

## Précis of *Unto Others*

DAVID SLOAN WILSON  
*Binghamton University*

ELLIOTT SOBER  
*University of Wisconsin, Madison, and London School of Economics*

Altruism has both an evolutionary and a psychological meaning. As the term is used in evolutionary theory, a trait is deemed altruistic if it reduces the fitness of the actor and enhances the fitness of someone else. In its psychological sense, the thesis that we have altruistic ultimate motives asserts that we care about the welfare of others, not just as a means of enhancing our own well-being, but as an end in itself. Since mindless organisms can be evolutionarily altruistic, evolutionary altruism does not entail psychological altruism. And since caring about the welfare of others can, unbeknownst to the actor, enhance the actor's fitness, neither does psychological altruism entail evolutionary altruism. If parents care about their children as ends in themselves, these parents are psychological altruists; but if this caring allows parents to be more reproductively successful, this is not an instance of evolutionary altruism. In *Unto Others* (hereafter UO), we consider both evolutionary altruism (Part I) and psychological altruism (Part II) from an evolutionary perspective.

It is a challenge to explain how evolutionary altruism can evolve by the process of natural selection, since altruists in a group will be less fit than the selfish individuals in the same group who receive benefits but do not make donations of their own. Darwin proposed a theory of group selection to solve this puzzle. Very simply, even though altruists are less fit than selfish individuals within any single group, groups of altruists are more fit than groups of selfish individuals. If a population is subdivided into many groups that vary in their altruistic tendencies, altruism will be favored at the level of selection among groups even as it is being disfavored at the level of selection among individuals within groups. Darwin's scenario became the basis for a theoretical framework called multilevel selection theory.

We think that Darwin was on the right track. Altruistic behaviors can evolve by group selection; the process of individual selection, on the other hand, because it favors individuals who are more fit over other individuals in

the same group who are less, favors selfishness, not altruism. Examples of altruism have been documented in many species and have been produced in the laboratory by artificial multilevel selection experiments. Furthermore, group selection was probably an exceptionally strong force in human evolution, accounting for our own groupish nature.

The history of scientific investigation of evolutionary altruism is not one of straightforward scientific progress. Group selection was rejected as an important evolutionary force by most biologists in the 1960's. Other theories were proposed to explain the evolution of seemingly altruistic behaviors in individualistic terms, such as kin selection theory (helping one's own genes in the bodies of others) and reciprocal altruism (helping in expectation of return benefits). It became so taboo to invoke group selection that, as we recount on p. 40 of *UO*, one elder statesman advised a young colleague that "there are three ideas that you do not invoke in biology—Lamarckism, the phlogiston theory, and group selection." Yet, upon closer examination, many of the arguments used against group selection appear to be confused, question-begging, or inconclusive, and the mechanisms that were thought to be alternatives to group selection turn out to be instances of group selection, properly conceived.

Thus, if our account of altruism as a product of group selection is correct, it amounts to a reversal of what many regard as a triumphant step forward in evolutionary thought. Establishing a new theory is difficult enough; establishing a heretical old theory is even harder, requiring a detailed historical reconstruction of ideas and a general consideration of why science sometimes departs from the smooth process it is often idealized to follow. These elements make *UO* relevant to historians and philosophers of science in general, apart from the specific topic of altruism. As we state on p. 13, "It is wonderful when science progresses by a straight march to the truth. But science is even more interesting when its progress is less direct. To understand the process of science, we need to consider not just the destination reached at the end of the journey but the false starts and detours encountered along the way."

Our presentation of multilevel selection theory expands the range of topics beyond altruism in other respects as well. Group selection favors any trait that causes some groups to survive and reproduce better than others. Altruism qualifies as such a trait but so also do other traits that do not appear altruistic, such as moralistic aggression and other forms of social control. As we state on p. 31 of *UO*, "the extreme altruism that cries out for explanation might be rare in nature for two very different reasons—(a) because group selection is seldom strong enough to evolve such behaviors, or (b) because there is usually a way to benefit the group without such extreme self-sacrifice." In short, the same theory that explains why evolutionary altruism exists also

explains why altruism shares the stage with many other traits that cause groups to function as adaptive units.

Human groups often appear to be functionally organized, even when they are large and composed of unrelated individuals. The founders of the social sciences were comfortable with the idea of societal organisms, complete with group minds, but this perspective was largely rejected by social scientists inclined to defend methodological individualism and by evolutionary biologists alike during the second half of the 20th century. Multilevel selection theory provides a new foundation for understanding the group-level functionality of human society in addition to the opposing forces that so often subvert group-level functionality from within. The mechanisms that enable human groups to function as adaptive units provide ample scope for cultural processes. Our theoretical framework is evolutionary, but it does not conform to the doctrine of genetic determinism that has so often been criticized. Indeed, we devote an entire chapter to a survey of 25 cultures, chosen at random from the Human Relations Area Files, an anthropological data base for cross-cultural research. Multilevel selection theory can help explain variation among cultures in addition to their commonalities and (partial) functionality above the level of the individual.

