19

THEORETICAL AND EXPERIMENTAL APPROACHES TO THE EVOLUTION OF ALTRUISM AND THE LEVELS OF SELECTION

Benjamin Kerr

All mankind . . . is one volume; when one man dies, one chapter is not torn out of the book, but translated into a better language.

JOHN DONNE, MEDITATION XVII

One of the central themes in Donne's meditation is the interconnectivity between human beings. Our lives are not stand-alone chapters from an edited volume (like the one you are reading), but more like chapters from an elaborate novel, each setting the stage for chapters to come while simultaneously depending on chapters already read. Biological systems, from subcellular biochemical networks to multispecies food webs, display striking forms of interconnectivity in their parts. How does the theory of biological evolution by natural selection handle this interconnectivity? The simplest description of natural selection starts by *ignoring* interdependence. For simplicity, individuals in a population are assumed to affect neither each other's fitness nor the form of their environment. As an example of the logic, consider the giant anteater. A standard story of natural selection would maintain that anteaters with longer, stickier tongues have been selected because they are able to gather more ants and termites. After all, it is these food resources that improve survival and the production of (long-sticky-tongued) progeny. However, tongue length of one anteater is assumed *not* to affect the fitness of another, and tongue stickiness is assumed *not* to influence the behavior or morphology of the insect prey. In effect, each anteater is treated as "an island, entire of itself" (Donne 1624/1839).

Under the "organism-as-island" incarnation of natural selection, there is an extreme premium placed on personal fitness. An individual with a phenotype that best solves current environmental challenges "cashes in" by earning the highest personal fitness. Within this perspective, what are we to make of the existence of individuals that appear to sacrifice personal fitness to improve the fitness of others? Of course, the answer is that such behavior is fundamentally backward and should be swiftly eradicated by the action of natural selection. Individuals that exhibit restraint and self-sacrifice for the benefit of others are foolishly eroding the precious commodity of personal fitness and should be undone by selfish counterparts. Nonetheless, the biological world is filled with examples of altruistic leanings, from slime mold cells that sacrifice themselves to form the somatic stalks of fruiting bodies (on which reproductive spores sit) to the nonreproductive workers of eusocial insect colonies.

A second reaction to the existence of altruists is that the organism-as-island version of natural selection is misleading when considering such behaviors, as compared with simpler behaviors, such as locomotion in an activity wheel (Rhodes and Kawecki this volume; Swallow et al. this volume). By its very nature, altruism is a social activity. The fitness of an individual depends on the behavior of others. Social organisms do not simply solve the challenges of an external environment—they *are* the selective environment (or at least part of it). The interconnectivity ignored by the organism-as-island model is now front and center. Indeed, how organisms are connected turns out to be *the* critical issue. Most explanations for the evolution of altruism depend on altruists disproportionately finding themselves in the company of other altruists. For instance, organisms may interact with relatives (kin selection), organisms may condition their own altruistic behavior on the cooperative behavior of partners (reciprocal altruism), or organisms may exclude

non-cooperating members from their pool of interactors (policing). If altruists are able to associate preferentially with other altruists, then the personal fitness cost can be viewed as the price of admission to a beneficial social milieu.

Sometimes cooperative interactions between the members of a group lead to a type of functionality and cohesiveness at the group level. For instance, eusocial insect colonies have been called "superorganisms" to underline the degree of interconnectivity between insects within the colony (e.g., Emerson 1939). However, does selection "act" at the level of groups in such cases? This question brings us to a subject that has occupied biologists and philosophers of biology for nearly half a century. While the existence of a biological hierarchy is uncontroversial (genes do line up on chromosomes, chromosomes are em-bedded in cells, cells do make up organisms, and organisms do interact in social groups), the level(s) targeted by natural selection is an extremely controversial subject. Indeed, a large literature has focused on this issue of levels of selection (e.g., Wynne-Edwards 1962; Maynard Smith 1964; Williams 1966; Price 1972; Uyenoyama and Feld-man 1980; Sober and Wilson 1998; Michod 1999).

A particularly contentious part of this levels of selection debate concerns whether selection can operate on groups of organisms. Perhaps the best-known proponent of group-level selection was the Scottish ornithologist V. C. Wynne-Edwards. Wynne-Edwards interpreted specific social behaviors of individuals as group-level adaptations. For instance, he viewed animal territoriality as a mechanism of spacing out individuals such that the population would not overexploit critical resources (Wynne-Edwards 1962). The idea was that populations without mechanisms to curb overexploitation would run a higher risk of extinction, and these "short-sighted" populations would be supplanted by other populations that possessed mechanisms of control. Central to Wynne-Edwards's argument was that the process of natural selection could operate on a population of groups.

A significant blow was dealt to Wynne-Edwards's version of group selection with the publication of Williams's classic *Adaptation and Natural Selection* (1966). Williams admitted that group selection was a logical possibility; indeed, he even cited one study on house mice (Lewontin and Dunn 1960) that provided what he called "convincing evidence" of its operation. Furthermore, he claimed that group selection was *required* to explain group-level adaptation. He simply felt that group-level adaptations did not, as a rule, exist. First, Williams argued that the conditions necessary for the operation of group selection were rarely realized (e.g., turnover of groups relative to the turnover of individuals they contained was too slow, migration between groups was too high to maintain intergroup variation, numbers of groups within metapopulations was too low, etc.). Second, Williams argued that in many cases a simpler explanation of the evolu-tionary origin of a social behavior entailed adaptation at the individual level. For in-stance, territoriality in animals could evolve because the *individual* territory holder is able to secure more resources *for itself*. Invoking Occam's razor, Williams argued that the simpler explanation was preferable. It is difficult to overstate the impact of

Williams's book on the levels of selection controversy (Wilson 1983; Sober and Wilson 1998). Due in no small part to this publication, group selection became taboo in the mainstream of evolutionary biology, and indeed, this is still the state of affairs in many quarters.

Altruism is a recurring character in the group selection debate. Williams (1966) suggested that testing for group selection should involve "finding adaptations that promote group survival but are clearly neutral or detrimental to individual reproductive survival in within-group competition."[1] Wynne-Edwards (1962) specifically invoked group selection to explain how "short-term advantages of the individual [that] undermine the safety of the race" could be eliminated. The so-called "forces" of individual and group selection appear to be opposed in the evolution of prosocial, self-sacrificial behavior. Specifically, altruists are always at a relative disadvantage within groups; however, groups with more altruists are more productive or long-lived. Some authors maintain that both individual and group selection operate simultaneously in such cases (Wilson 1983; Sober and Wilson 1998). From this multilevel perspective, the evolution of altruism depends on the relative strengths of these opposing, concurrent forces: within-group selection for selfishness and between-group selection for altruism.

Part of Williams's argument against group selection for altruism was that he felt that the "within-group force" was relatively strong. He illustrated this point with an example of how a population of robins that exercised restraint in the use of common resources could be invaded by a "selfish" variant that was less prudent (Williams 1971). Williams concluded that an inexorable force for selfishness within groups would lead to the decrease of altruists globally. Figure 19.1 is a rejoinder to Williams's logic that shows that altruist frequency can globally increase even if the altruist frequency decreases within *every* group! This can occur if there is positive covariance between group output and altruist frequency within the group (i.e., more altruistic groups are more productive—see Sober and Wilson 1998 for a full discussion of this point). The example in figure 19.1 illustrates that the biological details of the system (e.g., how groups come together and how reproductive output depends on group composition) are crucial to determining how likely it will be that unselfish behavior evolves.

There was theoretical interest in exploring the conditions favoring the evolution of altruism before Williams's famous publication (e.g., Wright 1945; Hamilton 1964; Maynard Smith 1964). However, following Williams's critique, a large set of theoretical papers appeared exploring the evolution of altruism (Eshel 1972; Wilson 1975, 1977; Cohen and Eshel 1976; Matessi and Jayakar 1976; Uyenoyama and Feldman 1980; Karlin and Matessi 1983; Matessi and Karlin 1984). Most of these models identify conditions (e.g., the form of population structure) in which altruism can evolve. Because altruism is defined differently in different models, these conditions do not always coincide exactly. Furthermore, while some evolutionary explanations for altruism are pitched as individualistic alternatives to group selection explanations (e.g., kin selection: Hamilton 1963; reciprocal altruism: Axelrod and Hamilton 1981), other authors (and sometimes
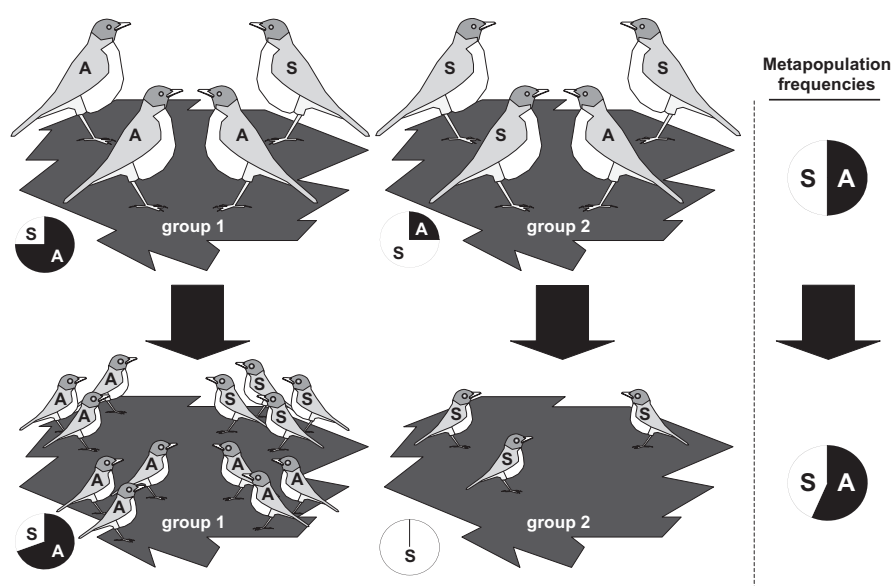
FIGURE 19.1

An illustration of the paradoxical result that altruists can decrease in frequency within groups but increase in frequency globally. Imagine two types of robins, self-restrained altruists and unrestrained selfish types. Suppose that the population of robins is split into two groups, where group 1 has three altruists and one selfish type and group 2 has three selfish types and a single altruist. Thus, the altruist frequency before selection occurs is 75 percent in group 1, 25 percent in group 2, and 50 percent globally (see pie charts). Over the selective episode, altruists have positive effects on their group mates and offspring are produced. We see that selfish types increase in frequency within both groups, but altruists increase in frequency globally (see pie charts). The reason behind this apparent paradox is that the group 1 (which started with more altruists) was more productive than group 2.

the same authors!) have found fundamental similarities between these "alternative" explanations (e.g., between kin selection and group selection: Hamilton 1975; Wade 1980c; Sober and Wilson 1998; Lehmann et al. 2007). Thus, for the casual observer, the precise connections between different theoretical ideas and the specific conditions favoring altruism may seem unclear.

In this chapter, I review some of the theoretical approaches to the study of altruism in an attempt to clarify some basic concepts. While I focus on the models and statistical tools that have appeared in the levels of selection literature, I attempt to make connections between different theoretical approaches in the process. With these theoretical ideas as a backdrop, I then proceed to describe laboratory and field experiments that have addressed the evolution of different forms of altruism, some directly and some indirectly. I end by discussing some philosophical issues in the debate over the levels of selection and the general impact of theoretical and empirical results for this debate.

THEORETICAL BACKGROUND

A SIMPLE TRAIT GROUP MODEL

Imagine an infinite population filled with two types of individuals, **A** types and **B** types. The **A** individuals are altruists, sacrificing personal fitness while increasing the fitness of others. The **B** individuals are selfish relative to the **A** types; they do not exhibit self-sacrifice. These individuals undergo the following life cycle: (1) individuals form groups of size n, (2) fitness-affecting social interactions (i.e., altruistic action) occur inside these groups and offspring are asexually produced, (3) the adults die, and (4) the groups dissolve. The offspring then form new groups and the cycle continues. This is a simple "trait group" scenario (Wilson 1975, 1980) used in several of the early models exploring the evolution of altruism (Cohen and Eshel 1976; Matessi and Jayakar 1976; Wilson 1977; Uyenoyama and Feldman 1980; Matessi and Karlin 1984). One noteworthy assumption of this model is that *discrete* groups are formed. Thus, the groups form well-defined entities and one question at hand is whether a (meta)population of these entities experiences a selective process.[1]

At generation t, the population-wide frequencies of **A** and **B** types (before groups are formed) are given by $\bar{p}(t)$ and $\bar{q}(t)$, respectively. The frequency of groups with i **A** types at generation t is given by $f_i(t)$. Note groups could form randomly, in which case the group frequency distribution would be binomial:

$$f_i(t) = \binom{n}{i}[\bar{p}(t)]^i[\bar{q}(t)]^{n-i} \tag{1}$$

for all $i \in \{0, 1, 2, \ldots n\}$. However, groups could also form nonrandomly (discussed later). Once groups have formed, the fitness of an **A** type in a group with i **A** types is given by α_i, while the fitness of a **B** type in a group with i **A** types is given by β_i.

DEFINING ALTRUISM

There are different definitions of altruism with interesting connections between them (see Nunney 1985; Wilson 1990; Kerr et al. 2004). In the context of our trait group model, we start with a popular definition of altruism, given by the following relations:

$$\alpha_{i+1} - \beta_i < 0 \tag{2}$$

$$(\alpha_{i+1} - \alpha_i)i + (\beta_{i+1} - \beta_i)(n - i - 1) > 0 \tag{3}$$

for all $i \in \{0, 1, 2, \ldots n - 1\}$. We set $\alpha_0 = \beta_n = 0$ for condition (3). To see the origin of conditions (2) and (3), consider a selfish **B** type in a group with i altruists. Suppose that this focal **B** individual switches types (a **B** $\rightarrow$ **A** conversion) so that there are $i + 1$ altruists in its group. The change in fitness of our focal individual is $\alpha_{i+1} - \beta_i$ and in order for altruism to be personally costly we require that this change is always negative (condition [2]). Now let's follow up on the change in fitness for the other $n - 1$ individuals (i altruists

and $n - i - 1$ selfish types) after the conversion of the focal. Their change in fitness is given by $(\alpha_{i+1} - \alpha_i)i + (\beta_{i+1} - \beta_i)(n - i - 1)$. In order for altruism to benefit *others,* we require that this change is always positive (condition [3]). Conditions (2) and (3) define "focal-complement altruism" (as costs are measured on a focal converting individual and benefits are measured on the complement of the focal).

A second definition of altruism requires the following conditions:

$$\alpha_i < \beta_i \qquad\qquad (4)$$

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

Condition (4) holds for all $i \in \{1, 2, \ldots n - 1\}$. Condition (5) holds for $i \in \{0, 1, 2, \ldots n - 1\}$, where we set $\alpha_0 = \beta_n = 0$. Condition (4) guarantees that in groups with both types, altruists have a lower fitness. Note that the cost of altruism is being gauged by *within-group* comparisons here. Condition (5) guarantees that the total reproductive output of the group increases with the fraction of altruists inside the group. Note that the benefit of altruism is measured by comparing the output of the *whole group* between groups that differ in altruist frequency. We term conditions (4) and (5) "multilevel altruism" (as costs are measured between individuals within groups and benefits are measured between groups). We note that neither focal-complement altruism nor multilevel altruism entails the other (Kerr et al. 2004).

