives are interchangeable, there is always the option of reformulating a model from the individualist perspective. And thus, despite the preferred definition of altruism, its evolution may be deduced from the graphical form of the fitness structure. This visual heuristic works when there is random formation of groups. However, when the fitness structure diagram shows that altruism will not invade under random group formation, some biological mechanism for “clumping” the group frequency distribution must be in place in order to explain stable non-zero frequencies of altruists. Sometimes these factors, too, can be represented in simple ways on the graphs. As reproduction within groups (before selection) is one particularly salient mechanism for achieving clumping of types, the graphical tools of the individualist perspective may be helpful in identifying biological systems in which kin selection may have played a role.

Like a number of other workers (Buss 1987; Maynard Smith and Szathmáry 1995; Michod 1999), part of our motivation for focusing on altruism is to understand the general nature of transitions between different levels of organization in evolution. Some might object that our showcasing of the individualist perspective exhibits a reductionist tendency that obscures the role of higher-level individuals in evolution. We suggest that the opposite is the case. Instead, the key to understanding the evolution of higher level organizational units is the ability to move constantly back and forth between a mode of analysis that focuses on lower-level entities and treats the social context of these entities as comprising part of their *environment*, and a mode of analysis that instead treats the collectives of lower-level entities as *wholes*.[30] So our aim is not to downgrade the importance of recognizing higher-level units, as Sober and Wilson legitimately urge, but to show how to switch back and forth between two perspectives on the evolution of these higher-level units.

We ended this paper with illustrations of how the individualist perspective may elucidate evolutionary outcomes in a number of biological systems. Some results from game theoretic analysis and population genetics become easily understood when utilizing the visual tools of the individualist perspective. The standard formulae for diallelic one locus population genetics are written within a multi-level framework (as recognized by Wilson). Perhaps the ease with which researchers use the multi-level framework for population genetics, while they remain skeptical about its use for groups of individuals, testifies to the perceived prominence of the individual organism in the gene-organism-group hierarchy. This is, of course, not difficult to understand given the position of the researcher in this hierarchy!

Acknowledgements

We thank Marc Feldman, Lisa Lloyd, Elliott Sober, Kim Sterelny, Karen Walag, David Sloan Wilson, and Rasmus Winther for many insightful comments on earlier versions of the manuscript.

Appendix: Proof of the Matessi-Jayakar Proposition

In this appendix, we offer a proof of the M-J proposition using the complete dynamic equations of the system. If $\overline{p}$ and $\overline{q}$ are the frequencies of altruists and selfish types in the population and we assume random formation of size n groups (i.e., (7) holds), the frequency recursions are

$$\overline{p}' = \frac{\sum_{i=1}^{n} \binom{n}{i} \overline{p}^{i} \overline{q}^{n-i} (\alpha_i) \left(\frac{i}{n}\right)}{\overline{w}}, \tag{$A1$}$$

and

$$\overline{q}' = \frac{\sum_{i=0}^{n-1} \binom{n}{i} \overline{p}^{i} \overline{q}^{n-i} (\beta_i) \left(\frac{n-i}{n}\right)}{\overline{w}}, \tag{$A2$}$$

with

$$\overline{w} = \sum_{i=1}^{n} \binom{n}{i} \overline{p}^{i} \overline{q}^{n-i} \alpha_i \left(\frac{i}{n}\right) + \sum_{i=0}^{n-1} \binom{n}{i} \overline{p}^{i} \overline{q}^{n-i} \beta_i \left(\frac{n-i}{n}\right).$$

We start by solving equations ($A1$) and ($A2$) for interior equilibria:

$$\hat{\overline{p}} = \frac{\sum_{i=1}^{n} \left[\binom{n}{i} \hat{\overline{p}}^{i} \hat{\overline{q}}^{n-i} (\alpha_i) \left(\frac{i}{n}\right) \right]}{\hat{\overline{w}}},$$

$$\hat{\overline{q}} = \frac{\sum_{i=0}^{n-1} \left[\binom{n}{i} \hat{\overline{p}}^{i} \hat{\overline{q}}^{n-i} (\beta_i) \left(\frac{n-i}{n}\right) \right]}{\hat{\overline{w}}}.$$

