
Biology and Philosophy **17:** 539–550, 2002.

Response

Group Fitness and Multi-level Selection: Replies to Commentaries

PETER GODFREY-SMITH
Philosophy Department
Stanford University
Stanford, CA 94305-2155
USA

BENJAMIN KERR*
Department of Biological Sciences
Stanford University
Stanford, CA 94305
USA

We thank all four critics for their thoughtful comments on what is (we realize) a rather complex paper. For the most part, we will discuss the commentaries in the following order: Sober and Wilson, Maynard Smith, Dugatkin. But at some points we will depart from this sequence so we can make connections between the commentaries.

1. Sober and Wilson, Maynard Smith

Sober and Wilson's main argument is that there is not a particularly tight connection between the main theoretical *ideas* of multi-level selection (MLS) theory and the π/ϕ *parameterization* that we discuss in our paper. (In this reply we will omit all "*i*" subscripts from parameters.) According to Sober and Wilson, MLS theory is free to use different kinds of parameters, including the α/β parameters that our paper associated with individualism. Because we overstate the connection between a parameterization and a theoretical perspective, we are said to misconstrue Sober and Wilson's arguments about the role of MLS theory. We are also called on to confront what Sober and Wilson regard as "the basic question" in this area: "can traits evolve by bene-

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

fitting whole groups, despite being selectively neutral or disadvantageous within groups?"

Those issues will be the focus of much of this response. Let us first look closely at the relation between MLS theory and the way that mathematical models can be set up. According to Sober and Wilson, what is fundamental to both MLS theory, and also to a proper consideration of "the basic question," is the idea of a mixture or combination of selection processes operating at different levels. Group selection occurs when there is "variation in fitness among the groups in a metapopulation." Individual selection occurs when there is "variation in fitness among individuals within groups." So the idea of variation in fitness among groups, or *group-level advantage*, is fundamental to Sober and Wilson's outlook. They think this idea is essential to understanding the evolution of altruism. We argue, in contrast, that the usefulness of thinking about these phenomena in terms of group advantage varies. In the trait-group cases modeled in our paper, there is a formal equivalence between descriptions given using MLS theory and descriptions given using a certain kind of "individualist" framework. Sometimes MLS concepts are especially useful for thinking about these cases, but sometimes the individualist framework is better at picking out key distinctions; thus our defence of "gestalt-switching." Further, although we did not discuss these cases in our paper, altruistic traits can also evolve in situations where the idea of group-level advantage is either difficult or impossible to even *apply* to the situation. That is one of the points made by Maynard Smith, in his discussion of trees in a forest where there is a well-defined *neighbor* relationship but no well-defined *groups*. (We will discuss this case in detail below.)

Our α/β parameterization is a way of assigning a complete fitness structure to a model in which *only* individuals are treated as the bearers of fitness properties. The α/β parameterization pays *attention* to the role of groups, but it does not assign *fitness values*, or anything similar to fitness values, to groups. Rather than the bearers of fitness, groups are considered as part of the *context* of each individual. Groups *affect* fitnesses of individuals, but groups do not *have* fitnesses. So groups are not being ignored or denied an important role, but they are being given a particular *kind* of role. They are, again, being seen as the context in which individuals live and interact, for at least part of the life cycle. So we call this a "contextualist" way of treating the role of groups. This "contextualist" way of thinking is not essentially tied to our particular mathematical apparatus. It is a mode of thinking described by Dugatkin and Reeve (1994) and Sterelny (1996), in different terms.

The π/ϕ parameterization, on the other hand, explicitly assigns fitness values (or fitness-like values) to groups themselves. To use the terminology we prefer, the π/ϕ parameterization treats groups in a collective, as opposed

to contextual way. Groups are seen as *fitness-bearing collective entities,* not just as factors that affect individual fitnesses. Now, as we said above, for Sober and Wilson the concept of group-level advantage is essential for expressing MLS theory and for addressing their "basic question." If questions about group-level advantage are to be asked at all, then something like a π parameter is needed. The parameter used might not be exactly our π, but it has to be something close to it. After all, π is measuring or representing a biological property that is absolutely central to MLS theory – group-level fitness.