STATISTICAL ASSOCIATION AND THE EVOLUTION OF ALTRUISM

We defined focal-complement altruism by focusing on fitness costs to the focal altruist and fitness benefits to the focal individual's $n - 1$ "neighbors" (the complement). Let $X(t)$ and $Y(t)$ be random variables giving the number of A types in the neighborhoods of a randomly chosen A type and B type, respectively, at generation t. If groups form randomly (i.e., equation [1] holds), then it can be shown:

$$\Pr\{X(t) = k\} = \Pr\{Y(t) = k\} = \binom{n - 1}{k}[\bar{p}(t)]^k[\bar{q}(t)]^{n-1-k} \qquad\qquad (6)$$

Thus, A and B individuals experience the same neighborhoods with the same probabilities if groups are randomly formed. This means that for any given generation, the selective environment experienced is, in a sense, constant across types. The altruist and selfish type are getting the same amount of average help from their neighbors; however, the altruist is giving up personal fitness, while the selfish type is not. Therefore, the logic of the organism-as-island version of natural selection would seem to apply. Specifically, it is as if different organisms are exposed to the same external selective environment and some "throw away" personal fitness. We expect those discarding fitness to be weeded out. Indeed, given random group formation, it can be proven (Cohen and Eshel 1976; Matessi and Jayakar 1976; Kerr and Godfrey-Smith 2002a) that condition (2) alone guarantees that altruists will be displaced by selfish types in our simple trait group framework.

For focal-complement altruism, we must have nonrandom formation of groups in order for altruists to stand a chance (this need not be the case for multilevel altruism). Specifically, altruists must tend to associate with other altruists (i.e., the group frequency distribution must be clumped). One way to measure clumping in the group frequency distribution is to ask how the actual variance in frequency of A types deviates from what the variance would have been if groups formed randomly (see Kerr and Godfrey-Smith 2002b). Let the actual variance in the frequency of A across groups at generation t be given by $\sigma_p^2(t)$, and let the variance in frequency of A if groups were to form randomly be $\nu^2(t) = \bar{p}(t)\bar{q}(t)/n$. Our "clumping index" is

$$\rho(t) = \frac{\sigma_p^2(t) - \nu^2(t)}{\nu^2(t)} \tag{7}$$

If groups actually form randomly, then $\sigma_p^2(t) = \nu^2(t)$, and the clumping index is zero. Large positive values of $\rho(t)$ correspond to frequency distributions with heavy tails, in which A types often encounter other A types and B types encounter other B types in their groups. Large negative values of $\rho(t)$ correspond to frequency distributions with heavy centers, in which many groups contain the same mixture of A and B types. Interestingly, the clumping index can also be written as

$$\rho(t) = E[X(t)] - E[Y(t)] \tag{8}$$

That is, this index also measures how many more altruistic neighbors an A type can expect in its group when compared to a B type.

Here we make use of our clumping index to illustrate the importance of association between altruists and to connect our framework to Hamilton's rule. We assume a simplified scenario in which each individual has a base fitness of z and each altruist provides a fitness benefit b to each of its $n - 1$ neighbors at a fitness cost c to itself. These assumptions give the following linear fitness functions:

$$\alpha_i = z - c + b(i - 1) \tag{9}$$

$$\beta_i = z + bi \tag{10}$$

When $c > 0$ and $b > 0$ (as assumed), then equations (9) and (10) satisfy conditions (2) and (3), and we are thus dealing with focal-complement altruism (incidentally, if $b(n - 1) - c > 0$, then the linear fitness functions also qualify as multilevel altruism). It can be shown (see Wilson 1980; Kerr and Godfrey-Smith 2002b) that altruists will increase in frequency if

$$\frac{b}{c} > \frac{1}{\rho(t)}. \tag{11}$$

Equation (11) states that the ratio of benefits to costs of altruism must be greater than the reciprocal of the degree of clumping. Our simple model makes some of the same assumptions that Hamilton made (e.g., additivity of fitness costs/benefits of altruism),

and equation (11) is structurally identical to Hamilton's famous rule for the increase of altruists (Hamilton 1963, 1964):

$$\frac{b}{c} > \frac{1}{r}, \tag{12}$$

where r is the coefficient of relatedness between actor and recipient. How is it that r and $\rho(t)$ are playing similar roles? Both of these quantities are measuring statistical association: how likely altruists are to interact with other altruists. The larger these measures of association, the greater the chances for altruistic behavior. Hamilton (1975) himself emphasized that it was association, rather than relatedness per se, that was critical: "It makes no difference if altruists settle with altruists because they are related . . . or because they recognize fellow altruists as such, or settle together because of some pleiotropic effect of the [altruistic] gene on habitat preference . . . correlation between interactants is necessary if altruism is to receive positive selection." More recent incarnations of Hamilton's rule (e.g., Queller 1985; Fletcher and Zwick 2006) emphasize that it is actually the positive association between altruistic genotypes and the helping *phenotypes* of interactants that is of most general relevance (which extends the application of Hamilton's rule to reciprocal altruism and interspecific mutualism—see Fletcher and Zwick 2006; Fletcher et al. 2006). Of course, interactions between relatives (as occurs when individual offspring are deposited in a nest or colony) may be a particularly common way to achieve this association (Nunney 1985).

THE FITNESS STRUCTURE AND CONTEXT FORMATION

The population-wide frequency of the altruistic type **A** at generation $t + 1$ is described by the following equation:

$$\overline{w}(t)\overline{p}(t + 1) = \sum_{i=1}^{n}\left(\frac{i}{n}\right)(\alpha_i)f_i(t), \tag{13}$$

where $\overline{w}(t)$ is average population-wide fitness of an individual at generation t:

$$\overline{w}(t) = \left\{\sum_{i=1}^{n}\left(\frac{i}{n}\right)(\alpha_i)f_i(t)\right\} + \left\{\sum_{i=0}^{n-1}\left(1 - \frac{i}{n}\right)(\beta_i)f_i(t)\right\}. \tag{14}$$

Thus, in order to predict change in the frequency of altruists, we must know at least two things: (1) the fitnesses of types in different social contexts (these are the α's and β's) and (2) the way that social contexts come to be (given by the f_i's). We label these two elements the *fitness structure* and *context formation*, respectively. The forms of both of these components will influence the prospects for altruism.

Equations (11) and (12) actually nicely separate terms representing the fitness structure (b and c) from terms representing context formation ($\rho(t)$ or r). We will see later that some experiments exploring evolution in structured populations can be categorized by whether they manipulate factors affecting the fitness structure or context formation or both.

STATISTICAL EQUATIONS AND THE LEVELS OF SELECTION

THE PRICE EQUATION

For simplicity, let us stick with the trait group model, but we will introduce slightly different notation. For each group of size n, let us arbitrarily number the individuals from 1 to n. Let w_{ij} be the fitness of the jth individual in a group with i **A** types (note that $w_{ij} = \alpha_i$ or $w_{ij} = \beta_i$). Let $w_{i\bullet}$ be the average fitness of an individual in a group with i **A** types (note that $w_{i\bullet} = (\alpha_i i + \beta_i(n - i))/n$). Let $\overline{w} = w_{\bullet\bullet}$ be the population-wide average individual fitness (given by equation [14]). Let p_{ij} be "the frequency of **A** types in the jth individual" in a group with i **A** types (note that $p_{ij} = 1$ if the jth individual is an **A** type and $p_{ij} = 0$ if the jth individual is a **B** type). Let $p_{i\bullet}$ be the frequency of **A** types in a group with i **A** types (thus, $p_{i\bullet} = i/n$). Finally, let $\overline{p} = p_{\bullet\bullet}$ be the population-wide frequency of **A** types, where we note

$$\overline{p}(t) = \sum_{i=1}^{n} \left(\frac{i}{n}\right) f_i(t). \tag{15}$$

Subtracting $\overline{w}(t)\overline{p}(t)$ from both sides of equation (13) and rearranging yields

$$\overline{w}\Delta\overline{p} = \mathrm{cov}_{f_i}(p_{i\bullet}, w_{i\bullet}) + E_{f_i}[w_{i\bullet}\Delta p_{i\bullet}], \tag{16}$$

where we have dropped time arguments from the quantities (each quantity refers to generation t except $\Delta\overline{p} = \overline{p}(t + 1) - \overline{p}(t)$ and $\Delta p_{i\bullet} = p_{i\bullet}(t + 1) - p_{i\bullet}(t)$), and the subscripts on cov and E indicate that these are weighted by the group frequency distribution. Although not strictly appropriate, we leave the subscripts on our variables in equation (16) and hereafter for clarity. Equation (16) is a manifestation of the famous Price equation (Price 1970). The covariance term is often taken to represent the effect of "between-group" selection (Price 1972; Hamilton 1975; Sober and Wilson 1998) measuring how group output co-varies with group composition. The expectation term is often taken to represent "within-group" (or individual-level) selection, measuring the (weighted) expected change in altruist frequency within groups over the selective episode.

If there is simple asexual reproduction without mutation, equation (16) can be rewritten as

$$\overline{w}\Delta\overline{p} = \kappa_{w_{i\bullet}p_{i\bullet}}\mathrm{var}_{f_i}(p_{i\bullet}) + E_{f_i}[\kappa_{w_{ij}p_{ij}}\mathrm{var}_j(p_{ij})], \tag{17}$$

where $\kappa_{w_{i\bullet}p_{i\bullet}}$ is the regression coefficient of average fitness of a group on altruist frequency within a group, and $\kappa_{w_{ij}p_{ij}}$ is the regression coefficient of individual fitness on individual type. Thus, the "group-level" term is nonzero only if there is variance in group composition and if there is a nonzero relationship between group output and its composition. And the "individual-level" term is nonzero only if there are both types in some of the groups and if there is a nonzero relationship between individual fitness and type within these same groups.

For instance, using the linear fitness functions given by equations (9) and (10), we find

$$\kappa_{w_{i\bullet}p_{i\bullet}} = b(n - 1) - c \tag{18}$$

$$\kappa_{w_{ij}p_{ij}} = -(b + c) \tag{19}$$

Since $b > 0$ and $c > 0$, then within any mixed group altruists are selected against (i.e., "individual-level" selection is said to work against altruists). However, if the total benefit provided by a single altruist to all its neighbors, $b(n - 1)$, outweighs the fitness cost of altruism, c, then more altruistic groups are more productive (i.e., "group-level" selection is said to work for altruists; see Wade 1980c). Whether altruists increase in frequency will be decided by the relative "strengths" of between-group selection for altruists (given by $\mathrm{cov}_{f_i}(p_{i\bullet}, w_{i\bullet}) = \kappa_{w_{i\bullet}p_{i\bullet}}\mathrm{var}_{f_i}(p_{i\bullet})$) versus within-group selection against altruists (given by $E_{f_i}[w_{i\bullet}\Delta p_{i\bullet}] = E_{f_i}[\kappa_{w_{ij}p_{ij}}\mathrm{var}_j(p_{ij})]$). For multilevel altruism (but not focal-complement altruism), Price's first term is guaranteed to be nonnegative, while Price's second term is guaranteed to be nonpositive. By defining altruism with conditions (4) and (5), the "between-group" term will never work against altruists, and the "within-group" term will never work for altruists.

Price's equation also illustrates the role for association in the evolution of altruism. As **A** types preferentially associate with other **A** types, the variance in **A** frequency across groups, $\mathrm{var}_{f_i}(p_{i\bullet})$, increases. Under certain conditions (e.g., multilevel altruism), this works for the evolution of altruism through the first term in Price's equation—namely, $\kappa_{w_{i\bullet}p_{i\bullet}}\mathrm{var}_{f_i}(p_{i\bullet})$. Note that $\sigma_p^2 = \mathrm{var}_{f_i}(p_{i\bullet})$, and we have already seen that a larger variance in altruist frequency across groups can improve the chances of the evolution of altruism with linear fitness functions (see equations [7] and [11]).

CONTEXTUAL ANALYSIS

Consider the following scenario (presented in alternative forms in Sober 1984; Nunney 1985; Heisler and Damuth 1987; Okasha 2006): groups of size n form, but there are no meaningful interactions between individuals. Thus, there is no altruism present, and **A** and **B** are seen as simply two different types of asocial individuals. Suppose that **A** always has two offspring and **B** always has a single offspring (i.e., $\alpha_i = 2$ and $\beta_i = 1$ for all relevant i). Then if there is variation in the frequency of **A** across groups, then Price's first term, $\mathrm{cov}_{f_i}(p_{i\bullet}, w_{i\bullet})$, will be positive. This is because groups that happen to have more **A** types are more productive. This is somewhat disheartening because Price's first term was supposed to capture group-level selection; yet here, where a description of pure individual selection seems apposite, the group-level term is nonzero. This appears to be a failure of the standard interpretation of Price's terms.

While Price's equation seems ill equipped to characterize this scenario, another statistical approach is ideal at handling it. Contextual analysis (Heisler and Damuth 1987; Goodnight et al. 1992; Okasha 2006) starts with the following linear regression model:

$$w_{ij} = \theta + \varphi p_{ij} + \Phi p_{i\bullet} + \varepsilon_{ij}. \tag{20}$$

Specifically, θ is the "base" fitness, φ is the partial regression coefficient giving the effect of individual type, p_{ij}, on individual fitness, w_{ij}, (controlling for group composition), and

Φ is the partial regression coefficient giving the effect of group composition, $p_{i\bullet}$, on individual fitness (controlling for individual type). The term ε_{ij} is the residual.

It can be shown that

$$\overline{w}\,\Delta\overline{p} = \Phi\,\mathrm{var}_{f_i}(p_{i\bullet}) + \varphi\,\mathrm{var}_{(f_{i,j})}(p_{ij}), \tag{21}$$

where the variance in individual types is given by

$$\mathrm{var}_{(f_{i,j})}(p_{ij}) = \sum_{i=0}^{n}\frac{f_i(t)}{n}\sum_{j=1}^{n}(p_{ij} - \overline{p})^2, \tag{22}$$

and the variance across groups in *A*'s frequency within groups, $\mathrm{var}_{f_i}(p_{i\bullet})$, is identical to its previous usage (see equation [17]). In equation (21), the first term is taken to give the effects of group composition on evolution (the "group-level" term), and the second term is taken to give the effects of individual type on evolution (the "individual-level" term). Again, for a nonzero effect in either term, both the partial regression coefficient and variance must be nonzero. For example, a group-level effect requires that group composition affects individual fitness (after controlling for individual type), $\Phi \neq 0$, and there must be variance in group composition, $\mathrm{var}_{f_i}(p_{i\bullet}) \neq 0$.

Revisiting our problematic case of constant *A* and *B* fitness, we see that the partial regression coefficient for group composition is zero. Given that $\Phi = 0$ there is no possibility for group-level effects. Thus, contextual analysis has yielded a result consistent with intuition about this case: evolution is wholly accounted for through individual-level effects.