Evolutionary altruism has been studied by biologists without much regard to the psychological component. However, the question of whether people are motivated to help others as an end in itself, rather than as a means for advancing self-interest, has a long history of its own within both philosophy and psychology. In Part II of *UO*, we evaluate the arguments, both conceptual and empirical, that philosophers and psychologists have advanced in an attempt to choose between psychological egoism and motivational pluralism. We conclude that these arguments are inconclusive. We suggest that it is illuminating to analyze these two motivational theories from an evolutionary perspective. In the past, psychological egoism has been regarded as a more parsimonious explanation of behavior than psychological altruism, resulting in an intellectual pecking order in which egoistic explanations are innocent until proven guilty. From an evolutionary perspective, psychological egoism and altruism are alternative proximate mechanisms that might have evolved to motivate adaptive behaviors. To see which mechanism is more likely to have evolved, we must consider both from the standpoint of reliability, availability, and efficiency. Psychological egoism is the inferior explanation when judged by these criteria. An exclusively egoistic motivation would constitute an unreliable and needlessly complex Rube Goldberg device when it comes to producing helping behavior. If helping behavior evolves among organisms that have minds as sophisticated as those exhibited by human beings, it most likely will be motivated at least in part by a desire to help others as a psychological end in itself.

UO uses altruism as a lens to study a wide variety of subjects, including foundational issues in evolutionary biology, the human social sciences, and philosophy. We are pleased that it is the subject of this book symposium in *Philosophy and Phenomenological Research* and we look forward to addressing some of the issues in more detail in our reply to the commentaries.

# Reply to Commentaries

DAVID SLOAN WILSON  
*Binghamton University*

ELLIOTT SOBER  
*University of Wisconsin, Madison, and London School of Economics*

**Barrett and Godfrey-Smith, Dennett, and Skyrms** focus on part I of UO, which concerns evolutionary altruism; **Jamieson** comments exclusively on part II, which is about psychological altruism. This 3-to-1 ratio is reflected in our reply.

## **Part I—Evolutionary Issues**

We are encouraged by the opening paragraph of each commentary and especially by **Barrett and Godfrey-Smith's** statement that "the book is entirely persuasive in its argument that attempts to marginalize group-selectionist ideas in the latter part of the 20<sup>th</sup> century were mistaken." At the very least, the commentators who addressed Part I of UO agree that the issues surrounding multilevel selection theory, however unsettled, are highly relevant to modern evolutionary research. This by itself is an important advance because for decades the idea of group selection has been portrayed as a settled issue that needs to be learned only to avoid an erroneous pattern of thought.

The substance of the commentaries, however, reveals considerable disagreement about how UO conceptualizes the idea of group selection. **Dennett** describes the issues as "mind-twistingly elusive and slippery" and hints that it is mere hype to say that group selection has been revived. **Barrett and Godfrey-Smith** discuss the problem of multiple perspectives at length and claim that we are too liberal in our definition of groups. We believe that these criticisms obscure the simplicity of the basic question that group selection theory is designed to answer and the historical continuity of that theory, from Darwin to the present.

## **The basic evolutionary question**

The question at the center of the group selection controversy is so simple that it can be stated in a single sentence: *Do traits evolve by benefiting whole groups, despite being selectively neutral or disadvantageous within groups?*

Obviously, posing this question requires a precise definition of groups, but if we have a trait in mind and can agree on what counts as a group, answering the question is in principle straightforward. First we need to measure the fitness of individuals possessing the trait, relative to the fitness of individuals possessing alternative traits in the same group. This information is necessary to confirm that the trait is indeed selectively neutral or disadvantageous within groups. If the trait evolves without being favored by selection within groups, something else is required to explain its evolution. That “something” could be genetic drift, pleiotropic effects, or a process of between-group selection. To confirm that the trait evolves by benefiting whole groups (the first part of the basic question), we need to measure how the trait varies among groups and whether this variation leads groups to make differential contributions to the total gene pool. Finally, we must do the math to show that the evolutionary force favoring the trait—selection among groups in the total population—is stronger than the opposing force of selection among individuals within groups. Carrying out these steps may involve methodological difficulties, but philosophically it is no more problematic to identify and measure the levels of selection than it is to measure, say, the relative importance of predation and sexual selection in the evolution of coloration in guppies. There is nothing “mind-twistingly elusive and slippery” about it. The problem becomes somewhat more difficult if an altruistic trait evolves to fixation, destroying the variation responsible for its selection. However, this problem is not unique to group selection and so we set it aside here.

Now that we have described the simplicity of the basic question, we need to establish its continuity with the understanding of group selection that was in place historically. This is important because **Dennett and Barrett and Godfrey-Smith** suggest that we have departed from past conventional usage. *UO* includes a more detailed history of the group selection controversy than any other published account (see also Wilson 1983). Here we will list six important events, from Darwin to the present, that help demonstrate historical continuity:

- 1) Darwin (1871, p. 166) proposed the original theory of group selection in part to explain the evolution of human morality: “It must not be forgotten that although a high standard of morality gives but a slight or no advantage to each individual man and his children over other men of the same tribe, yet that an increase in the number of well-endowed men and advancement in the standard of morality will certainly give an immense advantage to one tribe over another.” In this passage and elsewhere, Darwin clearly regarded features of human morality as selectively neutral or disadvantageous within groups and relied upon selection among groups to explain their evo-

lution, following the same logic that we described in our formulation of the basic question.