This relates to the issue of how altruism is defined. There are different definitions of altruism in the literature, with complicated relationships between them. And we don't find these different definitions being chosen randomly by authors. Rather, there is an association between how models are parameterized and how altruism is defined. This is because the parameterizations we discuss are not merely ways of attaching numbers to a system, they reflect different ways of thinking about the biological role of groups. Those who think about groups in a "collective" way tend to define altruism using a concept of group-level advantage, and tend to use a parameterization that is akin to our π/ϕ system. Those who think more individualistically, and treat groups in a contextual way, tend to use different parameters and different definitions of altruism (see also Kerr, Godfrey-Smith and Feldman forthcoming). The relation between parameterizations and definitions is not merely a matter of the amount of ink needed to write something down, as Sober and Wilson suggest several times. Even if ink is in plentiful supply, different parameterizations correspond to different ways of thinking about the roles of biological entities, especially groups.

Given this relationship between different parameterizations and different ways of thinking about groups, it is a very important fact that any system that can be modeled using π and ϕ parameters can also be modeled using α and β parameters. Moving to an individualist way of thinking, in cases of the kind we discuss, does not involve any loss of information or loss of representational power. Any system that can be modeled using π and ϕ – a parameter for group-level advantage and a parameter for change within groups – can be modeled without loss of information using α and β instead. That is, an equally informative model can be constructed that does *not* treat groups as fitness-bearing entities.

An α/β model might, of course, be *redescribed* in MLS terms. In cases with discrete groups, it is always possible to do a translation in either direction. As our paper says, the two parameterizations can be seen as ways to parcel the same information in different ways. In their commentary, Sober and Wilson also seem to treat individual mean fitness within a group as a proxy

for a group-level fitness. This is certainly possible, as individual within-group mean fitness is closely related to the π parameter. An individualist model does *implicitly* give a set of group fitnesses, via the rules given in Table 1 of our paper. But the concept of group-level advantage does not explicitly appear in an individualist model of the kind we describe and is not *needed* to describe evolutionary processes, even when altruism is being discussed. Sober and Wilson seem to think that even when multi-level fitnesses are not explicitly represented in a model, they in some sense correspond to what is *really going on* – these fitnesses pick out the crucial "processes" and "forces" in the situation. But Sober and Wilson have not shown anything about the differences between the individualist and multi-level perspectives that would justify this claim.

The general strategy in Sober and Wilson's commentary is to downplay the role of parameters. In our paper, we emphasize the connection between MLS theory and the assignment of fitness parameters to groups. Sober and Wilson say that this connection is not especially important; the advocate of MLS theory can use whatever parameters he or she finds convenient. If a parameterization that only ascribes fitnesses to individuals is useful in some particular case, the MLS advocate can happily use that parameterization while adding that this is entirely consistent with MLS theory. But Sober and Wilson are applying a double-standard here. In *Unto Others,* one of the most interesting parts of their defence of the MLS approach was their discussion of the success of tit-for-tat (TFT) in the iterated prisoner's dilemma (pp. 79–86). They want us to describe the success of TFT in terms of group selection acting on TFT pairs, not as a special kind of individual advantage enjoyed by TFT players. How do they mount this argument? In part, by giving us an alternative parameterization. They take the standard game-theoretical description of the iterated prisoner's dilemma, which is individualistic in its parameters, and they transform it into a multi-level description that includes a parameter representing group advantage. We are supposed to see the insights revealed by this new parameterization, and appreciate the power of MLS theory.[1]

Now compare this to their commentary on our paper. Our paper went to some length to argue that just as there are unexpected gains from moving to a multi-level parameterization in some cases, so there are also definite gains from moving to a certain kind of *individualist* parameterization in others. The α/β parameterization reflects a way of thinking about the relations between individuals and groups, and sometimes this contextualist way of thinking yields real insights. But when confronted with an argument from the usefulness of an individualist parameterization, Sober and Wilson adopt a very different attitude to the role of parameterizations. They now treat parameters

just as pieces of mathematical language, which are to be assessed in terms of convenience. This is plainly a double-standard.