Now let's return to the case of altruism and, in particular, to the linear fitness functions (9) and (10). Here, we can show that $\varphi = -(b + c)$ and $\Phi = bn$, and with $\theta = z$ and assuming $\varepsilon_{ij} = 0$, we have

$$w_{ij} = z - (b + c)p_{ij} + bnp_{i\bullet}. \tag{23}$$

As a quick check, by setting $p_{ij} = 1$ and $p_{i\bullet} = i/n$ in equation (23), we recover α_i from equation (9). Similarly, by setting $p_{ij} = 0$ and $p_{i\bullet} = i/n$ in equation (23), we recover β_i from equation (10). Now, for the linear fitness functions, Price's terms have the following forms:

$$\mathrm{cov}_{f_i}(p_{i\bullet}, w_{i\bullet}) = (\Phi + \varphi)\mathrm{var}_{f_i}(p_{i\bullet}) \tag{24}$$

$$E_{f_i}[w_{i\bullet}\,\Delta p_{i\bullet}] = \varphi(\mathrm{var}_{(f_{i,j})}(p_{ij}) - \mathrm{var}_{f_i}(p_{i\bullet})) \tag{25}$$

Comparing equations (24) and (25) to equation (21) shows that both the group-level effects and individual-level effects are more extreme under contextual analysis than Price's equation for altruism with linear fitness functions. The main point is that contextual analysis and Price's equation give two different statistical perspectives on evolution.

We started this section with an example that was handled well by contextual analysis, but inappropriately by Price's equation. Another example illustrates the converse. Imagine

a case of soft selection in which the average fitness of individuals within any group was constant ($w_{i\bullet} = w$). However, assume that individual fitness does vary across types within groups ($\alpha_i \neq \beta_i$). In this case, intuition suggests that terms measuring "between-group" selection (or group-level effects) should be zero (as groups do not vary in output). While Price's first term is indeed zero, the first term from the contextual analysis equation (21) may not be zero (because it is possible that $\Phi \neq 0$). Thus, each statistical equation seems to handle certain cases better. As we discuss experiments in the following section, we will refer back to these statistical approaches.

EXPERIMENTAL RESEARCH

How do the above theoretical predictions fare for actual biological systems? There are several different empirical ways to explore the evolution of altruism (see also Huey and Rosenzweig this volume; Futuyma and Bennett this volume). One could approach the study of altruism from a phylogenetic perspective, looking whether certain traits predicted to favor altruism are likely to have been present before the radiation of a clade exhibiting altruism. For example, this approach pointed to the existence of high levels of inbreeding (promoting high coefficients of relatedness) *prior* to the origin of a soldier caste in eusocial gall thrip species (McLeish et al. 2006). Another approach to studying altruism is to check whether certain groups exhibiting altruism satisfy one of the theoretical criteria for the evolution of altruism. This has been an active approach within the kin selection literature, where much effort has been invested in computing coefficients of relatedness (all else being equal, higher relatedness between interactants is predicted to work for the evolution of altruism). As recent examples, high coefficients of relatedness have been found between wild turkey males that exhibit cooperative breeding (Krakauer 2005), within multiple-queen ant colonies (Bargum and Sundstrom 2007), and within groups of slime mold cells that form fruiting bodies with reproductive division of labor (Gilbert et al. 2007). A third approach involves experimental manipulation of factors predicted to affect the payoffs or evolutionary success of altruism (and the monitoring of different types). In some cases, this can be done over multiple generations, and the evolutionary loss or *de novo* gain of altruism can be monitored in real time. In this section, I focus on this third approach. A growing collection of artificial selection experiments, laboratory experiments, and field experiments have shed much light on the evolutionary circumstances favoring altruism and have provided food for thought within the levels of selection debate.

ARTIFICIAL SELECTION EXPERIMENTS

It is no accident that the first chapter of *The Origin of Species* focuses on variation in plants and animals cultivated by humans (Darwin 1859). Darwin discussed morphologically divergent breeds of domesticated organisms coming from common ancestral stocks in order to convince the reader that *selection* (in this case, artificial) was a powerful

agent of change. In the same vein, artificial selection for various properties of *groups* of organisms (rather than properties of individuals) has been experimentally explored as a way of illustrating the potential power of group-level selection.

WADE'S CLASSIC EXPERIMENT

The first group selection experiment was initiated by Wade (1977), using the red flour beetle, *Tribolium castaneum*. The group property selected in this experiment was population size. Obviously, in order to apply group selection, *multiple* groups (that vary in the selected property) must be maintained. For each of his experimental treatments, Wade propagated forty-eight populations of beetles over several population growth cycles. Every population was initiated with sixteen individual beetles. Different experimental treatments were defined by the nature of selection applied at the end of each population growth cycle (thirty-seven days). I will discuss three of his four treatments here.

In one treatment (Group Selection for Productivity), at the end of each cycle, the largest population was divided up into as many groups of sixteen individuals as possible, and these groups were used to initiate the next generation of populations. Thus, these "packets" of sixteen beetles served as "propagules" that seeded the next set of populations. If there were not enough propagules from the largest population to seed all forty-eight new populations (i.e., if $L/16 < 48$, where L is the size of the largest population), then the second-most productive population was also used to supply propagules; and if there were still not enough, the third-most productive population was used; and so on. This same selective scheme was applied over several population growth cycles. A simplified version of this treatment is illustrated in figure 19.2a.

In a second treatment (Group Selection against Productivity), at the end of each cycle, the *smallest* population was divided up into as many propagules as possible. Then the second smallest population was divided into propagules, followed by the third smallest, and so on, until forty-eight propagules were obtained (enough to seed the entire set of future populations). Note that because smaller populations were selected in this treatment, the number of "parent" populations producing propagules was greater in this treatment than in the first treatment.

Finally, in a control treatment (Individual Selection), Wade allowed each population at the end of the population growth cycle to produce a single propagule. That is, each population in the next generation received all of its initial beetles from only a single "parent" population. This is a soft selection scheme, where different groups have the same *realized* productivity (despite their actual productivity at the end of the cycle). The reason this control treatment is termed "individual selection" can be illustrated with the Price equation described earlier. Given that Wade experimentally equalized group output across all populations, Price's first term ($\mathrm{cov}_{f_i}(p_{i\bullet}, w_{i\bullet})$) is zero. If this first term is taken to represent group-level selection (notwithstanding the issues raised earlier), then any evolutionary change is attributed to Price's second term (representing selection within groups between individuals). A simplified version of this control treatment is shown in figure 19.2b.

Interestingly, over only a handful of population growth cycles (and thus selective episodes), Wade recorded large differences between his treatments (figure 19.2c). In the Group Selection for Productivity treatment, average population size (at the end of a growth cycle) remained high, while in the Group Selection against Productivity and Individual Selection treatments, average population size decreased. After nine rounds of selection, average population size differed by 158 beetles between the Group Selection for Productivity and Group Selection against Productivity treatments. Thus, selecting groups to provide propagules on the basis of their population size affects the population size in the future groups.

In the Individual Selection treatment, average population size did dramatically decrease (by over 150 beetles). One factor later determined to contribute to this decrease was an increase in the rate of cannibalism of pupae by adults in the Individual Selection treatment, whereas adult beetles in the Group Selection for Productivity treatment displayed a slightly lower rate of cannibalism (Wade 1979).[2] Thus, the direction of group
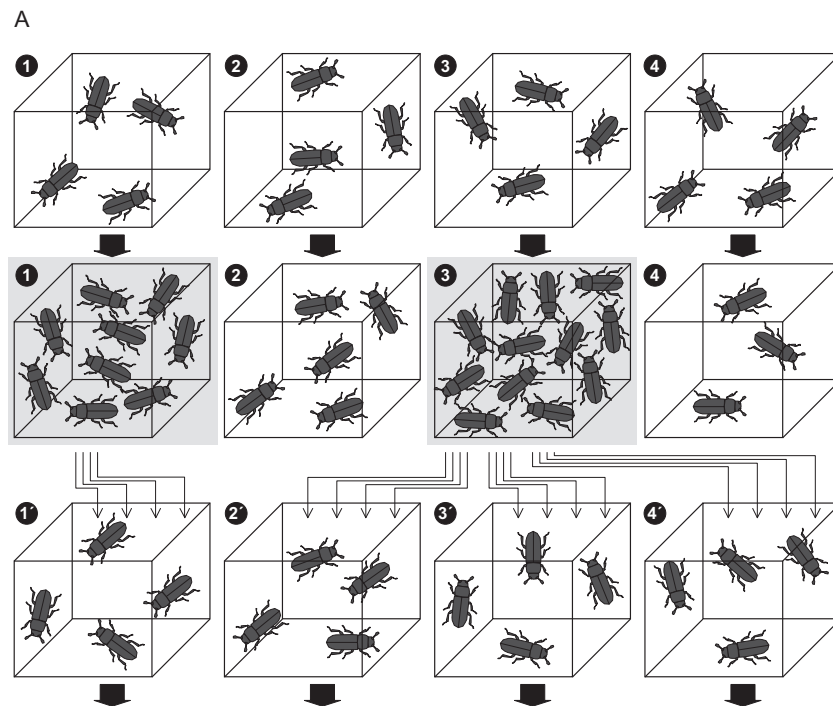
A

FIGURE 19.2

Wade's classic artificial group selection experiment. All subpopulations start with the same number of flour beetles and these subpopulations are incubated for a population growth cycle of thirty-seven days (the thick arrows in parts A and B). A, This schematic shows the Group Selection for Productivity treatment. The subpopulation producing the most individuals over a growth cycle is selected to "seed" the next set of subpopulations (in this case, it is subpopulation 3). Then the second-most productive subpopulation is used, and so on, until the next set of subpopulations have been filled and the next population growth cycle is initiated.
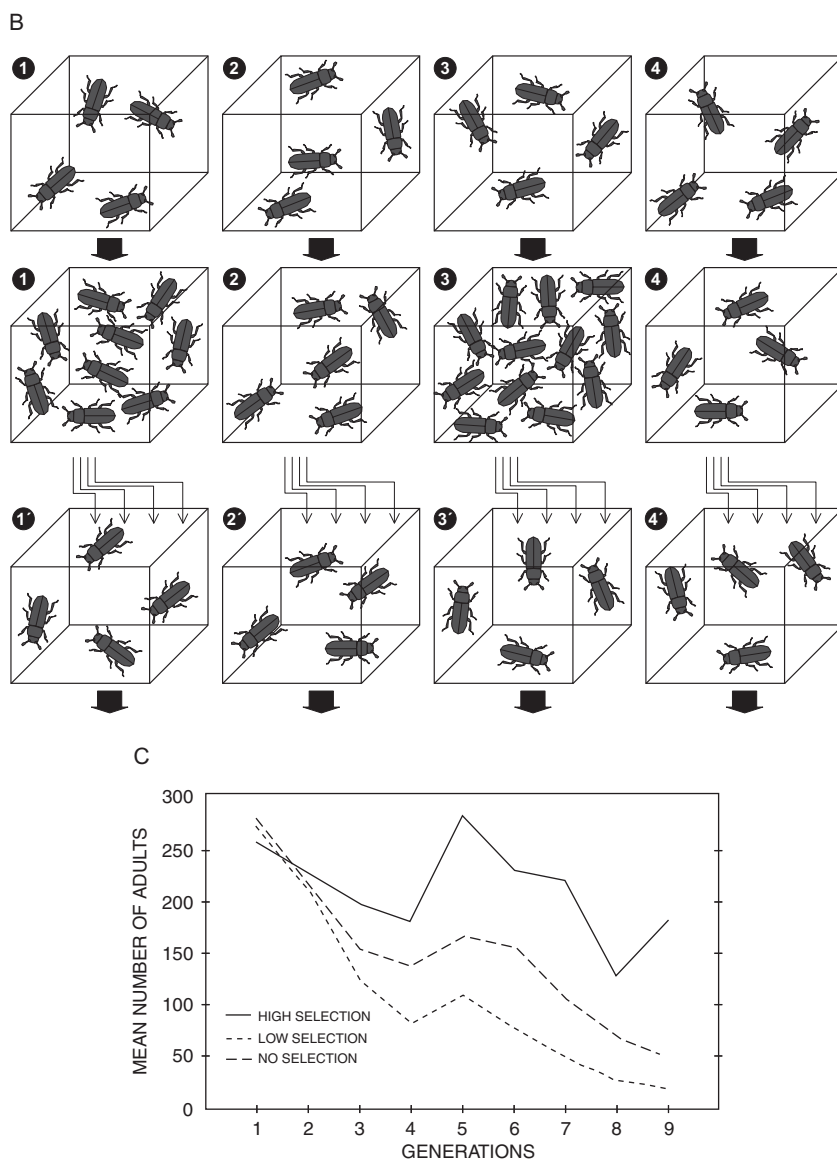
FIGURE 19.2 (*continued*)
B, This schematic shows the Individual Selection treatment. In this case, every group "seeds" one new group. Thus, there are no realized productivity differences between groups (despite differences that develop over the population growth cycle). C, The evolution of productivity in different selection treatments. High Selection is the Group Selection for Productivity treatment, Low Selection is the Group Selection against Productivity treatment; and No Selection is the Individual Selection treatment.

and individual selection in these cases appear to be in opposition. More specifically, when competition within groups occurs without realized productivity differences between groups (i.e., soft selection), some forms of cannibalism strengthen; however, when groups with higher actual productivity contribute more propagules to the next generation, some forms of cannibalism weaken.

An individual beetle that exercises restraint in its cannibalistic tendencies is an altruist (in this case, by the "multilevel" definition of conditions [4] and [5]). Here we analyze the evolution of this putative altruist under the Individual Selection and Group Selection for Productivity treatments using the Price equation. In each case, we measure evolutionary change starting from the set of populations before their growth cycle through their population growth, and ending with the next set of populations after a selective treatment is applied (the period illustrated from the top to the bottom of figures 19.2a and 19.2b). As mentioned, in the Individual Selection treatment, Price's first term, $\mathrm{cov}_{f_i}(p_{i\bullet}, w_{i\bullet})$, measuring *realized* productivity differences between groups, will be zero due to the nature of selection. This means that the only term contributing to evolutionary change is Price's second term, which will be negative for multilevel altruism. Thus, the frequency of restrained cannibals (the altruists) is expected to decrease, leaving populations in the Individual Selection treatment more cannibalistic.

In the Group Selection for Productivity treatment, populations that are productive over the population growth cycle disproportionately contribute to the next set of populations. If cannibalistic restraint is multilevel altruism, then actual group productivity will increase with the fraction of restrained beetles. In this case, the regression of average group output on frequency, $\kappa_{w_{i\bullet}p_{i\bullet}}$, is positive and Price's first term can now work for the altruists. In this case, the group selection (which is artificially applied) is expected to lead to a relative increase in the frequency of restrained cannibals, leaving the population less cannibalistic.