- 2) One of the first mathematical models of group selection is by Sewall Wright (1945), who considered a gene that decreases the relative fitness of the individual within groups by a factor  $(1-s)$  while increasing the fitness of the group by a factor  $(1+pb)$ , where  $p$  is the frequency of the gene in the group. Again, the use of between-group selection to explain the evolution of traits that are selectively disadvantageous or neutral within groups could not be clearer.
- 3) The rejection of group selection in the 1960's is attributed largely to G.C. Williams (1966) who wrote "It is universally conceded by those who have seriously concerned themselves with this problem that...group-related adaptations must be attributed to the natural selection of alternative groups of individuals and that the natural selection of alternative alleles within populations will be opposed to this development." The correspondence between this statement and our formulation of the basic question is obvious, and Williams himself comments on the historical continuity of thought on the subject. Williams' point here is not to change the theoretical framework of multilevel selection but merely to make the claim that group selection is invariably weak, compared to selection within groups. He defended this rejection of group selection empirically (e.g., in his discussion of sex ratio evolution), but, unfortunately, he also provided a set of more-or-less a priori arguments against the idea of group adaptation. We regard these a priori arguments (which Dawkins repeated in *The Selfish Gene*) as muddying the waters. It is pointless to formulate evolutionary hypotheses that can be refuted on a priori grounds. Group selection, like individual selection itself, may or may not be the right explanation of a given phenomenon, but surely this question has to be settled by consulting data. Whenever Williams talked about group selection in a way that could be tested empirically, as opposed to being decidable on a priori grounds, he followed the basic formulation outlined above.
- 4) Another mathematical model of group selection was formulated by George Price in the 1970's and has become influential in the modern multilevel selection literature. Price used a covariance approach to break total gene frequency change into two components; average gene frequency change within single groups, and gene frequency change caused by variation among groups. The two components of the Price equation obviously correspond to within- and between-

group selection, as previously understood. Thus, it is completely unsurprising that Price (and W.D. Hamilton) regarded his equation as a multilevel selection model.

- 5) Even modern-day biologists who reject group selection subscribe to our formulation of the basic question. In numerous editions of his influential textbook *Animal Behavior*, John Alcock discussed and dismissed group selection with the help of a Gary Larson cartoon showing a group of fabled lemmings running into the sea (ostensibly to regulate their population size); one of their number wears a sly smile and an inner tube. Alcock's caption to the cartoon is entitled "group selection's fatal flaw" and states "if most lemmings did commit suicide, wouldn't a mutant type that refrained from suicide quickly spread through the population?" The message of the cartoon and caption is that within-group selection is all that matters as far as natural selection is concerned. There is no need to consider multiple groups, variation among groups, and all the other paraphernalia associated with group selection. Like Williams, Alcock agrees with the basic formulation described above and merely asserts that group selection is too weak to matter. Everything that evolves by selection does so on the strength of within-group selection.
- 6) Biologists who now work within the framework of multilevel selection theory speak easily of selection within and among groups, as if there is no cause for controversy. In the context of his discussion of parasite evolution, Bull (1994, p. 1425) states: "both levels operate together in what may be regarded as a classic group selection hierarchy, the groups being the populations of parasites within hosts." In discussing sex-ratio evolution, Herre (1999, p. 217) states "the key to understanding how much greater or lesser degrees of female-bias are favored lies in understanding how among- and within-group selection are balanced." A recent article in the journal *Nature* on altruistic punishment in humans (Fehr and Gächter 2002) is accompanied by a commentary that states "A plausible explanation of the evolutionary success of this strategy is that groups with a high fraction of altruistic punishers would have sustained cooperation more successfully than groups with fewer punishers, and so would have prevailed over them" (Bowles and Gintis 2002). The similarity between this statement and Darwin's original scenario for the evolution of human morality is unmistakable.

We therefore continue to maintain that our understanding of what group selection amounts to has the virtue of both conceptual clarity and historical conti-

nunity with the problem as it has developed in biology. More details about the history are provided in UO. Since **Dennett and Barrett and Godfrey-Smith** think that we have broken with the past, they should support their claim with scholarship of their own.

### Defining groups

According to **Barrett and Godfrey-Smith**, we employ a liberal definition of groups that allows almost anything to be categorized as group selection. Again, the implication is that we have departed from past definitions of groups that were narrower, but this is not the case. One reason that early group selectionists were called naïve is that they saw adaptive groups everywhere, from ephemeral groups such as bird flocks and fish schools to more permanent groups such as entire species and ecosystems. More careful early discussions of group selection are similarly diverse. As we state on p. 92 of UO, “For Darwin (1871), groups were tribes that compete by direct conflict. For Haldane (1932), groups were tribes that compete by fissioning at different rates. For Wright (1945), groups were isolated populations that compete in the colonization of new groups. For Williams and Williams (1957), groups were sib-groups that last only a fraction of a generation. For Hamilton (1975), groups were any set of individuals that form for a period of the life cycle and influence each other’s fitness.”

There is a reason why multilevel selection theory—and all other evolutionary theories of social behavior—must be flexible in their definition of groups. The purpose of an evolutionary model is usually to predict and explain the evolution of a particular trait. When the trait is a nonsocial behavior that alters the fitness of the individual alone, we needn’t concern ourselves with groups. But when the trait is a social behavior, the fitness of an individual is determined by its own trait and by the traits of the individuals with whom it interacts. These individuals constitute the group, which must be identified accurately to calculate the fitnesses that determine the outcome of evolution.

Suppose that we are interested in the evolution of warning calls in birds. Birds that look for predators and call when one is spotted benefit everyone in earshot at their own expense. The fitness of an individual depends on whether it is a caller and on the number of other callers within earshot. Thus, something like a bird flock is the appropriate grouping for this particular trait. Callers have a lower fitness than non-callers within any given flock, but groups with more callers contribute more to the total gene pool than groups with fewer. This constitutes a multilevel selection model for warning calls in birds.