Indeed, this example reveals a general equivocation in Sober and Wilson's treatment of MLS theory. Sometimes they insist that *only* the concepts used in MLS theory are adequate for explaining the evolution of altruism. The interaction between group-level advantage and within-group selection is *the* fundamental idea, and any attempt to approach the issue differently is mistaken. In *Unto Others* the language of causation was used to make this point (p. 33), and there are many references to group selection as a "force" (pp. 102–104). This is the strong version of Sober and Wilson's defence of MLS theory. But at other times Sober and Wilson handle the issue differently. In their comments here, they describe the MLS approach as a "heuristic." They also say that individualist descriptions that do not engage in inappropriate averaging of fitnesses are fine.[2] They distance themselves from any "vague" claim about the superior causal or explanatory resources of MLS theory. They emphasize their endorsement in *Unto Others* of "pluralism of perspectives." This is the more moderate form of their defence. The more moderate view is not far from our own, although we emphasize the positive usefulness of "gestalt-switching" between perspectives as well as a tolerance of different frameworks.

So might we consider the moderate position to be the official view of Sober and Wilson? We cannot. Along with the mild claim that MLS theory is a "heuristic" we also find an insistence on what Sober and Wilson call "the basic question" about group selection. This question is so central that we are called on to answer it directly. And the *formulation* of the basic question itself presupposes the multi-level framework; the question is expressed as one about the evolutionary role of group-level fitness differences. So clearly MLS theory is not meant to be a mere heuristic or optional perspective. (We will say more about the "basic question" shortly.)

We now move to a more technical argument against Sober and Wilson's comments on the role of parameters. They say that MLS theorists can add α/β parameters to their stock of concepts, and combine these with the idea of group advantage. That is, why not use α/β parameters *and* the π parameter? Why not have α and β in MLS theory as a way to treat the within-group force, as opposed to the between-group force? This is impossible, however. One sign of the problem is the fact that we seem to have a choice between using two parameters, α and β, or a single parameter, ϕ, to do what is apparently the same job. How can two parameters be compressed into one? They can't, in a case like this. If we try to set up a model in which we independently assign α, β and π parameters, the result is an over-determined system. The assignment of α and β parameters *imply* all the π values. Here is another

way to think of the situation. If α and β were used in an MLS description, their role would be to represent the within-group processes, not the between-group processes. Then the *absolute* productivities of the types of individuals would not be relevant, only the relations between them. But that is exactly what the ϕ parameter measures. The α and β parameters, on the other hand, contain too much information. They are *already* a complete specification of the fitness structure. Here is yet another way to put the point. The α and β parameters are designed to do exactly what should not be done according to MLS theory; they characterize a complete fitness structure without making any distinction between levels.

Interestingly, it would be possible to construct a model using only the parameters α and π, or β and ϕ, or any other two of the four parameters we have discussed. But anything other than α and β, or π and ϕ, would yield a strange looking model. This is because the π/ϕ parameterization is a natural representation of one of the two standard perspectives on the situation, and the α/β parameterization is a natural representation of the other. The MLS theory idea of comparing group-level fitness differences with within-group fitness differences is captured by π and ϕ. Any other way of capturing this idea will be at least close to this way.

We turn, at last, to what Sober and Wilson call "the basic question." This, again, is: "Can traits evolve by benefitting whole groups, despite being selectively neutral or disadvantageous within groups?" Sober and Wilson explicitly call on us to answer "yes." But the question admits of two construals. Roughly speaking, we answer yes on one construal and no on another.

First, we are not sure whether Sober and Wilson really mean to use "can" in their formulation. If the question is whether it is *possible*, in *some* cases, for traits to evolve via group-level advantage, then our answer is, roughly, yes. More precisely, there are indeed systems that are naturally modeled using a multi-level parameterization in which traits that give rise to higher values π values evolve despite a local disadvantage.