Assume that Wade had used a selection scheme more closely modeled after the trait group life cycle described earlier. That is, all beetles from all populations would be mixed into a "migrant pool," and the next set of populations would be initiated with subsets from this pool. This is a "hard selection" scheme in which Price's first term is nonzero. That is, there are realized differences in group productivity (because more productive groups have a higher chance of contributing individuals from the migrant pool to the next generation of groups). Wade did not use such a migrant pool (but see note 4); rather, a set of propagules from the most productive groups were used in his Group Selection for Productivity treatment. This choice had two important consequences: (1) the magnitude of the regression coefficient $\kappa_{w_{i\bullet}p_{i\bullet}}$ will be larger if *only* the members of large groups (which tend to house more altruists) realize nonzero fitness; and (2) if there is variance (e.g., in frequency of altruists) between groups that are selected from the end of growth cycle t, then Wade's propagules will tend to maintain a higher variance between groups at the beginning of cycle $t + 1$ (because the mixing that occurs in a migrant pool will tend to equalize between-group variance in the next generation).[3]

ALTRUISM AND THE LEVELS OF SELECTION · 601

Again, we see that Price's first term ($\mathrm{cov}_{f_i}(p_{i\bullet}, w_{i\bullet}) = \kappa_{w_{i\bullet}p_{i\bullet}}\mathrm{var}_{f_i}(p_{i\bullet})$) is larger under Wade's propagule selection scheme than under a migrant pool scheme. Thus, choosing circumstances in which groups send founder propagules to initiate new groups (over a trait group mixing/reformation phase) stacked the deck in favor of altruistic evolution in Wade's experiment. Indeed, in an experiment similar to Wade's, but with a livebearing fish, the failure of group selection to produce effects was attributed, in part, to the abandonment of propagule initiation of groups for a type of mixed migrant pool (Baer et al. 2000).[4]

OTHER GROUP SELECTION EXPERIMENTS

Several other artificial group selection experiments on flour beetles followed Wade's pioneering work (Craig 1982; Wade 1979, 1980a, 1980b, 1982; Wade and McCauley 1980, 1984). Emigration rate and cannibalism rate were added as group properties under artificial selection. These studies also explored differences in population structure, random extinction, and propagule size on the effects of selection. As was the case for Wade's original experiments, all of these studies detected significant effects of artificial group selection.

Artificial group selection experiments were applied to the cress, *Arabidopsis thaliana*, by Goodnight (1985). Leaf area was the object of selection in this study. Goodnight mixed artificial individual selection and artificial group selection in a fully factorial manner. He selected groups of plants with the highest (or lowest) *mean* leaf area to serve as "propagule generators" for the next generation of groups. However, he also selected the individual plants (within the selected groups) that had the highest (or lowest) personal leaf area to serve as actual parents of plants in the next generation. "No selection" controls (i.e., picking random groups for propagules or random individual parents to fill those propagules) were included in the factorial design. Interaction between selection for individual and group properties could thus be gauged. While Goodnight detected strong positive responses to group selection, responses to individual selection were weak and in some cases negative. In particular, individual selection for high leaf area produced plants with *lower* leaf area when compared to the "no individual selection" control. Interactions between individual and group selection were also detected. Group selection was most effective in the absence of any individual selection. Specifically, group selection for increased mean leaf area within groups was muted when paired with individual selection for increased personal leaf area!

One of the explanations Goodnight offers for these unexpected results is that individual selection might select for plants that are able to interfere with the growth of their fellow group-mates. Picking the individuals with the highest leaf area (if this is associated with the most "interfering" type) might lead to a propagule filled with interfering types. This means that the group in the next generation coming from this propagule is filled with plants that are interfering with each other's growth. As a consequence, leaf area (along with general plant health) could decline. This could explain the result that individual

selection for increased personal leaf area leads to decreased area when compared to the control of "no individual selection." The basic idea here is that no plant is an "island": by selecting specific plants, one is simultaneously selecting specific social environments for the next generation.

Selection for groups of plants with high mean area in Goodnight's experiment may tend to pick out plants that interfere less (or even facilitate) the growth of their neighbors (thus leading to green groups). If group selection for increased mean leaf area is not paired with individual selection for increased personal leaf area, then the most interfering members of the greenest groups are not preferentially selected. This could explain why the effect of group selection was stronger without the presence of individual selection. In general, competitive restraint in plants (like cannibalistic restraint in beetles) could be viewed as a form of altruism; if so, within and between group selective effects would be in opposition. How much competition, cannibalism, or interference evolves would then depend on the relative strengths of these effects and, thus, on the nature of selective episodes.

A result similar to Goodnight's was found in an interesting practical application of artificial group selection within the poultry industry (Craig and Muir 1996; Muir 1996). In order to improve harvesting efficiency, many egg-laying chickens are currently kept in cages with multiple hens. Aggression between chickens can be substantial and beaks are trimmed to prevent injuries inflicted by cage mates. If one always selects the most productive chickens from cages, then average egg productivity within cages can actually *decrease* (Craig et al. 1975). Like the aforementioned plants, chickens in these cages are not islands: by selecting the most productive hen, one may be favoring a more aggressive social environment in which future egg production suffers.[5] Muir (1996) instead selected for egg productivity at the level of the multihen cage (hens from the most productive cages were used as parents to "seed" the next set of cages). This artificial group selection scheme produced dramatic results after a small number of generations, with annual egg production increasing 160 percent. Part of this improvement is due to lower aggression and improved stress coping abilities within cages (Dennis et al. 2006). Indeed, aggression was low enough in these group-selected lines that beak trimming was no longer necessary. Thus, this group-selected line of chickens has led to the potential for substantial cost savings through decreased hen mortality, abandonment of beak trimming, and increased egg production. As Sober and Wilson (1998) put it: "If this strain becomes widely used in the poultry industry, the projected annual savings will far exceed the money spent by the U.S. government for basic research in evolutionary biology."

ECOSYSTEM SELECTION EXPERIMENTS

The experiments described so far impose artificial selection on some property of groups of conspecific organisms. However, the same selection protocols can be applied to multispecies communities or ecosystems; and researchers have recently executed artificial

selection experiments at these higher levels (Swenson et al. 2000a, 2000b). These authors placed soil and aquatic microbial communities into a set of microcosms and then selected on the basis of various ecosystem properties. In the case of the soil communities, the authors selected microbial communities on the basis of aboveground biomass of *Arabidopsis thaliana* growing in the soil (Swenson et al. 2000b). In the case of aquatic communities, the authors selected microbial communities on the basis of microcosm pH (Swenson et al. 2000b) or the degree to which 3-chloroaniline (an industrial waste product) was degraded (Swenson et al. 2000a). Note, that in these cases, the criterion for selection was *not* a direct property of the organisms being selected; rather, these microbes were being selected on the basis of their effects on other organisms (*A. thaliana*) or their effects on the physical environment (pH or levels of 3-chloroaniline).

The actual selective protocols were similar to Wade's classic experiment. For the pH selection experiment, from the twenty-four aquatic microcosms, the six with the highest (or lowest) pH were used to send propagules to four microcosms each in the next generation. In the other selective experiments (plant biomass and 3-chloroaniline breakdown), the authors mixed and redistributed the selected microcosms across generations (instead of transferring propagules). For each of these experiments, responses to the selection schemes were recorded (either between high and low selected lines or between selected and control lines). The authors also found that the chemical and organismal composition of their microcosms came to differ under different selective regimes. Putting aside whether and how such experiments inform us about evolution in natural systems, there are obvious practical implications of such research (e.g., in bioremediation and agriculture).

The central idea being exploited in these experiments is the premise behind the theories of niche construction and ecosystem engineering: organisms affect each other and their environments (Jones et al. 1994; Odling-Smee et al. 2003). Organisms are not islands with respect to social interactions with conspecifics, but they are also not islands with respect to interactions with heterospecifics or shared environments. In some cases, selecting community players from a series of multispecies configurations on the basis of some ecosystem-level property may produce a response that does not occur when an organism selected for the same property (e.g., in isolation) is returned to a multispecies community.

QUASI-NATURAL SELECTION EXPERIMENTS

There have been several criticisms of the artificial selection experiments described in the previous section. First, some critics have argued that the selected properties seem arbitrary (e.g., community pH) and it is unclear how evolution would proceed with respect to such properties in natural systems. Second, critics have noted that a mechanistic understanding of responses to artificial selection is lacking in some cases. (Note, however, that this is a criticism that can be extended to many "individual-level" artificial selection

experiments, as well; see, e.g., Swenson et al. 2000b.) Third, critics have maintained that the experimental conditions used are extremely unlikely to apply outside the lab (Harrison and Hastings 1996). For instance, if migration is simultaneously the way that new groups form and a force that homogenizes intergroup variance, then the very means of producing offspring groups leads to a collapse in their variation (and thus the strength of group selection—e.g., as represented by Price's first term). Thus, even if an experimentalist can decouple these effects, through a combination of group extinction and propagule transfer, the operation of this process in nature is proposed to be limited. Fourth, some critics maintain that none of this should be called "group" or "ecosystem" selection. After all, isn't it the individuals within the groups that are changing? Can't we simply talk about this as a more complicated form of individual-level selection? I will return to this last criticism in the discussion.

The process of selection is actually quite clear in the experiments described in the previous section—indeed, it is the hand of the experimenter that picks *groups* based on group properties. An alternative approach is to create structured worlds in which groups thrive, split, mix and die "by their own devices." Specifically, the experiment is set up so that there is not any one *single* group property (e.g., productivity) that is being used to sort groups. Rather, groups and the individuals they contain prosper or fail based on the eco-evolutionary conditions that apply to the created world. This second approach is what Scheiner (2002) calls "quasi-natural selection" experiments. These experiments often take place in the laboratory, so it is not the environmental conditions that are "natural." What is taken to be more natural is the idea that selection could be simultaneously "operating" on a *variety* of properties within the system. Some of the criticisms of the artificial selection scheme play out differently for quasi-natural selection experiments. For instance, ostensibly arbitrary properties are generally not the basis of selection.

THE EVOLUTION OF COMPETITIVE RESTRAINT

A recent experiment illustrates the basic features of quasi-natural selection experiments. Kerr et al. (2006) explored evolution within a host-pathogen system consisting of the bacterium *Escherichia coli* (the host) and T4 bacteriophage (the viral pathogen). A meta-community was created by distributing the bacteria and phage into a large number of subpopulations (wells in multiwell plates). Using a high throughput liquid-handling robot, the entire metacommunity was serially propagated (transferring a small fraction of each well at the end of an incubation period to a corresponding well with fresh medium in a new plate to start the next incubation period). In addition, migration between wells occurred during these serial transfers. Within any subpopulation, phage T4 drives its bacterial host extinct over an incubation period; thus, the host and pathogen cannot coexist at a local scale. However, under moderate migration levels, the host and pathogen can coexist at the metapopulation scale due to dynamic asynchrony between wells. That is, at any point in time, some bacteria-filled wells are undergoing phage-driven

extinction, while some phage-filled wells are becoming empty due to dilution in the absence of hosts, while some empty wells are being restocked with bacteria through recolonization.

Within these microbial metacommunities, Kerr et al. manipulated the pattern of migration between subpopulations. Thus, altered migration schemes distinguished the different created worlds in this quasi-natural selection experiment. One treatment constrained migration to take place between neighboring wells within the multiwell plates, termed the Restricted Migration treatment (figure 19.3a). A second treatment allowed migration to take place between any two wells within the metacommunity, termed the Unrestricted Migration treatment (figure 19.3b). The rate of migration was held constant between treatments (only the topology of migration differed). After several transfers, the authors measured phage productivity (number of progeny produced per parent phage particle in a well with bacteria over an incubation period) and phage competitive ability (relative fitness of the evolved T4 strain when in competition for shared host cells with a marked T4 mutant). They found that phage from the Restricted Migration treatment evolved high productivity, but low competitive ability. Meanwhile, phage from the Unrestricted Migration treatment evolved high competitive ability, but low productivity (figures 19.3c and 19.3d). Since the same phage strain was used to inoculate all treatments, these results suggested the selection of *de novo* phage mutations.
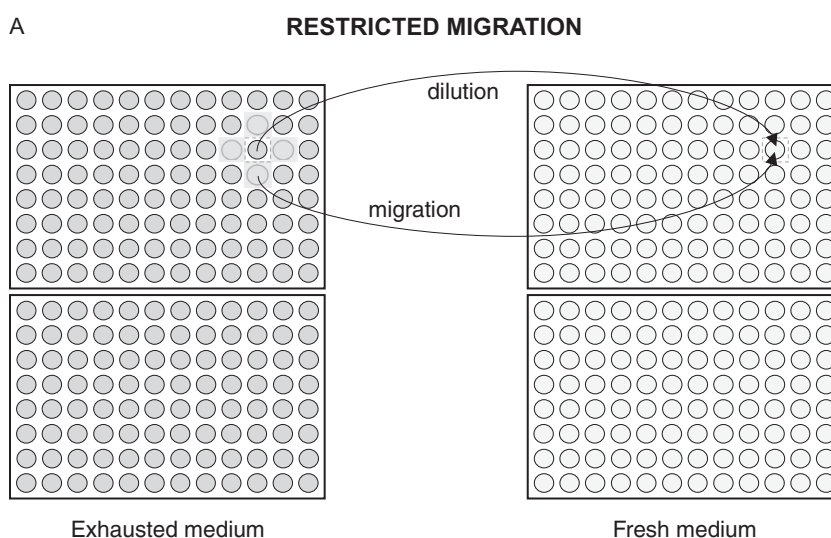

FIGURE 19.3

The metapopulation experiment of Kerr et al. (2006). Within two microtiter plates, 192 wells contain subpopulations of bacteria and phage. Every subpopulation is diluted into a well with fresh medium every twelve hours (a standard serial propagation at the metapopulation level). A, In the Restricted Migration treatment, immigration can occur into a focal well (boxed in a dashed line) from one of its nearest well neighbors (the wells that are highlighted to the north, south, east or west of the focal well).
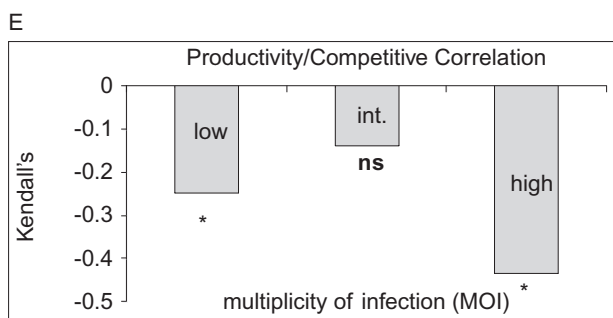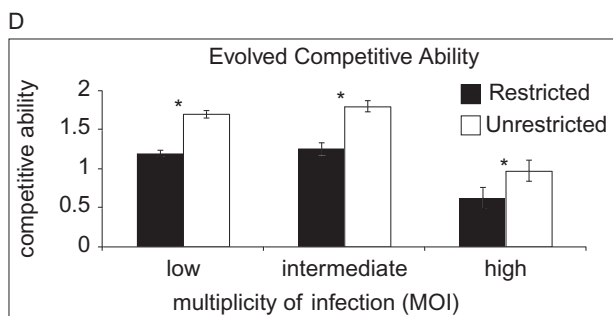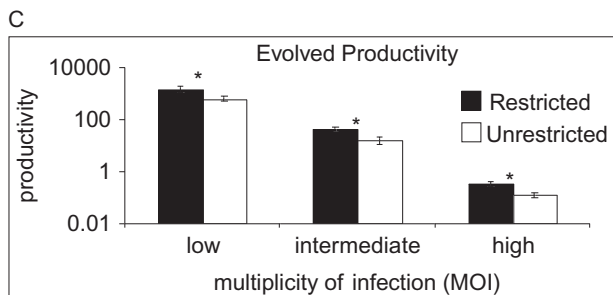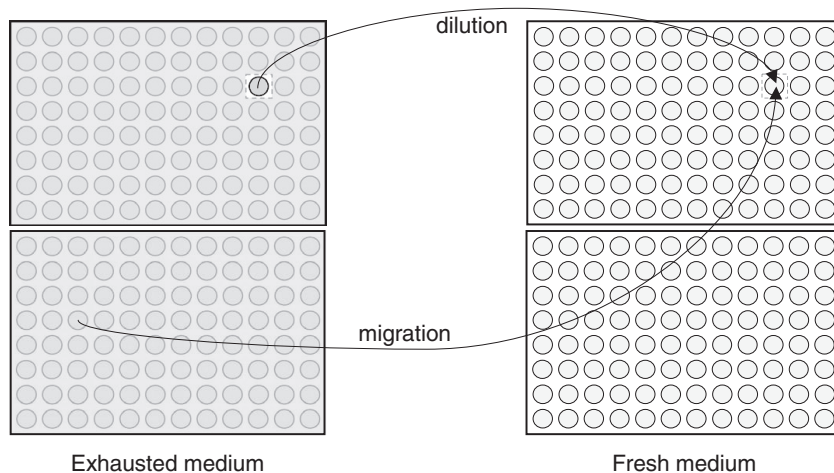
B **UNRESTRICTED MIGRATION**

dilution

migration

Exhausted medium   Fresh medium

C

Evolved Productivity

■ Restricted
☐ Unrestricted

productivity

multiplicity of infection (MOI)

D

Evolved Competitive Ability

■ Restricted
☐ Unrestricted

competitive ability

multiplicity of infection (MOI)

E

Productivity/Competitive Correlation

Kendall's

low   int.   high

ns

multiplicity of infection (MOI)

FIGURE 19.3 (*continued*)
B, In the Unrestricted Migration
treatment, immigration can
occur into a focal well from any
well in the entire metapopula-
tion. C, Productivity of evolved
phage isolates was measured at
three different ratios of phage
to bacteria (the "multiplicity
of infection"). In all three cases,
phage from the Restricted
Migration treatment were
significantly more productive. D,
Phage from the Unrestricted
Migration treatment competed
significantly better for common
host resources when paired with
a marked phage strain. E, After
pooling the data, negative corre-
lations between productivity and
competitive ability were discov-
ered (significant in two out of
three cases).

