A CRITICAL REVIEW OF PHILOSOPHICAL WORK ON THE UNITS OF SELECTION PROBLEM*

ELLIOTT SOBER†‡

*Philosophy Department
University of Wisconsin

DAVID SLOAN WILSON

*Biology Department
Binghamton University

The evolutionary problem of the units of selection has elicited a good deal of conceptual work from philosophers. We review this work to determine where the issues now stand.

1. Introduction. Philosophers have produced a large literature aimed at clarifying what a unit of selection is. Rather than launch immediately into the technical details of that literature, we begin with an informal description of what the problem of the units of selection is about. As the positivists used to say, the explicandum must be clarified before the adequacy of the explicans can be evaluated. Section 2 explains the issue. Sections 3–7 criticize some of the main ideas that have been introduced. Section 8 presents our own take on the problem. Section 9 extracts a consequence.

2. Back to Basics. The problem of the units of selection has engaged the attention of evolutionists ever since Darwin. It concerns whether traits evolve because they benefit individual organisms1 or because they are good for the group in which they occur. More recently, a third alternative has been proposed, which holds that traits evolve because they benefit the genes that code for them (Williams 1966, Dawkins 1976).

The choice that Darwin considered—between the group and the or-

---

*Received September 1993; revised July 1994.
†We thank the National Science Foundation for financial support (NSF Grant SBE92-12294). We also are grateful to Robert Brandon, Peter Godfrey-Smith, David Hull, Richard Lewontin, and the anonymous referee of this journal for useful suggestions.
‡Send reprint requests to Elliott Sober, Philosophy Department, University of Wisconsin, Madison, WI 53706, USA.
1As a terminological convenience, we use “individual” and “organism” interchangeably. This does not prejudice the substantive claims that species are individuals (sensu Hull 1988) or that groups are sometimes organisms (sensu Wilson and Sober 1989, 1994a).
ganism as units of selection—was important because of the issue of evolutionary altruism. An altruistic trait reduces the fitness of organisms that possess it while benefitting the group in which it occurs. Altruistic traits are bad for the organism but good for the group. If the organism is the exclusive unit of selection, then altruism cannot evolve. However, if the group is sometimes a unit of selection, altruism becomes an evolutionary possibility.

Two consequences of this standard pairing of altruism with the group as unit of selection and selfishness with the individual as unit of selection (Wilson 1990) are worth noting. First, altruism and selfishness are defined by the fitness effects of a behavior; they have nothing essentially to do with psychological motives. Second, altruism is not the same as helping. Parental care is a type of helping, but if parents who care for their offspring are fitter than parents who do not, then parental care is not an instance of altruism.2

In order to fix ideas, it will be useful to apply the contrast between the group and the organism as units of selection to a pair of examples. Our interest here is not in getting the biological details right, but in helping the reader to see the relevant conceptual contrast. Consider, first, why zebras run fast rather than slowly. The answer is that zebras who ran fast were more successful in surviving to reproductive age than were zebras who ran slowly. The trait of running fast evolved because it benefitted the organisms who possessed it. Compare this with the barbed stinger of the honeybee. When a honeybee stings an intruder to the nest, the bee disembowels itself. The barb did not evolve because it helped bees who had the barb. On the contrary, barbs evolved because they helped the group, and in spite of the fact that they harmed the organisms possessing them. Nests made of individuals with barbed stingers did better than nests made of individuals without barbs.3

If the biological details are as stated in these two examples, we should conclude that the individual organism is a unit of selection in the evolution of running speed in zebras, whereas the group is a unit of selection in the evolution of barbed stingers in honeybees. Generalizing from these two examples, we obtain the following definitions:

2This is most obvious when one considers species with uniparental reproduction. Also, the present point does not deny the possibility of parent/offspring conflicts of interest (Trivers 1972; Haig 1993).
3Readers who think that barbed stingers evolved by kin selection and that kin selection is not a kind of group selection are asked to grant this example for illustrative purposes only. We argue that kin selection is a type of group selection (in which the groups are composed of relatives) in Wilson and Sober (1989, 1994a). This also is the position taken by Seeley (1989); the title of his paper is instructive: “The Honey Bee Colony as a Super-organism”.
The organism was a unit of selection in the evolution of trait $T$ iff one of the factors that influenced $T$’s evolution was that $T$ conferred a benefit on organisms.