Note we require $0 < \hat{\overline{p}}, \hat{\overline{q}} < 1$. Rearranging the above gives:

$$n\hat{\overline{w}} = \sum_{i=1}^{n} \left[\binom{n}{i} \hat{\overline{p}}^{i-1} \hat{\overline{q}}^{n-i} (\alpha_i)(i) \right],$$

and

$$n\hat{\overline{w}} = \sum_{i=0}^{n-1} \left[\binom{n}{i} \hat{\overline{p}}^{i} \hat{\overline{q}}^{n-i-1} (\beta_i)(n-i) \right].$$

Therefore,

$$\sum_{i=1}^{n} \left[\binom{n}{i} \hat{\overline{p}}^{i-1} \hat{\overline{q}}^{n-i} (\alpha_i)(i) \right] = \sum_{i=0}^{n-1} \left[\binom{n}{i} \hat{\overline{p}}^{i} \hat{\overline{q}}^{n-i-1} (\beta_i)(n-i) \right],$$

$$\sum_{i=0}^{n-1}\left[\binom{n}{i+1}\hat{\overline{p}}^{i}\hat{\overline{q}}^{n-i-1}(\alpha_{i+1})(i+1)\right]-\sum_{i=0}^{n-1}\left[\binom{n}{i}\hat{\overline{p}}^{i}\hat{\overline{q}}^{n-i-1}(\beta_{i})(n-i)\right]=0,$$

$$\sum_{i=0}^{n-1}\left[\frac{n!}{i!(n-i-1)!}\hat{\overline{p}}^{i}\hat{\overline{q}}^{n-i-1}(\alpha_{i+1}-\beta_{i})\right]=0. \qquad (A3)$$

However, if for all $i \in \{0,1,2, \ldots n-1\}$, $\alpha_{i+1} < \beta_i$ (Class I fitness structure) or $\alpha_{i+1} > \beta_i$ (Class II fitness structure), then equation ($A3$) cannot hold and there exist no interior equilibria.

Given the current frequency of altruists, $\overline{p}$, the frequency of altruists in the next generation, $\overline{p}'$, is computed using equation ($A1$). For notational simplicity, let $\overline{p}' = g(\overline{p})$. Thus, $g(0) = 0$, $g(1) = 1$, and $0 < g(\overline{p}) < 1$ for $0 < \overline{p} < 1$. It can be shown that:

$$\left.\frac{\partial g}{\partial \overline{p}}\right|_{\overline{p}=0} = \frac{\alpha_1}{\beta_0}, \qquad (A4)$$

and

$$\left.\frac{\partial g}{\partial \overline{p}}\right|_{\overline{p}=1} = \frac{\beta_{n-1}}{\alpha_n}. \qquad (A5)$$

Under a Class I fitness structure, $\alpha_{i+1} < \beta_i$. This implies that $\left.\frac{\partial g}{\partial \overline{p}}\right|_{\overline{p}=0} < 1$ and $\left.\frac{\partial g}{\partial \overline{p}}\right|_{\overline{p}=1} > 1$. Since there are no interior equilibria under Class I and $g(\overline{p})$ is a *continuous* map, we have $g(\overline{p}) < \overline{p}$ for $0 < \overline{p} < 1$ (i.e., altruists always decrease in frequency). Therefore, given a Class I fitness structure, selfishness will invade, fix and remain stable to invasion by altruism.

Under a Class II fitness structure, $\alpha_{i+1} > \beta_i$. This implies that $\left.\frac{\partial g}{\partial \overline{p}}\right|_{\overline{p}=0} > 1$ and $\left.\frac{\partial g}{\partial \overline{p}}\right|_{\overline{p}=1} < 1$. Since there are also no interior equilibria under Class II, $g(\overline{p}) > \overline{p}$ for $0 < \overline{p} < 1$ (i.e., altruists always increase in frequency). Therefore, given a Class II fitness structure, altruism will invade, fix and remain stable to invasion by selfishness.