In this experiment, bacteria are resources for the phage. Phage exhibiting unrestrained use of this resource (e.g., entering and killing their host quickly) may outcompete restrained phage for common host resources (Abedon et al. 2003). However, phage exercising restraint (e.g., spending a longer period inside their host) may extract more from their resources and consequently may be more productive when competitors are absent. In this microbial community, there appears to be a trade-off between productivity and competitive ability (overall negative correlations were found between these properties—figure 19.3e). Thus, this system has the ingredients for a "tragedy of the commons" (Hardin 1968), which occurs when unrestrained individuals displace restrained ones, leading to overexploitation of shared resources and a lower group productivity. A microbial version of this tragedy occurs when rapacious phage outcompete their prudent cousins and end up lowering group productivity in the process.

Why would the pattern of migration within a metapopulation influence the resolution of this tragedy of the commons? As rapacious phage mutants are generated within a metacommunity, they will outcompete their prudent ancestors locally. However, these rapacious phage are less productive; therefore, given periodic dilution in the experiment, the rapacious phage are more extinction-prone. Prolonged survival of the rapacious phage depends on the virus continually finding fresh host bacteria to exploit. Such encounters are more likely in the Unrestricted Migration scheme. Furthermore, mixing of different phage types into the same subpopulation is more likely within the Unrestricted Migration treatment. Both of these factors contribute to the success of the unrestrained strategy given unrestricted migration. In the Restricted Migration treatment, limitation to host access leaves the rapacious phage vulnerable to extinction and the phage population remains prudent by default (Kerr et al. 2006; Prado and Kerr 2008). In essence, the form of migration can determine whether local tragedies of the commons can become global tragedies of the commons.

Some of the first models exploring the evolution of cooperation (e.g., Maynard Smith 1964) assumed the potential for tragedies of the commons at the subpopulation scale. The altruist is the type displaying prudent use of resources. Indeed, relative to rapacious types, prudent phage from the experiment just described are "multilevel altruists" (i.e., they satisfy conditions [4] and [5]). Looking back to the Price equation, the first term, $\text{cov}_{f_i}(p_{i\bullet}, w_{i\bullet})$, will be positive. That is, as the frequency of prudent phage in a well with hosts increases, the absolute productivity of the well increases. However, Price's second term, $E_{f_i}[w_{i\bullet}\,\Delta p_{i\bullet}]$, will be negative. That is, the prudent phage always loses in competition to the rapacious type in a mixed subpopulation. If these statistical terms are interpreted as group-level and individual-level "selective forces," then the outcome of evolution depends on the strength of group selection for prudence versus individual selection for rapacity (see the discussion for an exploration of the causal adequacy of this account). What will influence the magnitude of these terms? One factor is migration. Specifically, restrictions to migration will tend to produce a clumped group frequency distribution (with rapacious phage isolated in some subpopulations and prudent phage isolated in

others). This tends to tip the balance in favor of altruism (prudence). This is easiest to see when the fitnesses are linear (equations [9] and [10]), because the larger the "clumping index" (7), the more likely our generalized version of Hamilton's rule (11) is to hold. Thus, there are both theoretical and experimental reasons to think that the pattern of migration will influence the form of social evolution in metapopulations.

THE EVOLUTION OF VIRULENCE VERSUS BENEVOLENCE

At a fundamental level, disease-causing organisms occupy metapopulations since their hosts constitute discrete sets of resources (i.e., potential subpopulations). The manner in which the pathogen uses the resources of its host will influence the progression and nature of the disease. One property influenced by pathogen traits is disease virulence. Here, I will define *virulence* to be the increase in the rate of death of a host when infected by its pathogen. The conventional wisdom is that pathogens should always evolve *reduced* virulence and that highly virulent pathogens result from evolutionarily *recent* introduction into a new host (see Bull 1994 for a full discussion). All else being equal, a pathogen that more rapidly kills its host is at a selective disadvantage relative to one that preserves its host (and future opportunities for transmission). A complication occurs if all else is *not* equal. Specifically, there may be functional relationships between virulence and other properties that will affect the success of a pathogen. For instance, when more virulent pathogens are transmitted at higher rates, selection may not favor the lowest level of virulence (Anderson and May 1982).

In some cases, virulence is a product of the schedule of pathogen reproduction within the host, where rapid reproduction yields higher virulence. In this context, avirulent pathogens reproduce slower. In pair-wise competition for common host resources, the virulent strain of the pathogen is expected to displace the avirulent strain. However, hosts with more virulent strains die sooner (by definition). This description of virulence and avirulence conforms nicely to multilevel altruism (equations [4] and [5]). The avirulent pathogen exercises relative restraint in the use of host resources, a strategy that is benevolent (toward both the host and co-occuring pathogens).

What are the conditions favoring virulence versus benevolence? There have now been several quasi-natural selection experiments on this subject. It turns out that the nature of transmission (its mode, spatial scale, and timing) can strongly influence evolutionary outcomes. Bull et al. (1991) and Messenger et al. (1999) used the filamentous phage f1 that infects *E. coli* to explore the role of the mode of transmission on the evolution of benevolence. This bacteriophage is unusual in that it establishes a permanent infection in which progeny phage are secreted through the bacterial envelope without killing the host. Infection is detrimental to the growth of the bacterial host however. Thus, it is not virulence by the above definition (effect on death rate), but rather the "degree of harm" to host growth rate that was monitored in these experiments. As the rate of production of secreted phage progeny increases, so, too, does the degree of harm to its host (Messenger et al. 1999).

These authors manipulated the degree of vertical versus horizontal transmission in this system. (Vertical transmission occurs when the viral genome is inherited from a parent host cell to its offspring, whereas horizontal transmission occurs when viral progeny from one cell infect a previously uninfected, unrelated cell.) They found that vertical transmission promotes the evolution of benevolence, whereas the phage was more harmful to its host (but more productive in terms of secreted progeny) under a selection scheme involving higher levels of horizontal transmission.

The standard explanation for these findings is that vertical transmission promotes partner fidelity, which tends to work for cooperation in the host-pathogen interaction. A complementary way to explain these results is to realize that there are two ways a phage strain can increase its numbers, through (1) passing its genome copies to the offspring of its current host (offspring cells of an infected parent cell are also infected) and (2) secreting packaged phage progeny that successfully infect formerly uninfected hosts (note there is no superinfection in this phage). The trade-off between production of secreted progeny and reproduction of the infected host translates to a trade-off between these two components of fitness. As vertical transmission becomes common, there may be a premium placed on promoting the welfare of the host. As horizontal transmission opportunities become abundant, selection may favor strains that invest in producing many packaged infective particles. These findings are not restricted to microbial host-pathogen systems. For instance, vertical transmission of algal symbionts in jellyfish promoted benevolence, whereas horizontal transmission favored the evolution of *parasitic* algal symbionts with negative effects on jellyfish reproduction and growth (Sachs and Wilcox 2006). As with the phage-bacterium system, a positive relationship between degree of harm to the host and the rate of expulsion of the symbiont seems to underlie these results.

One factor that is predicted to influence the mode of transmission is the migration rates or movement patterns of hosts and their pathogens. Generally, as migration or movement becomes less restricted the opportunities for horizontal transmission improve. This will tend to select for more virulent pathogens if virulence trade-offs exist (e.g., Kerr et al. 2006). A recent experiment probed the effects of host movement on pathogen evolution. Boots and Mealor (2007) altered the viscosity of microcosm environments containing larvae of a phycitiid moth and a species-specific granulosis virus. They found that the virus evolved higher infectivity when hosts could move easily within the microcosm (less viscous food medium) as compared to microcosms in which the host was more restricted in its movement (more viscous food medium). Infectivity is the proportion of hosts infected after exposure to the virus. While infectivity is a different property than virulence, there is an obvious connection between the two. In the same way that restricted migration between hosts can favor pathogens that are more prudent *within* their hosts (i.e., less virulent), restricted movement of hosts can favor pathogens that are more prudent with the *set* of hosts to which they are exposed (i.e., less infectious).

Another quasi-natural selection experiment explored the role of the timing of transmission on the evolution of virulence (Cooper et al. 2002). These authors used a viral

pathogen of the gypsy moth. The experiment consisted of several cycles of transmission from infected larvae to uninfected larvae. Virus was obtained from infected larvae by homogenizing them at a specific time in their development (after which the virus was introduced to the next set of larvae to complete the transmission). In their Early treatment, virus was obtained from live infected larvae five days after infection. In their Late treatment, virus was obtained from live infected larvae nine days after infection. Virus evolved in the Early treatment was significantly more virulent than virus evolved in the Late treatment. An important component of their experimental protocol was that only *living* larvae could serve as sources of transmission at the time of transmission. While highly virulent viral strains outreproduce less virulent strains over short time periods (e.g., five days), these same virulent viral strains destroy their host over longer time periods (e.g., nine days). In the Late treatment, only those larvae containing less virulent viral pathogens would survive to the time of transmission.

The above study illustrates why virulence has popped up regularly in the discussions about levels of selection (Lewontin 1970; Sober and Wilson 1998; Wilson 2004). Of course, hosts often contain discrete *groups* of pathogens. When a pathogen's competitive ability within its host trades off with the longevity of the infected host, a multilevel perspective posits conflicting selective forces between and within groups. Specifically, the virulence of the pathogen determines the rate of group extinction (i.e., host death) such that groups of virulent pathogens are more extinction-prone. This is interpreted as a "between-group force" *against* virulence. On the flip side, when virulence is positively related to within-host competitive ability, more virulent pathogens should increase in frequency within hosts infected with different strains. This is interpreted as a "within-group force" *for* virulence. As before, the two terms from the Price equation nicely partition these components. Whether the pathogen becomes more virulent or more benevolent depends critically on how groups of pathogens form (e.g., do infections start with small numbers of pathogens that resist superinfection or does infection proceed with continual introduction of superinfecting pathogens?) and the exact nature of the relationships between virulence and other pathogen traits (i.e., trade-offs determining the fitness structure). As before, statistical association within hosts of benevolent pathogens will tend to work for lower virulence. It is interesting to note that Cooper et al. (2002) initialized each larval infection with a small number of viruses to "maintain within-host homogeneity." From a multilevel perspective, this has the effect of shrinking the within-group force, tipping the balance toward avirulence under conditions where there is a premium placed on group survival (in their Late treatment).

THE EVOLUTION OF CONFLICT RESOLUTION

In 1966, microbiologist K. Jeon noticed that his cultures of amoeba were infected with an intracellular bacterial parasite (which he termed X-bacteria). The X-bacteria harmed its host (in terms of compromised growth, reduced reproduction, and increased mortality).

However, Jeon continued to propagate some of the infected amoeba. After five years, the amoeba populations fully rebounded. However, this recovery was not due to the exclusion of X-bacteria. Indeed, the bacterial symbiont was still very much present. Through surgical and chemical manipulations (Jeon 1972; Jeon and Hah 1977), Jeon and colleagues found that, in some cases, if X-bacteria and its amoeba host were separated, the *amoeba* would not be viable! A parasitic relationship had evolved into a mutualism.

In order for the change that Jeon witnessed to occur, conflict between interacting parties must be resolved. (See Sachs and Bull 2005 for another fascinating example involving viruses.) The issue of conflict resolution arises both in social interactions between members of different species (as in Jeon's system) and between individuals of the same species. Effective conflict resolution carries consequences for the origin and maintenance of altruism. Specifically, the fundamental barrier to the emergence and persistence of altruists is the intrinsic conflict with *cheaters*. Cheaters are taken to be individuals that derive benefits from interactions with altruists but do not contribute socially. In this section, I will focus on experiments with microbes that highlight the ways social conflict can be resolved (see Travisano and Velicer 2004 for a full discussion).

In microbial systems, the evidence for the existence of cheating and conflict is widespread, with examples in viruses, prokaryotes, and eukaryotes. For instance, in phage Φ6, Turner and Chao (1999) evolved a "cheating" strain that was able to outcompete its ancestor within a host cell *(Pseudomonas phaseolicola)* infected by both strains. However, this cheat had lower reproductive output in pure infections, in comparison to the output of its ancestor in pure infections. Another "tragedy of the commons" scenario was discovered in strains of *Saccharomyces cerevisiae* that excrete an enzyme that breaks down extracellular sucrose facilitating subsequent sugar uptake. Yeast strains defective in exoenzyme production were functionally "cheats," able to free-ride on the public goods provided by producers without incurring the cost of production (Greig and Travisano 2004). Similarly, cheats were detected in biofilms of *Pseudomonas flourescens* (Rainey and Rainey 2003). In this system, "cooperating" bacteria excrete a sticky polymer that allows a mat to form at the interface of liquid medium and the air. Cheating strains do not contribute to the public good (the sticky matrix). As a consequence, the cheating strains have a growth advantage within the mat but lead to premature collapse of the mat due to its weaker integrity. Perhaps the most famous microbial examples of cheating come from social microbes, such as the bacterium *Myxococcus xanthus* and the slime mold *Dictyostelium discoideum*. For both organisms, starvation causes aggregation of single cells into multicellular fruiting bodies. In these bodies, some cells sacrifice themselves as supportive "somatic tissue," while other cells form reproductive spores (the "germ line"). Cheating occurs when a strain is able to achieve disproportionate representation in the spore pool when mixed with other strains, but this strain is compromised in its ability to form functional fruiting bodies when it occurs alone. Such cheaters have been described in both species (Velicer et al. 1998; Dao et al. 2000; Strassmann et al. 2000).