The group was a unit of selection in the evolution of trait $T$ iff one of the factors that influenced $T$’s evolution was that $T$ conferred a benefit on groups.

These two special cases generalize to yield the following formulation:

Objects at level $X$ were units of selection in the evolution of trait $T$ iff one of the factors that influenced $T$’s evolution was that $T$ conferred a benefit on objects at level $X$.

Although the first two definitions do not describe what it means for the gene to be a unit of selection, the third one does: If a trait evolved because it benefitted the gene that coded for it, then the gene was a unit of selection.4

We note two consequences of this proposal. First, different traits may evolve for different reasons; for example, the group may be a unit of selection for one trait but not for another. Second, the same trait may evolve for several reasons—several units of selection may be associated with the evolution of a particular trait. Although a monolithic solution to the units of selection problem is possible (e.g., “the gene is the one and only unit of selection for all traits”), such an approach must be argued for explicitly; it is not dictated merely by the problem’s formulation.

We believe that this simple schema is helpful as a point of departure for understanding the units of selection problem. Nonetheless, a number of problems of clarification remain, which we will address in due course.

3. Replicators and Interactors. Hull (1980, 1981, 1988) has argued that the distinction between interactor and replicator is central to understanding the debate over the units of selection. Hull’s ideas generalize themes explored by Dawkins (1976, 1982). Dawkins distinguished replicators and vehicles. Genes are examples of the former and organisms are examples of the latter. Hull substituted the term “interactor” for Dawkins’s “vehicle” because Hull took Dawkins’s terminology to be committed to the selfish gene point of view according to which the gene, not the organism or the group, is the one and only unit of selection. Hull wanted to formulate a more general framework than Dawkins had enunciated, one in which various positions concerning the units of selection problem could be stated and clarified. Other authors (e.g., Lloyd 1988,

4It may seem to follow from this that the gene is always a unit of selection. This will be discussed in section 7.
Brandon 1990) have endorsed Hull’s suggestions and have added more technical proposals of their own.

Hull defines a replicator as “an entity that passes on its structure directly in replication” and an interactor as “an entity that directly interacts as a cohesive whole with its environment in such a way that replication is differential” (1980, 318).

The most valuable part of Hull’s (and Dawkins’s) distinction is that it separates the issue of heredity from the issue of which causal processes underlie differential reproduction. That genes are the units of heredity has never been at issue in the units of selection problem. If an altruistic phenotypic trait evolves by a process of group selection, the genes coding for that trait also must evolve. The idea that genes are units of heredity—that they are replicators—is common ground.

Even though the replicator concept is not central to the units of selection problem, it merits philosophical scrutiny. If a replicator is defined as an object that passes on its structure directly in replication, what does “direct” mean, and what is replication? The process by which genes replicate is intricate. In what sense do parental genes create offspring genes directly?5

Dawkins holds that sexual organisms are not replicators, but genes are. In what sense do organisms fail to make copies of themselves, whereas genes succeed? Sexual organisms often exhibit less than perfect copying fidelity, although for canalsized traits, fidelity is often very high. Human parents have one heart and their children usually do too. In any event, if organismic reproduction involves imperfect fidelity, why does this mean that organisms are not replicators at all? Why not say, instead, that their replication is imperfect?

Dawkins (1976, 1982) stipulates that replicators obey Weismannian, rather than Lamarckian, principles—they cannot mediate the inheritance of acquired characteristics. This prohibition is illustrated in figure 3.1. When a mother giraffe lengthens her neck by stretching, this does not induce a mutation in the genes passed along to her offspring that allows them to have long necks without needing to stretch. Phenotypic traits

5The question of how to interpret the idea of directness becomes even more pressing when one considers the suggestion (Hull 1988, Dawkins 1982, and Williams 1992) that asexual organisms, populations, and/or species can be replicators.
acquired in development do not alter the genes passed along in reproduction.

We have no quarrel with this routine rejection of Lamarckian ideas, although we emphasize that it is an empirical question whether Weismannism is always correct. However, we do not see why the concept of replication should be burdened with the stipulation that replicators are Weismannian. If genes occasionally violated Weismannism, would this mean that genes are not replicators?