There are several ways in which this evolutionary outcome can fail to occur, of course. One is that altruism does not evolve. Another is that the system is not naturally or properly modeled with multi-level parameters like π and ϕ (see below). Interestingly, there is a third possibility. There can be a case where altruism evolves, the multi-level parameterization is available, but where it is questionable whether the Sober and Wilson description given above applies, because there is no real "benefit to whole groups." Consider a case where altruistic individuals randomly select *one* member of their group to receive a large donation of fitness. No one else in the group is helped by that altruist. This case will have the same evolutionary dynamics as the

case of equitable sharing of the donations of altruists. In a way, this is a case where there is "benefit to the whole group," but only in an *averaged* sense of "benefit." So if we construe the idea of group benefit in the basic question in a literal and narrow way, then this "lottery altruism" is another situation in which altruism evolves without contributing to group benefit.

We turn to a stronger construal of the basic question. Perhaps Sober and Wilson mean to ask something like this: "Is it the case that altruism *can* evolve, but *only* via group-level benefit?" If that is the question that Sober and Wilson regard as basic, then our answer to the question is no. We discussed an interesting borderline case of altruism without group-level benefit above. A more important class of cases are discussed by Maynard Smith, in his commentary. These are cases where a system has a spatial structure in which there are no well-defined groups but there is a well-defined *neighborhood* relation. Altruism can sometimes evolve via an exchange of benefits by neighbors. Maynard Smith's example is trees of the same species in a forest.

We did not discuss this important class of cases in our paper (though we will in a later paper). Maynard Smith is right to direct our attention to them. In these cases, an α/β style parameterization is easy to apply. Although we defined α and β in our paper for discrete groups only, the "contextual" way of thinking about individual fitness applies just as well when there are neighborhoods but no groups. So the α/β parameters can be used to model a case like this without any significant modification. On the other hand, it is very difficult, if not impossible, to apply a π/ϕ parameterization, or anything like it, to many cases of this kind. That has consequences for any attempt to describe such cases in the language of MLS theory.

Suppose we have a two-dimensional grid of squares, all occupied by individuals, so each individual has eight neighbors touching its own square at either an edge or a point. We have two types of individuals, ***A*** and ***B***, and the fitness of an individual depends on who its neighbors are. Both types benefit from having ***A*** neighbors rather than ***B*** neighbors, but ***B*** individuals have higher fitness in any given neighborhood-type than ***A*** individuals do. Let us look at how we might try to describe this case using the concepts of MLS theory. If we are to compare within-group selection to selection at the level of groups, we need to assign some kind of fitness value to groups as collectives. So we first need to work out where the groups begin and end. As we see it, there are three possibilities. One is to divide the grid into groups arbitrarily. Another is to say that each individual is in the center of a group composed of itself and its eight neighbors. A third is to say that whether or not there are groups (and hence whether or not MLS theory applies) depends on the particular case – perhaps on whether there are well-defined clumps of each type. Only the first two options would allow us to apply MLS theory in

a general way, and only the second option seems to be consistent with Sober and Wilson's general claims about how to locate groups.[3] That is, the second option is based on the idea that individuals who affect each other's fitness are in the same group. But this certainly yields some odd consequences. We now have as many groups as there are individuals. If you are in my "group" then I am also in your "group," but my group has members that yours does not have. What we have really done here is use the word "group" for something else, which is a given individual's set of neighbors.

Maybe that is OK; terminology per se is unimportant. The problem becomes worse when we try to assign group-level fitnesses. Here again we stress the role of parameters. It is one thing to talk about group-level advantage, but another thing to represent it precisely and measure it. So we need a parameter to represent group-level fitness. And now we find that the fitness of a group is no longer a function solely of the frequencies of ***A*** and ***B*** types within a group. If you are my neighbor, your fitness affects my group's fitness, as you are in my group. But your fitness depends on who *your* neighbors are, and these include individuals who are not members of my group. So the fitness of my group is affected by individuals who are not in my group, according to the neighborhood criterion for group-membership. Does that mean that all of my neighbor's neighbors should be placed in my group after all? But the argument will then extend to the next set of neighbors, and so on . . .