Given the ubiquity of cheating types in these and other systems, how is it that cooperation does not succumb to these antisocial influences? Quasi-natural selection experiments have provided some answers. I will present two classes of explanations here: (1) passive cheater control due to favorable social context formation and (2) active cheater control due to exclusion. For the first class, the experimenter manipulates factors that influence how social context forms. For instance, Greig and Travisano (2004) found that cooperative yeast strains, which produce an exoenzyme that degrades sucrose, had a competitive edge over defectors, which don't produce the exoenzyme, when the mixed population was grown at low density in a spatially structured habitat. In this case, the benefits of cooperators are disproportionately experienced by cooperators, because clumped growth occurs in the structured habitat. On the other hand, defectors receive fewer such benefits due to the limited diffusion of the exoenzyme. Similarly, in an experiment on metapopulations of *Pseudomonas aeruginosa*, Griffin et al. (2004) reported that cooperating strains (producers of siderophores, extracellular iron-gathering compounds) were able to displace cheats (nonproducers) when each new subpopulation was established from a single clone under a hard selection scheme. Barring mutation, cooperating bacteria were guaranteed to be interacting with other cooperators (namely, relatives), whereas defectors were isolated in their own subpopulations. This form of social context formation is optimal for cooperation by maximizing Price's first term, while minimizing the second. Finally, in the study on phage T4 described earlier, Kerr et al. (2006) found that prudent phage evolved when migration was restricted. Under such restricted movement, cooperators and defectors are less likely to compete for the same resources. Again, this type of social context formation works for cooperation.

In these experiments, the experimenter controls social context formation, by controlling density, founder size, and migration. The second class of cheater control mechanisms involves active exclusion of would-be defectors by the cooperators themselves. In an experiment with marked strains of *Dictyostelium purpureum*, Mehdiabadi et al. (2006) monitored aggregations of different strains. They found that fruiting bodies consisted primarily of *one* strain, suggesting that a form of kin discrimination occurs during fruiting body formation. Working with *D. discoideum*, Queller et al. (2003) showed that the product of gene *csA*, a protein embedded in the membrane that is involved in cell-cell adhesion, enabled slime mold cells to exclude *csA*-knockout cells from fruiting bodies under natural conditions. *csA* thus qualifies as a "greenbeard" allele: (1) it displays a phenotypic trait (a membrane protein), (2) it recognizes the same phenotype in other individuals (through homophilic binding), and (3) it provides a disproportionate benefit to like types (under natural conditions, the adhesion engenders overrepresentation in aggregation streams). Interestingly, under laboratory conditions (e.g., on agar plates) where both *csA* and *csA*-knockouts can coaggregate, these authors showed that *csA*-knockouts are overrepresented in the spores! Thus, the *csA*-knockout can be thought of as a cheater. However, under natural conditions, this would-be cheater is kept in check by active exclusion.

Finally, in a tour de force experiment with *Myxococcus xanthus,* Fiegna et al. (2006) discovered cheater control originating from a most unexpected place. These authors followed the evolution of an obligate cheater—a strain that could not make functional fruiting bodies in isolation but preferentially moved into the spore pool when mixed with other strains. Starting with a mixture of a cooperating strain and this obligate cheater, the cheat increased in frequency; however, as it rose to prominence, its inability to form functional fruiting bodies caught up with it, because the authors required the population to go through a spore stage every propagation cycle. Once in the majority, the cheat's density dropped precipitously, a perfect example of a tragedy of the commons. However, a mutant then arose that the authors dubbed "Phoenix," in reference to the mythical bird. Phoenix is not an obligate cheater; in fact, it is a superior cooperator, in that it makes functional fruiting bodies and can actively exclude its cheating progenitor in mixed populations. Amazingly, full genome sequencing revealed that this social back-flip was mediated by a single mutation. Thus, in a somewhat ironic twist of events, cooperation and cheater control evolved directly from the cheat itself.

NATURAL SELECTION EXPERIMENTS

The selection experiments of the previous section earned the adjective *quasi-natural,* because the arena of selection was often different from the natural conditions of the study organism. An argument could be made that, with respect to real-time evolution in the laboratory, the "artificial" conditions *were* the "natural" conditions. However, in understanding the traits of organisms in "the wild," it is sensible to study selection in the context under which those traits evolved. This means gathering data and performing experiments in the field.

In this section, I focus on a set of field experiments using plants. All of these field studies employed contextual analysis (discussed earlier; see also Heisler and Damith 1987; Goodnight et al. 1992; Goodnight and Stevens 1997). In contextual analysis, a relationship is sought between a proxy for the fitness of an individual (e.g., seed number or survival in a focal plant) and both "individual" and "contextual" traits. Individual traits belong to the focal individual (e.g., its size or height), whereas contextual traits belong to the group or neighborhood in which the focal individual is embedded (e.g., mean size or population density). A multiple-regression analysis teases apart the effects of individual properties and group properties on individual fitness. For instance, a statistically signif-icant partial regression coefficient of individual fitness on a group trait is taken to indi-cate the operation of group selection.

In an observational study of the jewelweed, *Impatiens capensis,* Stevens et al. (1995) found that larger individuals had higher survival, more chasmogamous (open-polli-nated) flowers, and more cleistogamous (self-pollinated) flowers. However, they also found that having smaller-sized neighbors led to higher survival and more cleistoga-mous flowers. Thus, according to the contextual analysis, there was individual selection

for larger size working against group selection for smaller size. The authors discuss small plant size as a potential type of altruism: "The altruistic plant may forgo the individual advantage of large size 'for the sake of' the increased survival rate and reproduction of the group." Interestingly, in stands of *Impatiens*, relatedness is likely to be high due to high levels of selfing and low levels of seed dispersal. Thus, the type of statistical clumping (described earlier) needed to favor altruistic phenotypes may be present (see Stevens et al. 1995).

In the Stevens et al. study, the partial regression coefficients of cleistogamous flower production on individual plant size and the mean of plant size in the group are opposite in sign but equal in magnitude. This is consistent with soft selection on plant size in their populations (i.e., constant yield from groups despite composition or density). When group output does not vary with group composition, then Price's first term is zero (in the case of linear fitnesses, the relationship between equally opposing regression coefficients and Price's first term can be seen clearly in equation [24]). Thus, Price's analysis suggests that there is no group selection in the *Impatiens* system, whereas contextual analysis detects group selection. I claimed earlier that the presence of a group selection effect under soft selection schemes was a weakness of contextual analysis. However, this "weakness" is completely dependent on the way that "group selection" is defined. For instance, Stevens et al. (1995) define group selection as "variation in the fitness of an *individual* due to properties of the group or groups of which it is a member" (my emphasis). Given that definition, contextual analysis is the perfect tool for detecting group selection, because it isolates the effects of group context on individual fitness while controlling for individual effects. However, others have defined group selection as differential fitness of *groups* (e.g., Sober and Wilson 1998). That is, groups, rather than individuals, are the bearers of fitness. Given this second definition of group selection, contextual analysis may detect a group-level effect when there are no differences between groups in their outputs (i.e., there is a classic frequency-dependent soft selection scenario). Given that different statistical approaches package selection differently, claims about the presence of group selection may be wholly dependent on the definition one is employing.

Using contextual definitions of group selection, other researchers have experimentally explored multilevel selection in other plant systems. For instance, Donohue manipulated relatedness (Donohue 2003) and density (Donohue 2004) in experimental stands of the Great Lakes sea rocket, *Cakile edentula*. She found that both of these factors influenced the strength of group selection, where the strongest group-level effects occurred with high relatedness between a focal plant and its neighbors (Donohue 2003) and when density was at an intermediate level (Donohue 2004). Group selection on plant size was found to operate in concert with individual selection in sibling groups (Donohue 2003); however, group and individual selection were opposed at intermediate densities (Donohue 2004): shorter, heavier plants growing with taller, lighter neighbors had higher fitness. Weinig et al. (2007) also found an effect of stand density on group selection in *Arabidopsis thaliana* with a contextual analysis approach. These authors found that, at

higher densities, the strength of group selection for size and branching patterns increased. In this study, group and individual selection effects were opposed, where individual selection favored an increase in size and elongation, while group selection favored a decrease. The idea of slower plant development as a form of altruistic restraint is particularly intriguing given the high levels of selfing and low levels of seed dispersal in natural populations of *A. thaliana*.

DISCUSSION

Even though the simplest version of natural selection treats organisms as islands, entire of themselves, it is clear that virtually all organisms experience meaningful interactions with other organisms. In this context, altruism is of particular interest, as it most clearly highlights the shortcomings of the "organism-as-island" model. Common explanations for the evolution of altruism are based on the idea that an altruist has a different social experience than a nonaltruist. Specifically, the altruist somehow manages to enjoy a more altruistic social circle. This can occur if the altruist interacts with kin (Hamilton 1964), conditions its behavior on the previous behavior of its partner (Trivers 1971; Axelrod and Hamilton 1981), or punishes/excludes noncooperators in its social sphere (Frank 1995). Using these theoretical predictions as inspiration, a collection of elegant experiments in the laboratory and the field have provided a fuller picture of the evolution of altruism in biological systems.

The first empirical finding is that altruism exists. Of course, this statement needs to be qualified by the operational definition of altruism. However, by any one of a few definitions, altruism (or the potential for altruism) has been uncovered in many systems, including viruses (Turner and Chao 1999; Kerr et al. 2006), bacteria (Velicer et al. 1998; Rainey and Rainey 2003), protists (Dao et al. 2000; Strassmann et al. 2000), fungi (Greig and Travisano 2004), plants (Goodnight 1985; Stevens et al. 1995), and animals (Wade 1979; Muir 1996). The second result is that altruistic traits can increase in frequency in real-time experimental evolution (see Rainey and Rainey 2003; Fiegna et al. 2006; Kerr et al. 2006). As the theory predicts, factors influencing social context formation can be critical. Experimental manipulations of population density (Greig and Travisano 2004), interactor relatedness (Griffin et al. 2004), population viscosity (Boots and Mealor 2007), and migration pattern (Kerr et al. 2006) all have effects on the evolution of altruistic traits.

Given the existence of altruism and its potential for evolution by natural selection, what can we say about the level(s) of selection? Does selection act on individual organisms only? On groups of individuals? On groups and individuals simultaneously? Given our discussion of theoretical foundations and experimental findings, what can we add to the debate over the levels of selection?

In the process of answering these questions, I develop an analogy. In 1915, W. E. Hill published the picture shown in figure 19.4. This cartoon is meant to delight its viewer in

FIGURE 19.4
W. E. Hill's famous cartoon.

capturing both the image of an old woman and a young woman. A conversation between someone who can only see the young woman and someone who can only see the old woman would be a frustrating experience for both parties. "It is clearly a young woman that Hill has drawn!" protests the first. "How can you be so blind?!" clamors the second, "it is an *old* woman!" (If the reader is sympathetic with one of these individuals, the following hint may help: the chin and cheek of the young woman is the nose of the old woman.) Hill's cartoon is interesting because it is *simultaneously* a young woman and an old woman. A pluralist would maintain that it is a matter of the viewer's perspective as to which is seen.

For some, the idea of pluralism when it comes to the levels of selection debate is distasteful. How can it be that a case of selection in a group-structured population can be seen as individual-level selection *or* group-level selection? Surely either group selection is occurring or it is not, right? A realist would claim that there is a single answer to the question: "At what level(s) is selection operating?" Note that this answer might be that selection is operating at the group and individual levels simultaneously, in that a realist can hold that multiple selective forces are in action. However, the realist would not claim that selection is operating at the group level *or* the individual level, depending on which way you look at it. This would be pluralist territory.

Why must we concern ourselves with the seemingly esoteric distinction between pluralism and realism? Part of the reason is that a good portion of the group selection debate has occurred between *realists.* Specifically, one set of realists maintains that group selection is not occurring in nearly all biological systems (e.g., Williams 1966), whereas

another set of realists maintains that group selection occurs in plenty of biological systems (e.g., Sober and Wilson 1998). Is it possible that different realists have been arguing over a selective equivalent of Hill's cartoon? Is it simply a matter of perspective, such that selection in group-structured populations can be equivalently understood in two different ways? While I don't believe the situation is so simple, I do think that a thorough consideration of pluralism in the group selection debate leads to a fresh angle on different realist positions and a new take on specific empirical results.

The first question is whether pluralism is even possible in cases of group-structured selection. Let us revisit our simple trait group model. We define π_i to be the fitness of a group with i A types:

$$\pi_i = \alpha_i i + \beta_i(n - i). \tag{26}$$

That is, π_i is simply the total productivity of a group with i altruists (note that $\pi_i = nw_{i\bullet}$). The altruist share of this productivity is simply

$$\phi_i = \frac{\alpha_i i}{\alpha_i i + \beta_i(n - i)}, \tag{27}$$

where $\phi_0 = 0$ and $\phi_n = 1$ always. Note that π's and ϕ's constitute another way to parameterize the fitness structure (see earlier discussion). That is, if all the α's and β's are known, then all the π's and ϕ's can be derived through equations (26) and (27). Similarly, another set of equations can be derived to compute α's and β's given π's and ϕ's (see Kerr and Godfrey-Smith 2002a). Thus, there are two interchangeable perspectives (parametrically speaking) on selection in the trait group framework. Equations (26) and (27) are analogous to telling a perplexed viewer of figure 19.4 that the old woman's mouth is the young woman's necklace, the old woman's eye is the young woman's ear, and so on. Thus, pluralism is certainly *possible* for some cases of group-structured selection.

When working within the multilevel selection framework, the π/ϕ parameterization is a natural choice (see also Wilson 1990). One way to see this is to reconsider the multilevel definition of altruism (equations [4] and [5]) with this new parameterization:

$$\phi_i < i/n \tag{28}$$

$$\pi_i < \pi_{i+1} \tag{29}$$

The π/ϕ parameterization makes it crystal clear that altruist frequency drops within groups (equation [28]), but groups with more altruists are more productive (equation [29]). In the α/β parameterization, only individuals are the bearers of fitness. Groups *affect* individual fitness, but groups do not explicitly *have* fitness. Thus, the α/β perspective is a type of individualist parameterization (see Dugatkin and Reeve 1994; Sterelny 1996; Kerr and Godfrey-Smith 2002a). Statements about *group* productivity require parametric manipulation of individual productivities under the α/β perspective (namely, equation [26]), whereas such statements are made in terms of untouched parameters

within the π/ϕ perspective. In this parameterization, groups are the immediate bearers of fitness, in the form of π quantities.

The realist may respond that it is all fine and good that one can represent the selective process in different ways, but it is an empirical issue whether between-group differences contribute to the evolution of a trait (Wilson and Wilson 2007). A realist might further claim that only one perspective accurately represents the causal structure of the system, while the other distorts it. Indeed, causal language has frequently popped up in the defense of realist positions (e.g., Sober and Wilson 1998 vs. Maynard Smith 2002). The issue of causality is murky, but it is helpful for us to attempt one possible "causal test" here. Given multiple interchangeable parameterizations (e.g., our α/β and π/ϕ parameters), we will say that one parameterization is more "natural" if fewer parameters need to change to accommodate a slight change in the system being modeled. That is, the more causally appropriate parameterization represents small changes to the system in a more isolated way. Put differently, the more natural parameters readily "grab" the changes that can actually occur. Note that this definition of the more natural perspective is relative to the nature of the variant imagined, although there will often be solid empirical reasons for considering certain variants. A forthcoming manuscript describes this near-variant test in detail (P. Godfrey-Smith and B. Kerr, unpublished manuscript).