A second question concerns the relationship of the concept of replicator to the notion of heredity. In biology, heredity is measured by the concept of heritability. When offspring phenotypically resemble their parents, this can be due to their sharing genes or living in similar environments (or both). A phenotypic trait has nonzero (narrow) heritability when parent/offspring resemblance is attributable, at least in part, to shared genes.

As such, heritability is a property of the phenotypic traits of organisms. The same concept applies straightforwardly to groups of organisms. If groups of organisms bud off daughter colonies, daughter colonies may resemble parental populations because of genetic similarities. Groups have heredity in the same sense that individual organisms do. (Maynard Smith 1987 and Ridley 1993 disagree; we pursue the point in Wilson and Sober 1989, 1994b.)

Heritability is essential for natural selection to cause evolution. If running speed in zebras is to evolve by individual selection, offspring organisms must resemble their parents. And if barbed stingers are to evolve in honeybees by group selection, daughter colonies must resemble their parents. Deciding “what the replicators are” is not important here, though ensuring that the traits are heritable is of the essence.\(^6\)

If organisms and groups both can possess heritable characters, what is so special about genes? Genes are, by definition, the objects that give the phenotypes of these higher level objects their (narrow) heritability. Nonetheless, reproduction and parent/offspring resemblance are hardly unique features of genes. Although being a replicator is not the same as being heritable, we see no harm in defining “the unit of heredity” as the gene.

A separate issue concerning the replicator concept is worth mentioning. Are the pages fed into a copying machine “replicators”? To be sure, copies are made of them. But do they make copies of themselves? Arguably, the answer is no. The pages are replicated, but they are not replicators. One implication of the term “replicator” is that replicators control their own destiny. They actively make copies of themselves; they are not passive entities of which copies are made. The idea that genes are replicators

---

\(^6\)This is one reason why Darwin was able to develop so many insights about natural selection even though his picture of the mechanism of heredity was completely erroneous.
may exaggerate the degree of encapsulation that the replication process possesses (Oyama 1985, Lewontin 1992).

We now turn to the interactor concept. In section 2, we introduced running speed in zebras and barbed stingers in honeybees as working examples of the organism and the group as units of selection. Does the concept of an interactor capture the requisite distinction? Let us see.

These two examples differ in a way that must now be made explicit. In our hypothetical example about running speed, we imagine that the individual organism, but not the group, is a unit of selection. However, in the case of the barbed stinger, we must recognize two units—the trait’s evolution is influenced both by the fact that barbed stingers are good for the group and by the fact that a barbless stinger is good for the organism. The group and the organism are units of selection in this instance.

If Hull’s proposal is to reflect these ideas, then it must be true, in the first case, that individual zebras, but not zebra herds, directly interact as cohesive wholes with their environment in such a way that replication is differential. In contrast, it must be true, in the second example, that both individual honeybees and the hives to which they belong directly interact as cohesive wholes with their environment in such a way that replication is differential.

Knowing how to judge an interaction’s “directness” is difficult, however. Presumably, zebras interact directly with the lions that kill them, just as bees interact directly with the bears they sting. But what if one zebra herd goes extinct because all its members are slow while another survives because all its members are fast? If this is not a case of groups interacting directly as a cohesive whole with their environments, how does this differ from one bee hive’s going extinct because it does not contain individuals with barbed stingers while another survives because its members have barbed stingers? For the group to be a unit of selection, more is required than the fact that some groups do better than others. How the idea of “direct interaction as a cohesive whole” supplies that further ingredient remains unclear.

4. The Analysis of Variance. Wimsatt (1980) proposed the following definition:

A unit of selection is any entity for which there is heritable context-independent variance in fitness among entities at that level which does not appear as heritable context-dependent variance in fitness (and, thus, for which the variance in fitness is context-dependent) at any lower level of organization. (P. 236)

How should the idea of “context-independence” be understood? Wimsatt explains that the idea of additivity, which has a clear meaning in the
statistical method known as the analysis of variance, is a special case of context-independence. Since we do not fully understand the wider meaning of “context-independence”, we focus on additivity.