Notes

[1] See Barrett and Godfrey-Smith (2002) for a more detailed discussion of different forms of pluralism in this area.

[2] Relatives of this "gestalt-switching" form of pluralism have also been proposed or hinted at by various other authors, including Maynard Smith (1987) and Waters (1991). Dawkins (1982) uses the metaphor of a Necker cube, to describe a kind of gestalt-switching pluralism in the context of his defense of the "gene's eye view." He does not apply the idea to group selection, and Dawkins' book is really more a defense of the gene's eye view than a defense of

pluralism. Sterelny and Kitcher (1988) present a more genuinely pluralist version of Dawkins' gene's eye view on evolution.

[3] The life cycle that we outline here follows the basic trait-group model proposed by Wilson (1975) and Hamilton (1975). To keep things as general as possible, we are being intentionally vague about where reproduction (or "copying") occurs in the life cycle. If the (group-dependent) fitness of an individual gives its number of offspring, then differential reproduction is the instrument of (fertility) selection. Alternatively, group composition could determine the probability of individual survival, in which case, (non-differential) reproduction may take place after (viability) selection – perhaps after the groups have dissolved. Lastly, as we shall see in the section on the brain worm, reproduction may play a critical role in the *formation* of groups before a selective episode. In such a case, the groups will generally consist of related individuals before (kin) selection occurs.

[4] The use of the term "particle" appears in Hamilton (1975).

[5] If the particle is a gene, then the term "broad-sense individualism" appears to be a misnomer, because the fitness parameters describe context-dependent *allelic* fitnesses with the diploid organism forming the *group*. To avoid confusion, we will use the term "contextual perspective" when considering population genetic systems in which fitnesses are assigned to the alleles. (However, note that the label "broad-sense individualism" could be used for any bi-level system as long as we associate the term "individual" with the lower-level replicating entity – i.e., the particle.)

[6] The distinction introduced here between contextual and collective approaches is roughly the same as the distinction between "conditional and conjunctive forms of accounting" used in Godfrey-Smith and Lewontin (1993), to address similar issues.

[7] Many theoretical biologists investigating selection in structured populations have used this individualist perspective in their models (see Cohen and Eshel 1976, Matessi and Jayakar 1976, Uyenoyama and Feldman 1980, and Matessi and Karlin 1984, as well as Wilson 1975, 1977, 1980, 1990).

[8] In cases where different individual types have different probabilities of surviving to reproduce (i.e., some form of viability selection), group productivity is taken to be the *expected* number of copies produced by all individuals within the group.

[9] Price (1972) discusses group selection using a similar framework. In his system, groups are labeled with an index, say j. Thus, n_j and p_j represent size and $\boldsymbol{A}$'s frequency, respectively, in group j before selection. Price uses primes to denote the state of the system after selection; consequently, n'_j and p'_j represent the same quantities after selection (see also Wade 1985; Dugatkin and Reeve 1994; Sober and Wilson 1998). If the index j is taken to be the number of $\boldsymbol{A}$'s in the group before selection, then, within our framework, $n_j = n$, $p_j = j/n$, $n'_j = \pi_j$ and $p'_j = \phi_j$. We choose to use new parameter names (rather than primes) to emphasize that these quantities form the *fitness structure* of the population.

[10] Since α_0 and β_n are undefined, we cannot use equations (1) and (2) for pure groups. However, we must have $\pi_0 = n\beta_0$, $\pi_n = n\alpha_n$, $\phi_0 = 0$, and $\phi_n = 1$. Note that for ϕ_i to be defined for all i, the following must hold: $\beta_0 > 0$, $\alpha_n > 0$, and, for all $i \in \{1,2,\dots n-1\}$, either α_i or β_i must be strictly positive.

[11] In the rest of this paper, we will number equations and inequalities with "a" and "b" when the same equation is written from both the individualist and multi-level perspectives. When we refer to the equation by its number alone (without the "a" or "b") then we are referring to the general relation expressed in the equation rather than its individualist or multi-level realization.