As an example of the near-variant test, let us revisit the scenario outlined earlier: a group-structured population in which no meaningful interactions occur between individuals; the **A** type has two offspring and the **B** type has a single offspring despite group context. We can describe this system using the α/β parameterization (i.e., $\alpha_i = 2$ and $\beta_i = 1$ for all i) or the π/ϕ parameterization (i.e., $\pi_i = n + i$ and $\phi_i = 2i/(n + i)$). Note that there are between-group differences in productivity (π changes with i). However, even the most ardent defenders of group selection would not call this a case of group selection. Why? Because these group-level differences are nothing more than products of individual level differences. (The argument that Price's first term is spurious in this case has similar roots; see Okasha 2006.) This notion is captured cleanly by the near-variant test. Imagine a slight change to the system: say, the **A** type produces three offspring. In the individualist parameterization, only the α's change, whereas both π's and ϕ's change in the multilevel parameterization. The α/β parameterization captures this variant in a more isolated way, because the parameters attach to the individuals and it is individual fitness that we envision as changeable. We *can* look at this selective episode from a multilevel perspective, but it seems more natural to do so from an individualistic one.

Similarly, there will be cases in which the multilevel perspective is more natural, even though the individualist perspective is available. It is interesting to note that game theory uses an individualist perspective that is identical to our α/β parameterization, while one-locus diploid population genetics uses a multilevel perspective, which is identical to our π/ϕ parameterization. This may reflect the tacit notion that context-dependent payoffs to individuals are changeable in many social games, while it is the fitness of groups of genes aggregated into genotypes that are changeable in many diploid population genetic cases

(see Kerr and Godfrey-Smith 2002a for a discussion). The upshot of the near-variant test is to offer a rigorous way to back up a realist claim. That is, in cases where the pluralist has offered multiple perspectives, the near-variant test may adjudicate among them.

To continue our analogy with Hill's cartoon: Imagine that you find out that Hill was actually drawing a young woman without a necklace, and his daughter placed a dripping cup of coffee on the diagram in such a way as to produce a marking resembling a necklace for the young woman. Hill sees that this necklace can double as a mouth of a previously unseen old woman. Knowing this history, we might claim that Hill's picture more *naturally* depicts a young woman. This is because if his daughter had placed the coffee cup in another location (a near-variant), then the young woman might change slightly, but the old woman entirely disappears. The point is that pluralism need not be antithetical to realism. Indeed, a careful description of multiple perspectives combined with some knowledge about the system allows a judgment about whether one perspective has causal priority, at least by the near-variant test. How do these considerations play out for the experimental work described above?

In Wade's classic experiment on flour beetles, groups are literally being selected by Wade on the basis of their productivity (π_i). Thus, a multilevel perspective is extremely natural in this case, as the different selective treatments are explicitly defined on the basis of effective group productivity. In this case, the realist could legitimately claim that group selection for increased productivity through decreased cannibalism (as gauged by a difference in π's) works against individual selection for decreased productivity through increased cannibalism (as gauged by a depression in ϕ's).

How about quasi-natural selection experiments, where groups are not chosen based on a specified group property? Wilson and Wilson (2007) discuss the results of the Kerr et al. (2006) study on the evolution of phage prudence as a clear case of group selection. Indeed, they call out the authors for not including the "*g*" word in their manuscript. That is, Wilson and Wilson see this case from a realist perspective, where group selection *is* occurring and a full understanding the evolutionary outcome depends on this recognition. In one way, I agree with Wilson and Wilson. It is not only possible to take a perspective that focuses on between-group differences (π's) and within-group skewing (ϕ's), but I find this multilevel perspective particularly illuminating for the Kerr et al. system.[6] However, the heuristic value of a particular representation is different from its causal adequacy. *Must* we take a multilevel perspective in order to properly represent evolution within this experiment? Here, I must confess that I am not as certain as Wilson and Wilson. I freely admit that between-group differences exist in this system, but the question is whether prudence evolves *because* of these differences. To answer this question, causal adequacy must be rigorously defined. Recall that between-group differences can exist in the case of groups of non-interacting individuals; however, we would hesitate to chalk up individual change to these between-group differences.

I have provided one tentative approach to making causal statements more rigorous: the near-variant test. Interestingly, in many models of altruism (using a trait group

framework), the near-variant test does *not* identify one parameterization as being more natural than the other (P. Godfrey-Smith and B. Kerr, unpublished manuscript). That is, multilevel and individualist parameterizations appear equally equipped to isolate changes. In one respect, this "gray zone" is reassuring. After all, it is these cases that have provoked much argument in the group selection debate, with some authors seeing these cases as clearly multilevel selection, while others see these cases as clearly individual-level selection. I would tentatively place the results of the Kerr et al. (2006) experiment and several other quasi-natural experiments in this gray zone. The multilevel perspective can be extremely helpful in understanding evolution, but there is another perspective available, one that does not focus on between-group differences, and it is unclear to me how to sort these different perspectives with regard to causal adequacy.

In the field experiments described earlier, group selection has a very specific meaning. These studies do *not* define differences in group productivity as group selection. Group selection is defined as the differences in an individual's fitness that are due to the collective traits of its group, after controlling for the effects of its own traits. These field studies are directly measuring the effects of social interaction on individual fitness. Thus, contextual analysis is actually pitched at the level of individuals. Altruism detected by contextual analysis could be called individual-centered altruism (see Kerr et al. 2004):

$$\alpha_i < \beta_i \tag{30}$$
$$\alpha_i < \alpha_{i+1} \tag{31}$$
$$\beta_i < \beta_{i+1} \tag{32}$$

Here, the cost of altruism is given by equation (30): as the focal individual switches to the altruistic type but maintains the composition of altruists in its group, its fitness drops. And the benefit is given by equations (31) and (32): as the frequency of altruists increase in the group, the unchanged focal individual improves in fitness. Note that this definition is most easily deployed in the α/β parameterization. Indeed, if groups are not discrete (e.g., as in Stevens et al. 1995) then it is hard to see how to define π_i and ϕ_i (see Godfrey-Smith (2008) for a discussion of this point). Thus, it is not immediately clear that group selection detected through contextual analysis will readily jibe with other definitions of group selection, such as differences in group fitness.

To illustrate this issue, let's revisit the case of frequency-dependent soft selection in a group-structured population. Here, π_i is constant for all values of i, whereas the ϕ_i is different from i/n. This means that all evolutionary change is captured by Price's second term, the term that is generally associated with individual-level selection within a multi-level perspective. However, a contextual analysis picks up a "group selection" effect in this case. Literally, the contextual analysis has picked up an effect of social context on individual fitness. Thus, it has correctly detected exactly what it was designed to detect. However, care should be exercised when communicating this form of "group selection" to an audience that may have something different in mind. It might be most straightforward to claim

that individual fitness depends on group composition, which is consistent with the presence of meaningful social interactions, when a nonzero contextual partial regression coefficient is discovered. Thus, in the *I. capensis* and *A. thaliana* systems, it was found that group composition can affect the fitness of individual plants (Stevens et al. 1995; Weinig et al. 2007). This need not go hand-in-hand with fitness differences at the level of plant groups. Nevertheless, the statistical approach is telling us something important: the potential role of social interactions in determining fitness.

In both laboratory and natural systems, social interactions are ubiquitous. The individualist and multilevel perspectives discussed here highlight different aspects of these interactions. In the spirit of contextual analysis, the individualist perspective explicitly lays out how individuals are affected by their group mates. In the spirit of the Price equation, the multilevel perspective scales up to the group-level effects of these interactions. Kerr and Godfrey-Smith (2002a) suggested that it might be helpful to keep both perspectives on the table when dealing with social evolution in group-structured populations. Indeed, developing the ability to "gestalt switch" may lead to a richer understanding of the evolutionary process.

There is one area where gestalt switching may be particularly appealing. This is in discussions of the so-called major transitions in evolution (Maynard Smith and Szathmáry 1995; Michod 1999; Okasha 2006). When we ask any question about the level of selection in a biological hierarchy, we are taking the existence of the hierarchy for granted. However, an extremely interesting topic is how the hierarchy came into existence in the first place (Okasha 2005, 2006). A major transition can involve the creation of a new level (e.g., prokaryotes associating to produce eukaryotes, single-celled organisms giving rise to multicellular organisms, asocial individuals forming societies). When discussing a transition, we must consider the evolutionary process whereby interactions between previously autonomous lower-level entities generate a higher-level entity, which often can be characterized by lower-level altruism, cohesive integration, and division of labor.

Let us consider such a major transition. At the beginning, suppose that the lower-level entities are completely autonomous and do not interact in any meaningful way. Here an exclusively individualist perspective is natural, where parameters are context-independent fitnesses of lower-level entities. However, suppose that meaningful interactions start to occur within collections of the lower-level entities. We now enter the aforementioned gray zone where *both* an individualist perspective, with parameters that are context-*dependent* lower-level fitnesses, *and* a multilevel perspective, with parameters that are higher-level fitnesses and lower-level skew, are natural options. As the collection becomes more integrated, we enter a situation in which the multilevel perspective is the most natural. Indeed, one notion of "common fate" is that different lower-level entities simultaneously experience fitness changes, and this may be best captured by a parameterization explicitly representing "joint" fitness (e.g., the multilevel perspective's π).

In this sense, it seems likely that selection at higher levels, as captured by differences in higher-level fitness, was crucial to the completion of these major transitions. How precisely such transitions are accomplished is an active area of research, engaging both theoreticians and empirical biologists. Indeed, approaches have been proposed to empirically explore specific transitions (e.g., see Rainey 2007). Some of the same factors thought to influence the evolution of altruism also reappear in discussions of the major transitions. In particular, factors affecting the formation of social context are often seen as important. Thus, experiments manipulating such factors may contribute to a better understanding of the major transitions. Clearly organisms are not islands, but theoretical and experimental approaches to the evolution of social interactions may give us much more than this simple insight—such research may speak to the very origins of organisms themselves.

SUMMARY

An orthodox perspective on natural selection maintains that any favored phenotype improves the fitness of the individual exhibiting it. From such a viewpoint, traits that involve personal sacrifice to increase the fitness of others should be swiftly eliminated. Despite this expectation, altruistic traits are found in many different natural systems, from the fruiting bodies of slime molds to the colonies of eusocial insects. This chapter explores the evolution of altruism using multilevel selection theory. Employing a simple trait group framework, the chapter reviews various definitions of *altruism* and some of the necessary conditions for the evolution of altruism under different definitions. In particular, it discusses the important role of association between altruists, which highlights certain connections between multilevel selection and kin selection. Different statistical approaches to partition selection within group-structured populations (e.g., the Price equation and contextual analysis) are also presented. The experimental evolution literature is explored next, with a survey of research ranging from artificial selection in the laboratory to natural selection in the field. Experiments with animals, plants, and microbes have provided much insight on the conditions promoting the evolution of altruism. The chapter ends with a reconsideration of the levels of selection controversy in light of the empirical and theoretical results.

ACKNOWLEDGMENTS

NOTES

1. For some authors, the discreteness of groups is critical in discussions of group selection (e.g., Maynard Smith 1964, 1976). As Maynard Smith (1976) notes: "For kin selection . . . it is necessary that relatives live close to one another, but it is not necessary (although it may be favorable) that the population be divided into reproductively isolated groups. . . . For group selection, the division into groups which are partially isolated from one another is an essential feature." By Maynard Smith's account, when interactions between relatives are diffuse and overlapping (e.g., when individuals are distributed spatially and interact with neighboring relatives), kin selection is a possibility, but group selection is not (see Godfrey-Smith 2008 for a full discussion of this issue).

2. In another study, McCauley and Wade (1980) investigated factors contributing to high and low group productivity in *T. castaneum* and found that differences in egg fertility, time of development, and the sensitivity to crowding were important. In addition, they also found that differences in egg and pupal cannibalism by adults and larvae influenced group productivity.

3. How "mixing" affects between group variance has been explored elegantly in a separate experiment on the evolution of egg cannibalism (Wade 1980). Wade allowed *T. confusum* to evolve in metapopulations where two components of population structure were manipulated in a full factorial design: (1) mating (breeding took place exclusively within groups or at random) and (2) social interaction (larvae were offered eggs to potentially cannibalize of varying degrees of relatedness). When breeding was constrained to occur within groups, the beetles evolved lower cannibalism rates when encountering eggs with a higher degree of relatedness—consistent with kin selection theory. Under random mating, no significant differences in evolved cannibalism rates were discovered across social interaction treatments. Wade suggested that within-group breeding tended to promote a higher between-group variance in productivity, whereas the random breeding tended to homogenize groups (see Wade and Breden (1981) for a theoretical treatment of these issues).

4. Incidentally, Wade and colleagues (Wade 1982; Wade and McCauley 1984; Wade and Goodnight 1991) did perform a series of metapopulation experiments in which some migration between subpopulations accompanied the founding propagules. While the metapopulation treatments without migration generally promoted the largest variance in population size between demes (this variance is necessary for "interdemic" selection), significant variation was also discovered between demes in treatments with substantial migration (e.g., where 25 percent of each founding subpopulation were migrants). Furthermore, selection for population productivity produced a response in these higher migration treatments.

5. Griffing (1976a, 1976b) labels the influence of group mates on a focal individual as "associate effects." Using a theoretical approach, he discusses the benefits of different breeding strategies in the presence of such effects.

6. Of course, there are other valuable approaches available. West et al. (2008) argue that the Kerr et al. (2006) experiment can be viewed from a kin selection perspective. While noting connections between kin and group selection approaches, West et al. (2007, 2008) argue that there are general advantages to kin selection methodologies. In response, Wilson (2008) and Wilson and Wilson (2007) defend the use of multilevel selection approaches. It is telling how each set of authors describes the Kerr et al. (2006) experiment. West et al. (2007) state, "A more efficient use of host resources is favored when there is a higher relatedness between

the phage infecting a bacterium—local migration leads to a higher relatedness and hence se-
lects for lower virulence," whereas Wilson (2007) states, "Biologically plausible migration rates
enabled 'prudent' phage strains to outcompete more 'rapacious' strains in the metapopula-
tion despite their selective disadvantage within each well, exactly as envisioned by Wynne-Ed-
wards." I would argue that West et al. are focusing on social context formation (how different
types find themselves in specific neighborhoods of interaction), while Wilson is focusing on
the fitness structure (within-group disadvantage and between-group advantage of the prudent
type). Of course, both social context formation and the fitness structure are critical to the
prospects for the evolution of restraint (as I suspect all of these authors would readily admit).
The original Kerr et al. (2006) manuscript mentioned neither group selection nor kin selection,
but it does make statements that resonate with each approach, in part because each approach
does focus on something important within this system.

REFERENCES

Abedon, S. T., P. Hyman, and C. Thomas. 2003. Experimental examination of bacteriophage latent-period evolution as a response to bacterial availability. *Applied and Environmental Microbiology* 69:7499–7506.

Anderson, R. M., and R. M. May. 1982. Coevolution of hosts and parasites. *Parasitology* 85:411–426.