So this is a system in which the concept of group-level advantage is difficult or impossible to apply in a general way. We have discussed the case rather briefly here, and intend to follow up in another paper. The points that are important for present purposes are two: (i) systems of this kind are easily modeled using an α/β parameterization, with slightly adjusted definitions, but are very difficult to model in a general way with any parameterization using group-level fitnesses; and (ii) altruism can evolve in cases of this kind, via the exchange of benefits across neighbors. So, as Maynard Smith says, this is a case where an individualist approach to the evolution of altruism is superior. On the stronger construal, the answer to Sober and Wilson's "basic question" is no.

Sober and Wilson are concerned at several points to indicate connections between our analysis and earlier work of theirs (Wilson's "weak and strong altruism," which is related to our Class I/II distinction, Sober's discussion of the analogy between evolution and deliberation, and Wilson's use of graphical methods, for example). We do not deny these connections; all were explicitly acknowledged in our paper, and we find this work very valuable. But in each case we have tried to go a step further, and in doing so have come to different

conclusions from them about the status of multi-level and individualist modes of thinking.

2. Maynard Smith, Dugatkin

Maynard Smith accepts that a pluralist attitude to these issues is often reasonable, and "gestalt-switching" between perspectives is often positively useful. We agree with Maynard Smith's comment that gestalt-switching is particularly helpful when thinking about the "major transitions" in evolution (the origins of eukaryotes, multicellularity, etc.). As Maynard Smith says, there are themes concerning cooperation, subversion, policing and high-level benefit that recur constantly when thinking about these phenomena. Throughout his commentary, Maynard Smith is generally willing to see *cooperative* intereractions between individuals as the basis for recognizing a higher-level unit (see his discussions of cells in multicellular organisms, and bird pairs dividing their labor). He is reluctant to make the same move in the case of competitive interactions, and we will discuss this below.

Maynard Smith discusses a series of five cases in which collective and contextual strategies of analysis can be compared. We discussed his example with trees in a forest above; we agree that this poses a problem for the collective mode of analysis. Maynard Smith accepts that both collective and contextual analyses are available for the cases of cooperative behavior in worker bees and meerkats. The example that Dugatkin discusses in detail, in which fish within schools engage in predator inspection, seems similar to the meerkat case. Both Maynard Smith and Dugatkin accept pluralism here.

Maynard Smith sees his other two examples as trickier: ordinary "individual level" adaptations like halteres and hearts, and contests between individuals for resources. In the case of ordinary "individual level" adaptations, Maynard Smith says that we have a standard mode of explanation available, but it works by *assuming* the integrity and relevance of a particular collective entity, the individual multi-cellular organism. We think that D. S. Wilson has always been right to emphasize the analogy between social groups of organisms, and groups of cells making up an organism. As Wilson says, there has always been one kind of "group selection" that orthodoxy has been willing to recognize – selection on multicellular organisms as units. This analogy is discussed in *Unto Others*, and also in our paper (see also Sterelny 1996). In both cases, however, the contextual perspective, as well as the collective one, is in principle available.

Maynard Smith's final example is non-cooperative interactions between individual organisms, like animal contests. As Maynard Smith notes, the history of debate over these cases hangs over the present discussion. Julian

Huxley explained ritualized contests in terms of species selection. For Maynard Smith, this was a very important error. And Maynard Smith is critical of any attempt to analyze animal contests as a form of group selection in which the pair of interacting individuals is seen as a temporary group.

We agree that *insisting* that animal contests should be understood in terms of group selection is a mistake. But should we insist that these cases *not* be analyzed via the concept of group benefit, and a multi-level process? For Maynard Smith, the causal structure of the situation is not compatible with this kind of analysis. We are, we confess, still uncertain about what to make of the role of causation in this debate. But our suggestion is that even when the gestalt-switch looks initially awkward, it may often be a useful exercise. Think about the role of cooperation in the "major transitions" in evolution. Such transitions might often involve a shift from simple and direct competition, through modulated or restrained competition (like the spiders signaling), through various stages of cooperation until we get a tightly integrated whole at a higher level. Our suggestion is that looking at both sides might be useful *all the way through* the process.