In contrast, suppose we are interested in the evolution of resource conservation in birds. Birds that eat moderately have fewer offspring than more glut-

tonous birds in the same group, but groups of birds that eat moderately conserve their resources while groups of gluttonous birds overexploit their resources and go extinct, as proposed by Wynne-Edwards (1962), the most vilified of the early group selectionists. A group in this case is the set of birds that share the same resources. This may include a much larger number of birds than those who are within earshot of each other. If so, we must expand the size of the group to include the individuals that influence each other's fitness in order to determine the outcome of evolution for this particular trait. This is why we emphasized in UO that what counts as a group must be determined on a trait-by-trait basis.

Against this background, let's consider **Barrett and Godfrey-Smith's** claim that our definition of groups is too liberal. As we have noted, if a theory is going to consider a variety of traits, its definition of groups *must* be liberal (=flexible) to encompass the groupings appropriate for each trait. However, for any particular trait, the definition of groups *cannot* be liberal; there is a single appropriate grouping and all others will simply come to the wrong conclusion concerning how the evolutionary process proceeds. This combination of liberality across traits and conservatism with respect to individual traits characterizes the multilevel selection literature, and it is fully warranted. Furthermore, all evolutionary theories of social behavior are bound by the same rules. Consider a model of fighting behavior in which one type of individual shares amiably while another type fights to take all. Fighters are more fit than sharers within any given group, but groups of sharers contribute more to the gene pool than groups of fighters. This is a multilevel selection model, identical in form to our models of warning cries and resource conservation. It is also a game theory model. In both cases, we must find the appropriate groups for fighting behavior. If individuals interact in pairs, then pairs are the appropriate groups. If individuals interact in triads or free-for-alls of 100, then those are the appropriate groups. N-person game theory is liberal in its definition of groups—and must be—in exactly the same sense that multilevel selection theory is liberal. Game theory models of warning cries, resource conservation, and fighting behavior must arrive at the same definition of groups as multilevel selection models of the same traits. Groups are determined by the biology of the organism, not the whim of the biologist or the name of the theoretical framework being employed.

So far we have defended our own conception of groups. What of **Barrett and Godfrey-Smith's** more narrow conception? They wish to restrict groups to highly functional groupings such as bee colonies and individual organisms (which are functional groupings of lower-level agents). As we discuss on p. 97 of UO, this definition fails to distinguish between the *product* and the *process* of multilevel selection. Bee colonies aren't groups because they are highly functional; rather, they are highly functional because

bees (or their ancestors) lived in groups before they were highly functional and selection among groups prevailed against selection within groups. Darwin did not say that human groups are groups because they are moral; rather, he said that morality evolved by group selection because more moral groups replaced less moral groups. Another defect in **Barrett and Godfrey-Smith's** proposal is that it does not allow for the existence of groups when group selection is too weak for group-level adaptations to evolve; bird flocks are the groups for warning calls, even if warning calls do not evolve or evolve only to an intermediate frequency. In general, we must have a definition of groups that describes the process of group selection and this definition of the *process* cannot be based on the assumption that a highly efficient group adaptation must be the *product* of that process. In addition, virtually all group selection models concern the evolution of single traits and take no stand on how other traits have evolved. A model for the evolution of warning calls remains agnostic about resource conservation and fighting behavior. It is interesting to ask how organismic units evolve that are adaptive with respect to many traits, but this is only one topic within multilevel selection theory; it cannot be used to provide a definition of groups for the whole theory.

Finally, we take issue with **Barrett and Godfrey-Smith's** statement that "causal description of evolution in individualistic terms ends as soon as *interaction between individuals begins* (italics theirs)." Multilevel selection theory bears its name because it carefully partitions natural selection into within- and between-group components. Plenty of interactions among individuals take place in the context of within-group selection and are properly classified as "not group selection". Group selection is a *subset* of frequency-dependent selection, not just another label for all frequency-dependent interactions.

### Pluralism

So far we have tried to establish in this reply (and in UO in greater detail) that multilevel selection is a coherent theoretical framework built around the question of whether traits can evolve that benefit whole groups, despite being selectively neutral or disadvantageous within groups. This involves a kind of pluralism concerning evolutionary forces—there are different kinds of selection processes that can and do operate in nature, with different traits evolving in response to different mixes of these processes. This pluralism has nothing to do with conventionalism, according to which it is up to us whether we say that group selection is present or absent in a given trait's evolution. In this sense, our pluralism might be viewed as a version of *scientific realism*.

To understand the issues involving pluralism raised by **Barrett and Godfrey-Smith, Dennett, and Skyrms**, we must consider the rejection of group selection in favor of other theoretical frameworks that occurred in

the 1960's. One important reason for group selection's rejection was empirical—the claim was advanced that among-group selection is invariably weak compared to within-group selection. Williams and other critics were not pluralists or conventionalists; they thought that groups could be identified and levels of selection measured precisely, thus providing the basis for their negative verdict on group selection. Subsequent theories were self-consciously developed as alternatives to group selection, not as different ways of defining or viewing the group selection process. However, as we have already suggested, it is impossible to study social behavior without assuming the existence of groups. Kin selection theory, evolutionary game theory, and selfish gene theory all had to assume the existence of groups to correctly predict the outcome of evolution, and the groups these different perspectives identify have to be the same for any particular trait. In UO we show that group selection was rejected in name only. When we identify the groups that are conceptualized within these alternative theories, we discover that traits frequently evolve by benefiting whole groups, despite being selectively neutral or disadvantageous within groups, exactly as postulated by multilevel selection theory but without mentioning the dreaded G-word.