Consider the fitness relationships that obtain between two loci, each of which has two alleles. Each organism is either AA, Aa, or aa at one locus and BB, Bb, or bb at the other. The fitness of each two-locus genotype may be represented as follows:

<table>
<thead>
<tr>
<th>B-locus</th>
<th>BB</th>
<th>Bb</th>
<th>bb</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>$w_{11}$</td>
<td>$w_{12}$</td>
<td>$w_{13}$</td>
</tr>
<tr>
<td>Aa</td>
<td>$w_{21}$</td>
<td>$w_{22}$</td>
<td>$w_{23}$</td>
</tr>
<tr>
<td>aa</td>
<td>$w_{31}$</td>
<td>$w_{32}$</td>
<td>$w_{33}$</td>
</tr>
</tbody>
</table>

If $w_{ij}$ is precisely halfway between $w_{ij}$ and $w_{ij}$, and $w_{ij}$ is precisely halfway between $w_{ij}$ and $w_{ij}$ ($i, j = 1, 2, 3$), then fitness relationships are additive and Wimsatt’s criterion judges the single gene to be the unit of selection.

Wimsatt’s criterion entails that the single gene is not the unit of selection in at least two circumstances. First, heterozygotes may fail to be precisely intermediate, though the relationships that obtain within one locus do not depend on what is true at the other. A hypothetical example is provided by the following table of viability fitnesses:

<table>
<thead>
<tr>
<th>B-locus</th>
<th>BB</th>
<th>Bb</th>
<th>bb</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>0.8</td>
<td>0.7</td>
<td>0.6</td>
</tr>
<tr>
<td>Aa</td>
<td>0.7</td>
<td>0.6</td>
<td>0.5</td>
</tr>
<tr>
<td>aa</td>
<td>0.3</td>
<td>0.2</td>
<td>0.1</td>
</tr>
</tbody>
</table>

Because the A-locus exhibits dominance in fitness in this case, Wimsatt’s criterion entails that the unit of selection is not the single gene, but the single locus genotype.

The second pathway by which the single gene may fail to be the unit of selection, according to Wimsatt’s criterion, involves epistasis in fitness. This occurs when the fitness relationships among the genotypes at one locus depend on what is true at the other, as is illustrated by the following hypothetical data set:

<table>
<thead>
<tr>
<th>B-locus</th>
<th>BB</th>
<th>Bb</th>
<th>bb</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>0.1</td>
<td>0.2</td>
<td>0.3</td>
</tr>
<tr>
<td>Aa</td>
<td>0.1</td>
<td>0.3</td>
<td>0.2</td>
</tr>
<tr>
<td>aa</td>
<td>0.2</td>
<td>0.1</td>
<td>0.3</td>
</tr>
</tbody>
</table>
Notice that the fitness ordering of the $B$-genotypes depends on what is true at the $A$-locus. Wimsatt’s criterion would conclude that the unit of selection in this case is the two-locus genotype, not the single-locus genotype, and not the single gene.\footnote{Wimsatt’s (1980, 1981) criterion was motivated by the last chapter in Lewontin (1974), which was called “The Genome as the Unit of Selection.”}

Wimsatt presented his criterion as a criticism of an argument that Williams (1966) advanced and Dawkins (1976) repeated, which aimed to show that the “meiotically dissociated gene” is the unit of selection. Here is Williams’s statement of the argument:

Obviously it is unrealistic to believe that a gene actually exists in its own world with no complications other than abstract selection coefficients and mutation rates. The unity of the genotype and the functional subordination of the individual genes to each other and to their surroundings would seem, at first sight, to invalidate the one-locus model of natural selection. Actually, these considerations do not bear on the basic postulates of the theory. No matter how functionally dependent a gene may be, and no matter how complicated its interactions with other genes and environmental factors, it must always be true that a given gene substitution will have an arithmetic mean effect on fitness in any population. One allele can always be regarded as having a certain selection coefficient relative to another at the same locus at any given point in time. Such coefficients are numbers that can be treated algebraically, and conclusions inferred from one locus can be iterated over all loci. Adaptation can thus be attributed to the effect of selection acting independently at each locus. \citep[1966, 56–57]{williams1966}  

Wimsatt contends that although attending to the frequencies and fitness values of single genes may be useful as a “bookkeeping” device, this point is irrelevant to whether the single gene is the unit of selection. The appropriate criterion, Wimsatt maintains, is the additivity criterion we have just discussed, which leads to quite different conclusions.