[12] Although we are considering the group frequency distribution as a "given," it is extremely

interesting to contemplate the evolution of the group frequency distribution. Biological processes influence the manner by which individuals come together (e.g., reproduction within groups before a selective episode, recognition systems that allow individuals to exclude certain types from groups, etc.). Thus, in reality, the group frequency distribution not only influences the evolutionary outcome of selection in structured populations, but also is itself influenced *by* selection. Wilson and Sober (1989) and Skyrms (1994) address this point nicely.

[13] See Chapter 4 of Lloyd (1988) for a good review of theoretical models addressing population structure.

[14] The essential components to understanding selection in structured populations outlined here generalize Sober's (1992) claim that the evolution of altruism in structured populations "depends on the costs and benefits of the behaviors considered and the degree of correlation that obtains among the interacting individuals." The more general statement would be that evolution in a structured population depends on the fitness structure (which determines the "payoffs" to different types in different contexts) and the group frequency distribution (which describes how interacting individuals come together).

[15] The exception can be found in the second row of Figure 3.

[16] A related visual framework appears in Charnov and Krebs (1975) and Wilson (1977, 1980). However, these authors do not use the graphs to distinguish between Class I and Class II structures. In a discussion of the relationships between evolution and deliberation, Sober (1998) does visually make the Class I/Class II distinction.

[17] Last member analysis is a heuristic device related to an argument used by Nunney (1985). He considers a focal individual and defines the remaining $n - 1$ members of its group as its "neighborhood." If groups form at random in an infinite population, then both a focal ***A*** and a focal ***B*** see the same neighborhoods with the same probabilities. In equations (12) and (13), the term i in the subscripts *defines the neighborhood* (i.e., the focal individual encounters i altruists). While Class I and Class II relations are comparisons of fitness between different types over *different groups*, these relations compare the fitness of different types over the *same neighborhoods*. Given a Class I structure, ***B*** has the advantage in any given neighborhood (read relation (12) with i altruists filling the $n - 1$ member neighborhood) and with a Class II structure, ***A*** has the advantage in any given neighborhood (read relation (13) similarly). Given that the two types see the same neighborhoods with the same probabilities, the M-J proposition results. To get at this rigorously, let us say that we pick an individual at random from the structured population. We let I_A be the event that the chosen individual is type ***A***, I_B be the event that the selected individual is type ***B***, and N_i be the event that the individual possesses a neighborhood with i altruists. As we show in our companion paper, random group formation guarantees $Pr\{\mathrm{N}_i|\mathrm{I}_A\} = Pr\{\mathrm{N}_i|\mathrm{I}_B\} = Pr\{\mathrm{N}_i\}$. Using Skyrms (1994) terminology, random group formation entails that the JEFFREY expected utility (which uses conditional probabilities) equals the SAVAGE expected utility (which uses unconditional probabilities). This is of course not true if there is non-random formation of groups. Skyrms gives a general treatment of correlation in structured populations in his discussion of the relationship between decision theory and evolutionary theory. Last member analysis is also closely related to what Sober (1998) calls the "heuristic of personification" (or HOP). HOP claims that a certain trait will evolve over alternatives if and only if a rational agent, interested in maximizing fitness, would choose that trait over the alternatives. In his discussion of the Prisoner's Dilemma, Sober shows that HOP can fail when there is non-random formation of groups.

[18] For instance, Nunney (1985) suggests that Wilson's "weak altruism" be considered as a case of "benevolence," while Grafen (1984) prefers the phrase "self-interested refusal to be spiteful."

[19] Wilson (1990) explores this graphical representation of the multi-level framework.

[20] Note that if one uses conditions (8), (10), and (11) to define altruism, then the visual framework of the individualist perspective is actually more helpful than the visuals from the multi-level perspective.