For the benefit of those who have not yet read UO, the following example shows how a model can include all the elements of multilevel selection while seeming to deny a role for group selection. Returning to our example of warning cries in birds, consider a flock with one caller and nine non-callers. Everyone has a low fitness in this flock because only one bird is looking out for predators; however, this bird has the lowest fitness of all. Let us say that the chance of surviving predators is 50% for the non-callers and 25% for the caller. A second flock of birds has nine callers and one non-caller. Everyone has a high fitness in this flock because so many members are looking out for predators; however, the shirking non-caller has the highest fitness of all. Let us say that its chance of surviving predators is 100% compared to 75% for the callers. When we compare the fitness of callers and non-callers within each group, we see that callers are the losers in both cases. However, the group with more callers fares better than the group with fewer callers. This is the classic group selection scenario that began with Darwin. Now for a subtle shift in perspective—rather than comparing fitnesses within and among groups, let's simply calculate the average fitness of callers and non-callers across the groups. One non-caller has a survival probability of 100% and nine have a survival probability of 50% for an average of 55%. One caller has a survival probability of 25% and nine have a survival probability of 75% for an average of 70%. The average caller is more fit than the average non-caller, so why not say that calling evolves by individual selection? As in a magician's trick, the concept of group selection seems to have vanished! However,

the disappearance is just an illusion. The need for multiple groups and variation among groups is absolutely essential for the calling behavior to evolve.

This, in a nutshell, is the perspective shift that makes the group selection controversy seem so “mind-twistingly elusive and slippery.” Readers should consult UO to see how averaging the fitnesses of lower-level units across higher-level units enables theories such as kin selection, game theory, and selfish gene theory to assume the existence of groups while avoiding the appearance of group selection.

Three questions can be asked of this averaging approach. First, does it provide a different answer to the basic question that group selection theory was designed to answer? We believe that it does not. After all, the basic question explicitly asks whether traits evolve by benefiting whole groups, despite being selectively neutral or disadvantageous within groups. There is no way to answer this question without comparing fitnesses within and among groups. The averaging approach can tell us whether a behavior will evolve, but not whether it evolves because of within- or between-group selection. It cannot provide an answer at all, let alone a different answer. We find no room for pluralism here. If a trait evolves on the strength of group selection, despite being selectively neutral or disadvantageous within groups, then averaging the fitness of individuals across groups does not change any of these facts.

Second, does the averaging approach permit one to construct a true causal description of evolution, different from but no less true than a description based on multilevel selection? For example, can the outcome of a group selection process also be explained by describing the effects that an individual or a gene’s traits have when that individual or gene is embedded in different contexts (i.e., groups)? This is the thesis advanced by Sterelny and Kitcher (1988) and in a forthcoming paper by Kerr and Godfrey-Smith. We find some potential for pluralism here, though once again identifying “multiple descriptions of the causal facts” does not mean that it is a matter of convention whether or not group selection occurs.

Third, is the averaging approach heuristically useful? Does it provide insights about evolution despite its inability to address the basic question that multilevel selection theory was designed to answer? We find much potential for pluralism here, because we are happy to acknowledge that there has been progress in evolutionary research during group selection’s dark age. Who can deny that game theory and kin selection theory have provided insights, even when they have been erroneously interpreted as alternatives to the theory of group selection? However, we also think that the revival of multilevel selection theory and the unification of evolutionary theories of social behavior will result in even more progress in the future.

This third kind of pluralism is philosophically mundane, as **Dennett and Barrett and Godfrey-Smith** acknowledge. There is much to discuss about the second kind of pluralism but it is even more important to distinguish the second kind from the first kind. It is remarkable how often philosophical discussion of group selection departs from the basic question that group selection theory was designed to answer. Sterelny and Kitcher's (1988) influential article entitled "Return of the Gene" said nothing about groups as vehicles of selection, which is how group selection must be discussed within selfish gene theory. Kerr and Godfrey Smith's forthcoming article, which will be accompanied by commentaries by us and others, also discusses multilevel selection theory and pluralism at great length without addressing the basic question. Thus, whatever Kerr and Godfrey-Smith mean by pluralism, it falls under the second question that we have outlined above. We conclude that the first and most basic question associated with group selection has a definitive and non-pluralistic answer: Traits can and do evolve by benefiting whole groups, despite being selectively neutral or disadvantageous within groups.

### Correlation

**Skyrms and Barrett and Godfrey-Smith** think that correlation provides the most general explanation of altruism. We see this as a restatement of the basic problem. If a trait is selectively advantageous within its own locality, then correlation among interactors is not required for its evolution. Thus, the reason that correlation is required for altruism is because it is selectively disadvantageous within its own locality — within groups. Correlation means that the frequency of altruists is not exactly the same in all locations—there is variation among groups. What has been said that was not already contained in the original formulation?

Correlation is only part of the multilevel selection process. In a multi-level selection model, correlation is synonymous with variation among groups. It is true that increasing variation among groups increases the strength of group selection, relative to selection within groups, but other factors can make group selection strong despite weak variation among groups or weak despite strong variation among groups (Wilson 1990, 1997). In the précis we stated that group selection explains the evolution of altruism but also entails that altruism must share the stage with many other traits that benefit the group without appearing overtly altruistic. Most individual-level adaptations evolve on the strength of random genetic variation among individuals. Is it not reasonable to expect many important group-level adaptations to evolve on the strength of random genetic variation among groups? This insight follows directly from multilevel selection theory but is obscured by

focusing exclusively on correlation, which erroneously gives the impression that group selection requires above-random variation among groups.