We agree with Wimsatt’s criticism of Williams’s argument, but disagree with the additivity criterion that Wimsatt proposed. Before explaining the disagreement, we will elaborate on Wimsatt’s important critique.

If evolution is defined as change in gene frequency, then evolution by natural selection entails that genes will differ in fitness, \textit{regardless of what the unit of selection is}. If group selection causes an altruistic gene to evolve, that gene will be fitter than the selfish allele it displaces. This means that the units of selection problem is not settled by the mere fact that the different alleles in a population have fitness values that can be “treated algebraically” \citep{sober1982, sober1984, sober1993; see
also Godfrey-Smith and Lewontin 1993 on the irrelevance of a model’s dimensionality to the units of selection question).

Let us now apply Wimsatt’s additivity criterion to the examples of organismic and group adaptations introduced in section 2. The first point is that the issue of additivity plays no role in explaining why running speed in zebras is an adaptation that evolved for the good of the individual organism. The genes that influence running speed may or may not exhibit dominance or epistasis. These questions are relevant to how fast running speed will evolve (as we learn from Fisher’s 1958 fundamental theorem). But they simply cut no ice with respect to the problem of whether the organism is the unit of selection in this instance.

We believe that the same conclusion should be drawn when the additivity criterion is applied to the problem of defining what it means for the group to be a unit of selection. To explain why, we need to say a little more about the concepts of altruism and selfishness. Figure 4.1 represents two fundamental facts about these evolutionary concepts. No matter what mix of altruism and selfishness is found in a group, selfish individuals are fitter on average than altruists. Second, increasing the frequency of altruism found in a group raises the fitnesses of altruists and selfish individuals alike. If we define the fitness of a group as the average
fitness ($\bar{w}$) of the individuals in the group, then this second point means that groups in which altruism is common are fitter than groups in which altruism is rare.

Figure 4.1 is a standard representation of the fitness relationships of evolutionary altruism and selfishness. When an ensemble of populations, each containing its own mix of altruistic and selfish individuals, satisfies certain further conditions, altruism can increase in frequency by the process of group selection. This suffices for the group to be a unit of selection.

Note that the fitness functions depicted in the figure are *straight lines*. The fitness of the group is here an additive function of the proportion of altruists it contains. However, this additive relation does not prevent the group from being a unit of selection. Of course, it is easy to model the nonadditive case. For example, just bend the fitness functions in figure 4.1 so that groups benefit from additional altruists according to a rule of diminishing returns. We then have an analog of dominance in fitness, but this makes no difference as to whether the group is a unit of selection.

Sober (1984) and Lloyd (1988) agreed with Wimsatt that absence of dominance and epistasis is criterial for the gene to be a unit of selection. Lloyd (1988) and Mayr (1990) also used additivity as a criterion for the group to be a unit of selection, whereas Sober (1984) resisted this conclusion. The schizophrenia implicit in Sober’s treatment of the issue was pointed out by D. S. Wilson in conversation and was independently identified by Walton (1991).

We suggest that additivity is wrong through and through. If running speed in zebras evolved because it benefitted individual organisms, this says nothing about the details of how genotypes code for phenotypes. Similarly, group selection does not require that group phenotypes be “emergent”, if emergence entails nonadditivity. The group as unit of selection embodies a kind of holism (Sober 1981, Wilson 1988, Wilson and Sober 1994a), but it is a holism that does not demand emergentism.

5. The Pattern of Variation in Fitness. Additivity describes a relationship that can obtain among fitness values. Even if the additivity criterion misdescribes what a unit of selection is, it is worth asking whether some other relationship defined on the observed variation in fitness can be used to characterize what a unit of selection is.

We believe that an argument presented in Sober (1984) shows that the answer is *no*. Suppose we investigate a set of populations, each internally

---

homogeneous for height. In the first population, all the organisms are one unit tall. In the second, all are two units tall, and so on. When we measure the fitnesses of these individuals, we find that height is perfectly correlated with fitness. In this case, there is no within-group variance in fitness; all the variation is between groups.