[21] Last member analysis aids understanding of this result. If your partner is a defector, you should defect (since $\alpha_1 < \beta_0$) and if your partner is a cooperator, you should still choose to defect (since $\alpha_2 < \beta_1$). These comparisons can also be made using Table 2. In the table, we restrict attention to one of the two columns (i.e., "fixing" what the partner is doing). Then we seek the row with the higher payoff. As can be seen in Table 2, despite which column we condition on, this "highest payoff row" will always correspond to the focal individual defecting.

[22] The case of extreme clumping here can be seen as an example of Skyrm's (1994) claim that fixation on a strictly efficient type (a type, when self-matched, that does better than any other self-matched type) is a globally stable equilibrium when there is extreme positive correlation in pairing (i.e., every individual only pairs with another of its type).

[23] This result is particularly easy to interpret using last member analysis. If your partner is a sneaker, you should choose to be an investor (since the expected payoff is higher: $\alpha_1 > \beta_0$). Likewise, if your partner is an investor, you should still choose to be an investor (since the expected payoff is higher: $\alpha_2 > \beta_1$).

[24] An evolutionary stable strategy is generally defined in populations where groups are formed *randomly*. In his lucid discussion of adaptive ratifiability, Skryms (1994) develops a generalization of the ESS concept to deal with cases of non-random group formation (in our case, where $f(t)$ is *not* a binomial distribution).

[25] In reality, the dove is, at best, a "borderline" altruist using Sober and Wilson's definition, since $\pi_1 = \pi_2$ (as opposed to $\pi_1 < \pi_2$). However, if we imagine that there is a small cost to the hawk in *preparing* to fight for the resource, then $\pi_1 < \pi_2$ and the dove is an altruist. We will ignore this minor issue here.

[26] As mentioned above, the term "individualist" is confusing in the case of population genetics. This is because the gene is playing the part of the "individual" and the diploid organism is the group. To avoid confusion, we will use the more general term "contextual" for the remainder of this section.

[27] To avoid confusion, we must point out that in labeling α_i and β_i as fitnesses of alleles, we are diverging from the definition of "allelic" or "genic" fitness found in authors such as Sober and Lewontin (1982), Sober (1984, 1992), Maynard Smith (1987), Lloyd (1988) and Godfrey-Smith and Lewontin (1993). These authors define the fitness of an allele as its *marginal* fitness rather than as its context-dependent fitness within a specific diploid organism. For instance, in their scheme the fitness of the ***A*** allele is given most generally by $\{(1/2)\ f_1(t)\alpha_1 + f_2(t)\alpha_2\}/\{(1/2)\ f_1(t) + f_2(t)\}$, where $f_1(t)$ and $f_2(t)$ are the frequencies of ***AB*** heterozygotes and ***AA*** homozygotes, respectively (see our companion paper). This reduces to the standard marginal $\overline{q}(t)\alpha_1 + \overline{p}(t)\alpha_2$, given that $f(t)$ is Bin $(\overline{p}(t), 2)$ (i.e., given that random mating occurs). Our treatment of genic/allelic fitnesses here is similar to Waters' treatment (1991) and even earlier Williams' treatment (1966); an allele is assigned fitnesses for each of the genotypic environments in which it can appear (see also Sterelny and Kitcher 1988). Within our framework, the genotype fitness cannot be treated as identical to an allelic fitness in context, because the genotype fitness is no more than a π_i term, whereas the allelic fitness in context includes both π_i and ϕ_i terms (this is most clear in the discussion over meiotic drive).

[28] Wilson (1990) discusses some of the cases shown in Figure 8. However, he employs the visual representation of the multi-level framework, where evolutionary outcomes are more

difficult to distinguish visually. The ease of visually discriminating evolutionary outcomes is a major heuristic advantage for the contextual perspective.

[29] Maynard Smith (1987) compares the Hawk-Dove game to a single locus diallelic population genetic system with overdominance. While the dynamical *behavior* of Hawk-Dove games and overdominant systems are very similar, the Hawk-Dove *fitness structure* is actually slightly different than that of overdominance. Specifically, while both alleles of the heterozygote claim the same fitness (given Mendel's rules), hawks have a higher fitness than doves in a mixed pair. Thus, the structure of the Hawk-Dove game that we consider here is actually closer to the meiotic drive system in Figure 8d.