Another reason that an exclusive focus on correlation will miss important aspects of the selection process is that altruism can fail to evolve despite strong correlations. **Skyrms** states that fitness should be computed by using the appropriate conditional probabilities instead of the unconditional probabilities appropriate for random interactions. However, this statement assumes that the fitness of a given type, averaged across the total population, accurately predicts its average change in frequency in the total population; in physics, this is called a mean field approximation (Bar-Yam 2000). Interactions with neighbors in a fixed spatial structure, mentioned by **Skyrms** as an important area of future research, is a situation in which this assumption can easily fail. When dispersal is limited and offspring are always deposited close to parents, patchiness spontaneously develops, leading to high correlations that might seem to favor the evolution of altruism. However, group selection requires competition among groups, and the same limited dispersal that creates the patches also prevents the altruistic patches from exporting their productivity to other regions of the landscape. On the other hand, the advantages of selfishness are local, which allows selfish individuals to invade and take over altruistic patches. Altruism fails to evolve, despite high correlations and fitnesses based on conditional probabilities that would seem to favor the evolution of altruism. There is more to multilevel selection than correlation (Wilson et al. 1992, Mitteldorf and Wilson 2000).

### **The need for historical scholarship**

Some subjects in evolutionary biology have been examined in great detail by historians of science. The group selection controversy has not yet received this kind of attention, but we suspect that this will change. Schwarz (2000) has made a start by examining the relationship between George Price and W.D. Hamilton, including their private correspondence. It is crystal clear that Hamilton initially regarded his theory of inclusive fitness as an alternative to group selection and then changed his mind on the basis of the Price equation, as we mentioned earlier and recount in detail in UO. This event took place in the early 1970's, at the same time that Wilson (1975) independently developed his model of group selection. We are unsure what **Dennett** means by Hamilton's "alternative" to group selection or when he describes as a scandalous and unsubstantiated rumor the idea that Hamilton "went over" to the side of the group selectionists; Hamilton's publications and private correspondence with Price in addition to Wilson show that this is true. Price, Hamilton, and Wilson all converged on a view of multilevel selection that was theoretically robust, in contrast to previous models. The mystery is why so many evolutionary biologists found it difficult to change their minds about

group selection, even after the models and evidence were at hand. We suspect that the atmosphere surrounding group selection had grown so repressive that few had courage to question the dogma. Gadagkar (1997, p. 38-39) recalls that during the 1960's and 70's, he "... sat in many seminars where a question from a member of the audience was loudly dismissed by other members of the audience shouting 'but that's group selection!' even before the speaker had a chance to understand the question." Not only is it hard to think clearly under these circumstances; even the collective memory of the subject begins to fade as it is replaced by a melodrama, replete with heroes, villains, and great battles. Dawkins (1982, p. 6) writes that "we painfully struggled back, harassed by sniping from a Jesuitically sophisticated and dedicated neo-group-selectionist rearguard, until we finally regained Darwin's ground, the position that I am characterizing by the label 'the selfish organism'." Scholarship is required to retire this melodrama and restore an accurate collective memory of what multilevel selection is and has always been about. Darwin was the father of the idea of group selection. Hamilton began as a critic but then endorsed the hypothesis thirty years ago. Even Williams (1992, Williams and Nesse 1991) has endorsed the hypotheses for at least some traits. These icons of the anti-group selection brigade need to be read more carefully. Thirty years have gone by since the scene described by Gadagkar. The commentaries in this issue of *PPR* and the more general reaction that UO has elicited suggest that we are almost, but not quite, out of the woods.

### Odds and ends Concerning Part I

We organized the previous sections to address the most important issues raised by the commentators concerning Part I of UO. In this section we briefly address a residue of miscellaneous issues.

**Barrett and Godfrey-Smith** confuse kin recognition with the "green-beard effect," in which genetically unrelated altruists are supposed to recognize each other on the basis of a shared second trait. Dawkins (1976, p. 96) invented this imaginary example to show how altruism might in principle evolve among nonrelatives, if altruists have green beards and focus their donations on others with green beards. Maynard Smith (1964) did not distinguish kin selection from group selection on the basis of kin recognition, at least originally. He equated group selection with multigenerational groups, as we discuss on pp. 67-71 of UO, and kin selection with groups of relatives that hold together for the space of one-generation (or less).

**Dennett** says we commit a "sleight of hand" in our description of the Berkeley sex discrimination example. As he says, both the behavior of departments and the statistical treatment of women across the whole University are relevant. Indeed, as we point out, it is the former that explains the latter. The point is that fairness at the department level (men and women

applicants to a department have the same chance of gaining admission) is part of the explanation of why women get admitted less often than men in the University as a whole. The situation is precisely parallel to the fates of altruism and selfishness. Within groups, selfish individuals are fitter than altruists, but averaging across groups, the reverse can be true. We don't see that any sleight of hand is involved in our exposition; we're just explaining a well-known example of Simpson's paradox.