If pattern of variation somehow determined what the unit of selection is, the above information would settle definitively what the unit of selection is in this case. But it does not; two different hypotheses are consistent with the information given. The first proposes that there is individual selection for being tall, in which case the individual is the unit of selection. The second says there is group selection favoring groups with higher average height, which would mean that the group, not the individual, is the unit of selection.

Although pattern of actual variation in fitness does not determine what the unit of selection is in the example of the tall and the short, an experiment that would provide useful evidence is not hard to describe. Suppose we create some heterogeneous groups and then measure the fitnesses of the organisms in them. If tall individuals are equally fit, regardless of the kind of group they inhabit, this is evidence that selection is at the level of individuals. And if tall and short individuals in the same group have the same fitness regardless of their individual phenotypes, this favors the hypothesis of group selection.

The conclusion that Sober (ibid.) drew about this example—that the actual pattern of variation in fitness does not define what the units of selection are—was challenged by Lloyd (1988) and by Griesmer and Wade (1988). They argued that biologists use background information that allows them to use the observed variation in fitness to infer the units of selection in a given case. To some extent, Sober and his critics were talking past each other. Sober argued that facts about within- and among-group variation, by themselves, do not uniquely determine what the units are. The critics argued that those facts, plus other assumptions, settle the matter. Obviously, these two assertions are compatible.

Even if the actual pattern of variation in fitness does not determine the units of selection, a related criterion might be considered. The new idea is that group selection occurs precisely when the fitness of organisms depends on the kind of group they inhabit. Sober (1984) argued that this criterion is too permissive. In many cases of individual selection, individual fitnesses depend on group composition. For example, suppose that in the evolution of traits A and B, the advantage goes to the common trait; within each population in which the traits are exemplified, individual selection proceeds according to this frequency dependent rule. Now imagine two populations; in the first A is common while in the second A
is rare. These two populations will evolve in different directions. But this is not an instance of group selection.

Even if this critique is correct, we must recognize that the pattern of variation in fitness exhibited within and among populations is important in several ways. First, pattern of variation may be evidence for the existence of different sorts of selection processes. Second, the pattern of variation in fitness helps predict how the system will evolve. And finally, pattern of variation does play a defining role in a limiting case. Selection at a given level requires variation in fitness at that level. If groups do not vary in fitness, the group cannot be a unit of selection. If the organisms under study do not vary in fitness, the organism cannot be a unit of selection.

Group selection requires more than groups varying in fitness. And it is not enough that they vary in fitness and differ in their rates of extinction and colonization. Rather, what is required is that this pattern of variation obtains because of their different traits.

6. Screening-Off. Brandon (1984, 1990) has argued that the statistical concept of screening-off can be used to clarify the concept of a unit of selection. \( Y \) is said to screen off \( X \) from \( Z \) precisely when \( P(Z|X\&Y) = P(Z|Y) \neq P(Z|X) \). When this relation obtains, \( Y \) and \( X \) are related asymmetrically to the task of predicting \( Z \); if you know \( Y \), knowing in addition that \( X \) is true would not change your prediction. According to Brandon,

\[
\text{[s]election occurs at a given level (within a common selective environment) if and only if (1) there is differential reproduction among the entities at that level; and (2) the “phenotypes” of the entities at that level screen-off properties of entities at every other level from reproductive values at the given level. (1990, 88)}
\]

Brandon applies this criterion in two contexts. First, he argues that it explains why selection standardly acts on an organism’s phenotype, not on the genes the organism contains. Second, he argues that it elucidates what group selection is, and how it differs from selection at the level of the individual organism. We consider these in turn.

Mayr (1963, 184) and Gould (1980, 90) emphasize that selection acts “directly” on the organism’s phenotype, and only “indirectly” on its genes. Gould takes this to undermine the selfish gene point of view—the idea

---

9 Brandon uses “level of selection” to talk about what we call the units of selection problem. He reserves the term “unit” for another use. We use “level” and “unit” interchangeably.