[30] The general availability of these two modes of analysis and the reasons why one mode will sometimes be preferable to the other, are also discussed in section 5 of Godfrey-Smith and Lewontin (1993).

References

Alexander, R.D. and Borgia, G.: 1978, 'Group Selection, Altruism, and the Levels of Organization of Life', *Annual Review of Ecology and Systematics* **9**, 449–474.

Axelrod, R.: 1984. *The Evolution of Cooperation*, Basic Books.

Axelrod, R. and Hamilton, W.D.: 1981, 'The Evolution of Cooperation', *Science* **211**, 1390–1396.

Barrett, M. and Godfrey-Smith, P.: 2002, 'Group Selection, Pluralism, and the Evolution of Altruism: A Review of Unto Others', *Philosophical and Phenomenological Research*.

Buss, L.: 1987, *The Evolution of Individuality*, Princeton University Press, Princeton.

Charnov, E.L. and Krebs, J.R.: 1975, 'The Evolution of Alarm Calls: Altruism or Manipulation?', *American Naturalist* **109**, 107–112.

Cohen, D. and Eshel, I.: 1976, 'On the Founder Effect and the Evolution of Altruistic Traits', *Theoretical Population Biology* **10**, 276–302.

Dawkins, R.: 1982, *The Extended Phenotype*, Oxford University Press, Oxford.

Dugatkin, L.A. and Reeve, H.K.: 1994, 'Behavioral Ecology and Levels of Selection: Dissolving the Group Selection Controversy', *Advances in the Study of Behavior* **23**, 101–133.

Eshel, I.: 1972, 'On the Neighbor Effect and the Evolution of Altruistic Traits', *Theoretical Population Biology* **3**, 258–277.

Godfrey-Smith, P. and Lewontin, R.: 1993, 'The Dimensions of Selection', *Philosophy of Science* **60**, 373–395.

Grafen, A.: 1984, 'Natural Selection, Kin Selection, and Group Selection', in J.R. Krebs and N.B. Davies (eds.), *Behavioral Ecology: An Evolutionary Approach*, Sinauer, Sunderland, pp. 62–84.

Hamilton, W.D.: 1975, 'Innate Social Aptitudes in Man: An Approach from Evolutionary Genetics', in R. Fox (ed.), *Biosocial Anthropology*, Wiley, New York, pp. 133–155.

Karlin, S. and Matessi, C.: 1983, 'Kin Selection and Altruism', *Proceedings of the Royal Society, London B* **219**, 327–353.

Keller E.F. and Lloyd, E.A.: 1992, *Keywords in Evolutionary Biology*, Harvard University Press, Cambridge.

Lewontin, R.C.: 1969, 'The Meaning of Stability', *Diversity and Stability in Ecological Systems, Brookhaven Symposia in Biology* **22**, 13–24.

Lloyd, E.A.: 1988, *The Structure and Confirmation of Evolutionary Theory*, Princeton University Press, Princeton.

Matessi, C. and Jayakar, S.D.: 1976, 'Conditions for the Evolution of Altruism under Darwinian Selection', *Theoretical Population Biology* **9**, 360–387.

Matessi, C. and Karlin, S.: 1984, 'On the Evolution of Altruism by Kin Selection', *Proceedings of the National Academy of Sciences, U.S.A.* **81**, 1754–1758.

Maynard Smith, J.: 1964, 'Group Selection and Kin Selection', *Nature* **201**, 1145–1147.

Maynard Smith, J.: 1976, 'Group Selection', *The Quarterly Review of Biology* **51**, 277–283.

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

Maynard Smith, J.: 1987, 'How to Model Evolution', in J. Dupre (ed.), *The Latest on the Best: Essays on Evolution and Optimality*, MIT Press, Cambridge, pp. 119–131.