All three commentaries allude to the history of the group selection debates. Dugatkin goes further, and makes a short-term projection into the future. He conjectures that our paper will not change the tone of the discussion or resolve the disagreements. His conjecture might be based partly on the fact that he tried, in his paper with Reeve, to suggest a similar pluralist or compatibilist solution back in 1994; and the debate still continues. So our own paper, longer and more complicated than theirs, will probably not help much either. Die-hard individualists (our term, not his) will "scream bloody murder," and see it as a retrograde attempt to resurrect ideas that were supposed to have had stakes put through their hearts long ago. Mad-dog MLS theorists (also our term) will see it as "yet another sellout," an attempt to downgrade the importance of group selection while coverty embracing it.

We are not sure whether Maynard Smith counts as a die-hard individualist. Certainly he might be considered a candidate for this role, by reputation. But his commentary shows a good deal of receptiveness to both gestalt-switching pluralism, and the importance of sometimes treating groups in a collective way. This is important in the major transitions, cooperative social groups like the meerkats, and in cooperative interactions more generally. Maynard Smith is even supportive of the conjecture that sex is a species-level adaptation. On the multi-level side, Sober and Wilson's commentary is less positive, though inclined to treat many of our points as old hat, rather than as a sellout. Sober and Wilson do not call our π/ϕ parameterization a "mathematical smokescreen," as Dawkins once called ML-revivalist modeling, but they do

seem to think it lays distracting mathematical smoke (or ink) over what they regard as the "basic question."

In closing, we are grateful to *Biology and Philosophy* for collecting such a terrific group to help us work through these issues. In our minds, some big issues remain unresolved (for example, the role of concepts like "cause" and "force" in this area). Although, as Dugatkin suggests, peace may not immediately follow, dialogue of this kind will surely help to elucidate both the real differences between positions and the common ground.

Acknowledgment

We thank Michael Weisberg for comments on a draft of this reply.

Notes

[1] The issue of parameterization emerges again in *Unto Others*, albeit tacitly, when Sober and Wilson argue for a powerful role for the Price equation:

> If Hamilton had presented his theory in the form of the Price equation in 1963, evolutionary biologists would have been forced to conclude that group selection is a significant evolutionary force that partially justifies the interpretation of groups as organismic units. (p. 100)

As we mentioned in note 9 on our original paper, Price employed (what we call) a multi-level parameterization. That is, Price considered groups as collectives which are explictly assigned fitness-like values (Price 1972; see also our companion paper in this volume). So it is not surprising that his statistical analysis is friendly towards MLS theory.

[2] ESS analysis is endorsed as a useful form of analysis in the final paragraph of Sober and Wilson's commentary. Here Sober and Wilson do *not* argue that ESS concepts are naturally used within their multi-level approach; ESS analysis, unlike the "individualism" described in our paper, is regarded by them as genuinely different from MLS theory. However, the structure of ESS theory is explicitly contextual, in exactly the sense of our paper. Note that the ESS condition $w(\boldsymbol{A},\boldsymbol{A}) > w(\boldsymbol{B},\boldsymbol{A})$ they discuss is identical to the condition $\alpha_2 > \beta_1$ (when $n = 2$), a Class II relationship. ESS analysis is, in effect, a simplified version or a useful fragment of the individualist dynamical model we discuss in our paper.

[3] Sober and Wilson discuss spatially continuous populations in *Unto Others* (pp. 94–96) and say the following: "it is still true that altruism is locally disadvantageous and requires the differential productivity of patches to evolve." It is unclear whether a "patch" refers to a focal individual and its neighbors or a relatively uniform "clump" of larger size (though their comments about the "fuzzy" and "amorphous" nature of such patches makes us think the latter interpretation applies). However, it is clear that they think the MLS approach is both suitable and helpful for understanding the evolution of altruism in such populations. We contend that the explicit assignment of patch-level (or group-level) productivity is at best extremely difficult in many of these cases.

References

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

Kerr, B., Godfrey-Smith, P. and Feldman, M.: (forthcoming), 'What is Altruism?'.

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

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

Sterelny, K.: 1996, 'The Return of the Group', *Philosophy of Science* **63**, 562–84.