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

Baer, C. F., J. Travis, and K. Higgins. 2000. Experimental evolution in *Heterandria formosa*, a livebearing fish: Group selection on population size. *Genetical Research* 76:169–178.

Bargum, K., and L. Sundstrom. 2007. Multiple breeders, breeder shifts and inclusive fitness returns in an ant. *Proceedings of the Royal Society B, Biological Sciences* 274:1547–1551.

Boots, M., and M. Mealor. 2007. Local interactions select for lower pathogen infectivity. *Science* 315:1284–1286.

Bull, J. J. 1994. Perspective: Virulence. *Evolution* 48:1423–1437.

Bull, J. J., I. J. Molineux, and W. R. Rice. 1991. Selection of benevolence in a host-parasite system. *Evolution* 45:875–882.

Cohen, D., and I. Eshel. 1976. On the founder effect and the evolution of altruistic traits. *Theoretical Population Biology* 10:276–302.

Cooper, V. S., M. H. Reiskind, J. A. Miller, K. A. Shelton, B. A. Walther, J. S. Elkinton, and P. W. Ewald. 2002. Timing of transmission and the evolution of virulence of an insect virus. *Proceedings of the Royal Society of London B, Biological Sciences* 269:1161–1165.

Craig, D. M. 1982. Group selection versus individual selection: An experimental analysis. *Evolution* 36:271–282.

Craig, J. V., M. L. Jan, C. R. Polley, A. L. Bhagwat, and A. D. Dayton. 1975. Changes in relative aggressiveness and social dominance associated with selection for early egg-production in chickens. *Poultry Science* 54:1647–1658.

Craig, J. V., and W. M. Muir. 1996. Group selection for adaptation to multiple-hen cages: Beak-related mortality, feathering, and body weight responses. *Poultry Science* 75:294–302.

Dao, D. N., R. H. Kessin, and H. L. Ennis. 2000. Developmental cheating and the evolutionary biology of *Dictyostelium* and *Myxococcus*. *Microbiology—UK* 146:1505–1512.

Darwin, C. 1859. On the Origin of Species. London: Murray.

Dennis, R. L., W. M. Muir, and H. W. Cheng. 2006. Effects of raclopride on aggression and stress in diversely selected chicken lines. *Behavioural Brain Research* 175:104–111.

Donne, J. 1624. Meditation XVII. Pages 574–575 *in* H. Alford, ed. *The Works of John Donne* III. London: Parker, 1839.

Donohue, K. 2003. The influence of neighbor relatedness on multilevel selection in the Great Lakes sea rocket. *American Naturalist* 162:77–92.

———. 2004. Density-dependent multilevel selection in the Great Lakes sea rocket. *Ecology* 85:180–191.

Dugatkin, L. A., and H. K. Reeve. 1994. Behavioral ecology and the levels of selection- dissolving the group selection controversy. Pages 101–133 *in Advances in the Study of Behavior, Vol. 23. Advances in the Study of Behavior.*

Emerson, A. E. 1939. Social coordination and the superorganism. *American Midland Naturalist* 21:182–209.

Eshel, I. 1972. On the neighbor effect and the evolution of altruistic traits. *Theoretical Population Biology* 3:258–277.

Fiegna, F., Y. T. N. Yu, S. V. Kadam, and G. J. Velicer. 2006. Evolution of an obligate social cheater to a superior cooperator. *Nature* 441:310–314.

Fletcher, J. A., and M. Zwick. 2006. Unifying the theories of inclusive fitness and reciprocal altruism. *American Naturalist* 168:252–262.

Fletcher, J. A., M. Zwick, M. Doebeli, and D. S. Wilson. 2006. What's wrong with inclusive fitness? *Trends in Ecology & Evolution* 21:597–598.

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

Gilbert, O. M., K. R. Foster, N. J. Mehdiabadi, J. E. Strassmann, and D. C. Queller. 2007. High relatedness maintains multicellular cooperation in a social amoeba by controlling cheater mutants. *Proceedings of the National Academy of Sciences of the USA* 104:8913–8917.

Godfrey-Smith, P. 2008. Varieties of population structure and the levels of selection. *British Journal for the Philosophy of Science* 59:25–50.

Goodnight, C. J. 1985. The influence of environmental variation on group and individual selection in a cress. *Evolution* 39:545–558.

Goodnight, C. J., J. M. Schwartz, and L. Stevens. 1992. Contextual analysis of models of group selection, soft selection, hard selection, and the evolution of altruism. *American Naturalist* 140:743–761.

Goodnight, C. J., and L. Stevens. 1997. Experimental studies of group selection: What do they tell us about group selection in nature? *American Naturalist* 150:S59–S79.

Greig, D., and M. Travisano. 2004. The Prisoner's Dilemma and polymorphism in yeast SUC genes. *Proceedings of the Royal Society of London B, Biological Sciences* 271:S25–S26.

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

Griffing, B. 1976a. Selection in reference to biological groups. V. Analysis of full-sib groups. *Genetics* 82:703–722.

———. 1976b. Selection in reference to biological groups. VI. Use of extreme forms of nonrandom groups to increase selection efficiency. *Genetics* 82:723–731.

Hamilton, W. D. 1963. Evolution of altruistic behavior. *American Naturalist* 97:354–356.

———. 1964. Genetical evolution of social behavior I. *Journal of Theoretical Biology* 7:1–16.

———. 1975. Innate social aptitudes in man: An approach from evolutionary genetics. Pages 133–155 *in* R. Fox, ed. *Biosocial Anthropology.* New York: Wiley.

Hardin, G. 1968. Tragedy of the commons. *Science* 162:1243–1248.

Harrison, S., and A. Hastings. 1996. How effective is interdemic selection? Reply. *Trends in Ecology & Evolution* 11:299–299.

Heisler, I. L., and J. Damuth. 1987. A method for analyzing selection in hierarchically structured populations. *American Naturalist* 130:582–602.

Jeon, K. W. 1972. Development of cellular dependence on infective organisms: Microsurgical studies in amoebas. *Science* 176:1122.

Jeon, K. W., and J. C. Hah. 1977. Effect of chloramphenicol on bacterial endosymbiotes in a strain of *Amoeba proteus. Journal of Protozoology* 24:289–293.

Jones, C. G., J. H. Lawton, and M. Shachak. 1994. Organisms as ecosystem engineers. *Oikos* 69:373–386.

Karlin, S., and C. Matessi. 1983. Kin selection and altruism. *Proceedings of the Royal Society of London B, Biological Sciences* 219:327–353.

Kerr, B., and P. Godfrey-Smith. 2002a. Individualist and multilevel perspectives on selection in structured populations. *Biology & Philosophy* 17:477–517.

———. 2002b. On Price's equation and average fitness. *Biology & Philosophy* 17:551–565.

Kerr, B., P. Godfrey-Smith, and M. W. Feldman. 2004. What is altruism? *Trends in Ecology & Evolution* 19:135–140.

Kerr, B., C. Neuhauser, B. J. M. Bohannan, and A. M. Dean. 2006. Local migration promotes competitive restraint in a host-pathogen "tragedy of the commons." *Nature* 442:75–78.

Krakauer, A. H. 2005. Kin selection and cooperative courtship in wild turkeys. *Nature* 434:69–72.

Lehmann, L., L. Keller, S. West, and D. Roze. 2007. Group selection and kin selection: Two concepts but one process. *Proceedings of the National Academy of Sciences of the USA* 104:6736–6739.

Lewontin, R. C. 1970. The units of selection. *Annual Review of Ecology and Systematics* 1:1–18.

Lewontin, R. C., and L. C. Dunn. 1960. The evolutionary dynamics of a polymorphism in the house mouse. *Genetics* 45:705–722.

Matessi, C., and S. D. Jayakar. 1976. Conditions for the evolution of altruism under Darwinian selection. *Theoretical Population Biology* 9:360–387.

Matessi, C., and S. Karlin. 1984. On the evolution of altruism by kin selection. *Proceedings of the National Academy of Sciences of the USA* 81:1754–1758.

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

———. 1976. Group selection. *Quarterly Review of Biology* 51:277–283.

———. 2002. Commentary on Kerr and Godfrey-Smith. *Biology & Philosophy* 17:523–527.

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

McCauley, D. E., and M. J. Wade. 1980. Group selection: The genetic and demographic basis for the phenotypic differentiation of small populations of *Tribolium castaneum. Evolution* 34:813–821.

McLeish, M. J., T. W. Chapman, and B. J. Crespi. 2006. Inbreeding ancestors: The role of sibmating in the social evolution of gall thrips. *Journal of Heredity* 97:31–38.

Mehdiabadi, N. J., C. N. Jack, T. T. Farnham, T. G. Platt, S. E. Kalla, G. Shaulsky, D. C. Queller et al. 2006. Kin preference in a social microbe: Given the right circumstances, even an amoeba chooses to be altruistic towards its relatives. *Nature* 442:881–882.

Messenger, S. L., I. J. Molineux, and J. J. Bull. 1999. Virulence evolution in a virus obeys a trade-off. *Proceedings of the Royal Society of London B, Biological Science*s 266:397–404.

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

Muir, W. M. 1996. Group selection for adaptation to multiple-hen cages: Selection program and direct responses. *Poultry Science* 75:447–458.

Nunney, L. 1985. Group selection, altruism, and structured-deme models. *American Naturalist* 126:212–230.

Odling-Smee, F. J., K. N. Laland, and M. W. Feldman. 2003. *Niche Construction: The Neglected Process in Evolution.* Princeton, NJ: Princeton University Press.

Okasha, S. 2005. Multilevel selection and the major transitions in evolution. *Philosophy of Science* 72:1013–1025.

———. 2006. *Evolution and the Levels of Selection.* Oxford: Oxford University Press.

Prado, F., and B. Kerr. 2008. The evolution of restraint in bacterial biofilms under nontransitive competition. *Evolution* 62: 538–548.

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

———. 1972. Extension of covariance selection mathematics. *Annals of Human Genetics* 35:485–490.

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

Queller, D. C., E. Ponte, S. Bozzaro, and J. E. Strassmann. 2003. Single-gene greenbeard effects in the social amoeba *Dictyostelium discoideum. Science* 299:105–106.

Rainey, P. B. 2007. Unity from conflict. *Nature* 446:616–616.

Rainey, P. B., and K. Rainey. 2003. Evolution of cooperation and conflict in experimental bacterial populations. *Nature* 425:72–74.

Sachs, J. L., and J. J. Bull. 2005. Experimental evolution of conflict mediation between genomes. *Proceedings of the National Academy of Sciences of the USA* 102:390–395.

Sachs, J. L., and T. P. Wilcox. 2006. A shift to parasitism in the jellyfish symbiont *Symbiodinium microadriaticum. Proceedings of the Royal Society of London B, Biological Sciences* 273:425–429.

Scheiner, S. M. 2002. Selection experiments and the study of phenotypic plasticity. *Journal of Evolutionary Biology* 15:889–898.

Sober, E. 1984. *The Nature of Selection.* Chicago: University of Chicago Press.

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

Sterelny, K. 1996. The return of the group. *Philosophy of Science* 63:562–584.

Stevens, L., C. J. Goodnight, and S. Kalisz. 1995. Multilevel selection in natural populations of *Impatiens capensis. American Naturalist* 145:513–526.

Strassmann, J. E., Y. Zhu, and D. C. Queller. 2000. Altruism and social cheating in the social amoeba *Dictyostelium discoideum. Nature* 408:965–967.

Swenson, W., J. Arendt, and D. S. Wilson. 2000a. Artificial selection of microbial ecosystems for 3-chloroaniline biodegradation. *Environmental Microbiology* 2:564–571.

Swenson, W., D. S. Wilson, and R. Elias. 2000b. Artificial ecosystem selection. *Proceedings of the National Academy of Sciences of the USA* 97:9110–9114.

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

Trivers, R. L. 1971. Evolution of reciprocal altruism. *Quarterly Review of Biology* 46:35–&.

Turner, P. E., and L. Chao. 1999. Prisoner's dilemma in an RNA virus. *Nature* 398:441–443.

Uyenoyama, M., and M. W. Feldman. 1980. Theories of kin and group selection: A population-genetics perspective. *Theoretical Population Biology* 17:380–414.

Velicer, G. J., L. Kroos, and R. E. Lenski. 1998. Loss of social behaviors by Myxococcus xanthus during evolution in an unstructured habitat. *Proceedings of the National Academy of Sciences of the USA* 95:12376–12380.

Wade, M. J. 1977. Experimental study of group selection. *Evolution* 31:134–153.

———. 1979. Primary characteristics of *Tribolium* populations group selected for increased and decreased population size. *Evolution* 33:749–764.

———. 1980a. An experimental study of kin selection. *Evolution* 34:844–855.

———. 1980b. Group selection, population growth rate, and competitive ability in the flour beetles, *Tribolium* species. *Ecology* 61:1056–1064.

———. 1980c. Kin selection: Its components. *Science* 210:665–667.

———. 1982. Group selection: migration and the differentiation of small populations. *Evolution* 36:949–961.

Wade, M. J., and F. Breden. 1981. Effect of inbreeding on the evolution of altruistic behavior by kin selection. *Evolution* 35:844–858.

Wade, M. J., and C. J. Goodnight. 1991. Wright's shifting balance theory: An experimental study. *Science* 253:1015–1018.

Wade, M. J., and D. E. McCauley. 1980. Group selection: The phenotypic and genotypic differentiation of small populations. *Evolution* 34:799–812.

———. 1984. Group selection: The interaction of local deme size and migration in the differentiation of small populations. *Evolution* 38:1047–1058.

Weinig, C., J. A. Johnston, C. G. Willis, and J. N. Maloof. 2007. Antagonistic multilevel selection on size and architecture in variable density settings. *Evolution* 61:58–67.

West, S. A., A. S. Griffin, and A. Gardner. 2007. Social semantics: Altruism, cooperation, mutualism, strong reciprocity and group selection. *Journal of Evolutionary Biology* 20:415–432.

———. 2008. Social semantics: How useful has group selection been? *Journal of Evolutionary Biology* 21:374–385.

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

———. 1971. *Group Selection.* Chicago: Aldine Atherton.

Wilson, D. S. 1975. Theory of group selection. *Proceedings of the National Academy of Sciences of the USA* 72:143–146.

———. 1977. Structured demes and evolution of group-advantageous traits. *American Naturalist* 111:157–185.

———. 1980. *The Natural Selection of Populations and Communities.* Menlo Park, CA: Benjamin/Cummings.

———. 1983. The group selection controversy: History and current status. *Annual Review of Ecology and Systematics* 14:159–187.

———. 1990. Weak altruism, strong group selection. *Oikos* 59:135–140.

———. 2008. Social semantics: Toward a genuine pluralism in the study of social behaviour. *Journal of Evolutionary Biology* 21:368–373.

Wilson, D. S., and E. O. Wilson. 2007. Rethinking the theoretical foundation of sociobiology. *Quarterly Review of Biology* 82:327–348.

Wilson, R. A. 2004. Test cases, resolvability, and group selection: A critical examination of the myxoma case. *Philosophy of Science* 71:380–401.

Wright, S. 1945. Tempo and mode in evolution: A critical review. *Ecology* 26:415–419.

Wynne-Edwards, V. C. 1962. *Animal Dispersion in Relation to Social Behavior*. New York: Hafner.