Maynard Smith, J. and Szathmáry, E.: 1995, *The Major Transitions in Evolution*, Oxford University Press, Oxford.

Michod, R.E.: 1999, *Darwinian Dynamics: Evolutionary Transitions in Fitness and Individuality*, Princeton University Press, Princeton.

Nunney, L.: 1985, 'Group Selection, Altruism, and Structured-deme Models', *American Naturalist* **126**(4), 212–230.

Price, G.R.: 1970, 'Selection and Covariance', *Nature* **227**, 520–521.

Price, G.R.: 1972, 'Extension of Covariance Selection Mathematics', *Annals of Human Genetics* **35**, 485–490.

Skyrms, B.: 1994, 'Darwin Meets *The Logic of Decision*: Correlation in Evolutionary Game Theory', *Philosophy of Science* **61**, 503–528.

Sober, E.: 1984, *The Nature of Selection: Evolutionary Theory in Philosophical Focus*, University of Chicago Press, Chicago.

Sober, E.: 1992, 'The Evolution of Altruism: Correlation, Cost, and Benefit', *Biology and Philosophy* **7**, 177–187.

Sober, E.: 1998, 'Three Differences Between Evolution and Deliberation', in P. Danielson (ed.), *Modeling Rationality, Morality and Evolution*, Oxford University Press, Oxford, pp. 408–422.

Sober, E. and Lewontin, R.C.: 1982, 'Artifact, Cause and Genic Selection', *Philosophy of Science* **49**, 157–180.

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

Sterelny, K. and Kitcher, P.: 1988, 'The Return of the Gene', *Journal of Philosophy* **85**, 339–361.

Sterelny, K.: 1996, 'The Return of the Group', *Philosophy of Science* **63**, 562–584.

Uyenoyama, M.K. and Feldman, M.W.: 1980, 'Theories of Kin and Group Selection: A Population Genetics Perspective', *Theoretical Population Biology* **17**(5), 380–414.

Uyenoyama M.K. and Feldman, M.W.: 1992, 'Altruism: Some Theoretical Ambiguities', in E. Fox Keller and E.A. Lloyd (eds.), *Keywords in Evolutionary Biology*, Harvard University Press, Cambridge, pp. 34–40.

Wade, M.J.: 1978, 'A Critical Review of the Models of Group Selection', *The Quarterly Review of Biology* **53**(2), 101–114.

Wade, M.J.: 1985, 'Soft Selection, Hard Selection, Kin Selection and Group Selection', *American Naturalist* **125**(1), 61–73.

Waters, C.K.: 1991, 'Tempered Realism About the Force of Selection', *Philosophy of Science* **58**, 553–573.

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

Williams, G.C. and Williams, D.C.: 1957, 'Natural Selection of Individually Harmful Social Adaptations Among Sibs with Special Reference to Social Insects', *Evolution* **11**, 32–39.

Wilson, D.S.: 1975, 'A Theory of Group Selection', *Proceedings of the National Academy of Sciences, U.S.A.* **72**(1): 143–146.

Wilson, D.S.: 1977, 'Structured Demes and the Evolution of Group-Advantageous Traits', *American Naturalist* **111**, 157–185.

Wilson, D.S.: 1980, *The Natural Selection of Populations and Communities*, Benjamin/ Cummings, Menlo Park.

Wilson, D.S.: 1983, 'The Group Selection Controversy: History and Current Status', *Annual Review of Ecology and Systematics* **14**, 159–187.

Wilson, D.S.: 1990, 'Weak Altruism, Strong Group Selection', *Oikos* **59**(1), 135–140.

Wilson, D.S.: 1997, 'Altruism and Organism: Disentangling the Themes of Multilelvel Selection Theory', *American Naturalist* **150**, S122–S134.

Wilson, D.S. and Sober, E.: 1989, 'Reviving the Superorganism', *Journal of Theoretical Biology* **136**, 337–356.

Wright, S.: 1945, 'Tempo and Mode in Evolution: A Critical Review', *Ecology* **26**(6), 415–419.