**Dennett** also notes that we use terms like "voluntary" and "forced" to describe possible interpretations of the schooling fish in the cartoon on p. 14 of UO. Fish on the periphery of the school are altruistic on the assumption that their position enhances the fitness of fish in the school's interior at a cost to those on the periphery. Whether the fish on the periphery are there because they are forced there, or do so voluntarily does not affect this point. However, it is relevant to understanding how the existence of fish with peripheral locations evolved. If fish go to the periphery voluntarily, it is pretty clear that group selection is involved. But what if they go there because they are forced to do so? In this case, we must shift our focus from the fish on the periphery to the fish who do the forcing. These fish are providing a public good and are themselves altruistic, as compared to the strong fish who stay at the school's interior but do not force other fish to the periphery. We discuss this point about primary and secondary behaviors on pp. 143-149 of UO.

**Dennett's** question about what it means to say that the fish at the periphery go there "voluntarily" or are "forced" to do so is easy to address without attributing mental states to these organisms. Here is an experiment that would help one decide whether fish on the periphery are "forced" to be there or assume that dangerous location "voluntarily": Place fish that swim on the periphery of the group into a new school of robot "fish" that will move to the periphery if the introduced fish go to the interior of the school. If the introduced fish go to the interior in this experiment, we use the term "forced" to describe their behavior in nature. If they go to the periphery, we use the term "voluntary." There is no need to speculate about the internal mechanism (mental or otherwise) that makes them do this. As **Dennett** says, "the absence of "force" can only be a matter of there not having been any costly move ... by the other fish that has the effect of inducing or enforcing that dangerous choice."

## **Part II: Psychological Altruism**

We thank **Jamieson** for his interesting comments on Part II of UO. Interest in the group selection controversy often seems to overshadow the equally important issues surrounding psychological altruism. We sense that he

broadly agrees with us but thinks we work too hard to argue against positions that can be dismissed more easily.

**Jamieson's** greatest reservations concern the use of belief/desire explanations of behavior. We noted in UO that the debate between psychological egoism and motivational pluralism (the view that people have both altruistic and egoistic ultimate desires) presupposes that people have desires and that these desires cause behavior. **Jamieson** points out that people ascribe beliefs and desires to others for reasons other than their wanting to explain and predict behavior. This is fine with us; as long as beliefs and desires are causes of behavior, the problem of egoism versus altruism can be addressed. Also, it doesn't follow from the fact that people have a variety of motives for ascribing beliefs and desires that "literal descriptivism is not much more plausible in folk psychology than it is in ethics." We want to emphasize that it wasn't a goal of UO to defend folk psychology. As long as scientific psychology postulates representational states that can be regarded as beliefs and desires (or as relevantly similar to those states), the problem can be posed. If not, not. As we say, if Churchland-style eliminativism is correct, there is no problem here to address. The same point applies to **Jamieson's** suggestion that human parental care might be explained by a behaviorist account that does not invoke beliefs and desires at all. We doubt that behaviorism has the resources to provide this kind of explanation. In any event, part 2 of UO begins with the assumption that mentalism is true; only in that setting does the debate between psychological egoism and motivational pluralism make sense.

**Jamieson** notes that we formulate our evolutionary argument for motivational pluralism by considering hedonistic egoism as the alternative hypothesis. **Jamieson** agrees with Kyle Stanford (2001) that it is a defect in our argument that we did not consider pluralistic (nonhedonistic) egoism in this context. The reason we focused on hedonism is that we think that hedonism is the version of egoism that is hardest to refute. Furthermore, we don't see how a nonhedonistic egoism would help the defender of egoism to avoid the argument we construct. Stanford and **Jamieson** need to explain why the version of egoism they are thinking about rescues egoism from our evolutionary argument.

**Jamieson** says that we are "scornful of appeals to plausibility." What we say is that agreement with "common sense" is no argument in favor of motivational pluralism. We deny that the "burden of proof" is on those who challenge common sense. The approach we tried to implement in UO is *to let the evidence speak*; if it can be shown that motivational pluralism is better supported by the evidence than egoism is, fine. And if it can be shown that the opposite conclusion is true, that also is fine. But if the evidence fails to discriminate between the hypotheses, then it is a mere prejudice to insist that one's pet theory is true and that the other theory is "implausible." **Jamieson**

says that psychological altruism is “the view to which most of us are pre-analytically committed” and that he rejects “the idea that ordinary ascriptions of altruism require ... scientific justifications.” For **Jamieson**, “such ascriptions are grounded in a different way than straightforwardly scientific claims.” We confess that we don’t understand what the ground-rules are here, if they are not scientific.

**Jamieson** says that “the tortured and ad hoc explanations offered by hedonistic egoism are more like those given by the Ptolemaic view than those provided by relativity theory or quantum mechanics.” What is the substantive claim behind this analogy? Ptolemaic astronomy is often described as less parsimonious than Copernicanism. Although we agree that parsimony can be epistemically relevant, and not just an aesthetic frill (Forster and Sober 1994), this doesn’t help one make an argument for motivational pluralism. The reason is that egoism (a monistic theory) is *prima facie* simpler than motivational pluralism. We discuss this issue in UO.

**Jamieson** describes the example of Kieran, who believes that two actions would be equally pleasurable, but still has values and preferences that lead her to choose one action over the other. Kieran, so described, is a counterexample to hedonistic egoism. Jamieson does not explain why this description should be taken at face value. The defender of hedonistic egoism will want to find a difference in Kieran’s assessment of the pleasures that each activity will provide. Why is this a mistake? **Jamieson** says he thinks that this would be implausible, but why is it implausible?

**Jamieson** criticizes hedonistic egoism on the ground that it “crudely conceptualizes people as simple, one-dimensional, decision makers, seeking to realize only one value (pleasure).” We don’t wish to defend hedonistic egoism, but the point needs to be made that any theory of decision, egoistic or otherwise, must render different values commensurable. This is why the concept of utility is needed in decision theory. We assume that **Jamieson** would not want to criticize motivational pluralism for assigning commensurable utility values to both the state of self and the state of others. Why, then, is hedonistic egoism subject to this criticism?

**Jamieson** concludes his commentary by pointing out that egoism is sometimes used ideologically; it is often used to defend a political vision of the future. If people are essentially selfish, isn’t it hopelessly utopian to try to advance the goals of cooperation and equality? We agree that egoism has been and continues to be used in this way. However, this is perfectly consistent with its also being a hypothesis that we should try to assess on scientific grounds. For those caught in the grip of an ideology, what is the best way to break its hold? More ideology, but of a different sort? Or a consideration of the evidence that bears on this question?

## Concluding Comment

We thank the commentators and organizers of this symposium for their efforts. Science made great strides in understanding the material world long before it addressed the subject of altruism, but we think that a scientific theory of altruism is emerging and we are pleased that it is getting the attention it deserves.

## REFERENCES

- Alcock, J. (1989). *Animal Behavior* (4th ed.). Sunderland, MA: Sinauer Ass.
- Bar-Yam, Y. (2000). Formalizing the Gene-Centered View of Evolution. *Advances in Complex Systems*, 2, 277-281.
- Bowles, S., & Gintis, H. (2002). Homo reciprocans. *Nature* 415, 125-128.
- Bull, J. J. (1994). Virulence. *Evolution*, 48, 1423-1437.
- Dawkins, R. (1976). *The Selfish gene* (1st ed.). Oxford: Oxford University Press.
- Dawkins, R. (1980). Good strategy or evolutionary stable strategy? In G. W. Barlow & J. Silverberg (Eds.), *Sociobiology: beyond nature/nurture?* (pp. 331-367). Boulder, CO: Westview Press.
- Darwin, C. (1871). *The descent of man and selection in relation to sex*. New York: Appleton.
- Fehr, E., & Gächter, S. (2001). Altruistic punishment in humans. *Nature* 415, 137-141.
- Forster, M. and Sober, E. (1994): How to Tell When Simpler, More Unified, or Less *Ad Hoc* Theories Will Provide More Accurate Predictions. *British Journal for the Philosophy of Science* 45: 1-36.
- Gadagkar, R. (1997). *Survival Strategies*. Cambridge, MA: Harvard University Press.
- Haldane, J. B. S. (1932). *The causes of evolution*. London: Longmans Green & Co.
- Hamilton, W. D. (1975). Innate social aptitudes in man, an approach from evolutionary genetics. In R. Fox (Ed.), *Biosocial anthropology* London: Malaby Press.
- Herre, E. A. (1999). Laws governing species interactions? Encouragement and Caution from Figs and their Associates. In L. Keller (Eds.), *Levels of Selection in Evolution* (pp. 209-237). Princeton: Princeton University Press.
- Kerr, B., & Godfrey-Smith, P. (forthcoming). On the relations between individualist and multi-level models of selection in structured populations. *Biology and Philosophy*.
- Maynard Smith, John (1964). Group Selection and Kin Selection. *Nature* 201: 1145-1146.

- Mitteldorf, J. J., & Wilson, D. S. (2000). Population viscosity and the evolution of altruism. *Journal of Theoretical Biology*, 204, 481-496.
- Price, G. R. (1970). Selection and Covariance. *Nature*, 277, 520-521.
- Schwartz, J. (2000). Death of an altruist. *Lingua franca*, 10(5), 51-61.
- Stanford, K. (2001): "Review of *Unto Others*." *Journal of Philosophy* 98: 43-47.
- Sterelny, K., & Kitcher, P. S. (1988). The Return of the Gene. *Journal of Philosophy*, 85, 339-360.
- Williams, G. C. (1966). *Adaptation and Natural Selection: a critique of some current evolutionary thought*. Princeton: Princeton University Press.
- Williams, G. C. (1992). *Natural selection: domains, levels and challenges*. Oxford: Oxford University Press.
- Williams, G. C., & Nesse, R. M. (1991). The dawn of Darwinian medicine. *Quarterly Review of Biology*, 66, 1-22.
- Williams, G. C., & Williams, D. C. (1957). Natural selection of individually harmful social adaptations among sibs with special reference to social insects. *Evolution*, 11, 32-39.
- Wilson, D. S. (1983). The group selection controversy: History and current status. *Annual Review of Ecology and Systematics* 14, 159-187.
- Wilson, D. S. (1990). Weak altruism, strong group selection. *Oikos*, 59, 135-140.
- Wilson, D. S. (1997). Altruism and organism: disentangling the themes of multilevel selection theory. *American Naturalist*, 150, S122-S134.
- Wilson, D. S., & Dugatkin, L. A. (1997). Group selection and assortative interactions. *American naturalist*, 149, 336-351.
- Wilson, D. S., Pollock, G., & Dugatkin, L. A. (1992). Can altruism evolve in purely viscous populations? *Evolutionary Ecology*, 6, 331-341.
- Wright, S. (1945). Tempo and mode in evolution: a critical review. *Evolution*, 26, 415-419.
- Wynne-Edwards, V. C. (1962). *Animal dispersion in relation to social behavior*. Edinburgh: Oliver & Boyd.