10 A better formulation of Reichenbach’s (1956) idea would treat \( X, Y, \) and \( Z \) as variables that come in states. Let “\( X = a \)” mean that \( X \) is in state \( a \). Then \( Y \) screens off \( Z \) from \( X \) precisely when, for all \( i, j, k, P(Z = i \mid Y = j \& X = k) = P(Z = i \mid Y = j) \neq P(Z = i \mid X = k) \).
that the gene is the one and only unit of selection. Brandon suggests that
the Mayr/Gould point about the directness of selection can be captured
via the notion of screening-off, and that the criterion stated above explains
why selection typically acts at the level of organismic phenotypes, not at
the level of the gene.

Mayr and Gould’s causal claim is illustrated in figure 6.1. Even though
this diagram fails to represent the causal role of the environment, the
point is that the organism’s phenotype is a more proximal cause and its
genome is a more distal cause of the organism’s survival and reproductive
success.

In many causal chains, the proximal cause screens off the distal cause
from the effect. Is this true in the case at hand? Often it is. If a zebra’s
fitness is determined by its speed, then fixing a zebra’s running speed
allows (probabilistic) prediction of its survival and reproductive success;
adding information about the genes that endow the zebra with the running
speed it has will not alter the prediction.

The simple case of genetic dominance, however, provides an excep-
tion. Suppose that individuals with the AA and Aa genotypes are
phenotypically indistinguishable, but that both differ from individuals who
are aa. Let AA and Aa have the same chance of surviving from egg to
adult, and aa individuals have a lower viability. Then individuals who
are AA and Aa have different prospects for reproductive success. The
reason is that AA individuals never produce aa offspring, whereas Aa
individuals sometimes do. Success in reproduction is not measured merely
by number of offspring, but by the number of viable, fertile offspring.
In this instance, phenotype does not screen off genotype from reproduc-
tive success. But, as argued in section 3, the organism can be a unit of
selection even when there is dominance.

\[11\] This happens often, but not always. When the chain is deterministic, or when it fails
to include all factors that play a causal role, screening-off can fail. See Sober (1992) for
discussion.

\[12\] We owe this observation to Marsha Ensor, Julie Faulhaber, and Jennifer Hoepner.

\[13\] Brandon says that “the notions of genotype and phenotype are not mutually exclusive.
The genotype of an organism is part of its phenotype. Thus my claim commits me to the
position that any change in genotype that does lead to a change in reproductive success
must also be a change in the organism’s phenotype” (1990, 84–85). This stipulation would
save Brandon’s proposal from the problem posed by dominance. However, this proposal
endangers Brandon’s whole enterprise. If phenotype is to screen off genotype, then phe-
notype cannot include genotype. Violation of this requirement would mean that some of
We turn now to Brandon’s application of the screening-off criterion to the task of distinguishing group from individual organism as units of selection. Assuming that the groups in question exhibit differential reproduction, Brandon proposes that the group is a unit of selection precisely when there is “some group property (the group ‘phenotype’) that screens-off all other properties from group reproductive success” (1990, 87). Modifying Brandon’s notation slightly, his idea is that the group is a unit of selection precisely when

$$\text{Exp}(n \mid G&P) = \text{Exp}(n \mid G) \neq \text{Exp}(n \mid P).$$

Here \(\text{Exp}(n \mid \_\_\_)\) means that \(n\) is the expected number of propagule groups that a group produces (conditional on \(\_\_\_\_)\),\(^{14}\) \(G\) is the group phenotype, and \(P\) is a specification of the phenotypes of the organisms in the group. Brandon remarks that it is not inevitable that this two-part condition be satisfied. He describes a case in which the equality is true, but the inequality is not, the latter because “the phenotype of each individual within the group would determine that individual’s adaptedness, and the adaptedness values of each member of the group would determine the adaptedness value of the group” (ibid.).

We are puzzled as to why the inequality demanded by this criterion should ever be true, since the unary and relational properties of individuals evidently determine the properties of the group. Consider the case of the honeybee’s barbed stinger. Here the group is a unit of selection; groups benefit because they contain individuals who have barbed stingers. Yet, in this instance, the group phenotype is determined by the phenotypes of the individuals in the group.

One possible solution is to restrict what one means by a “group property” and by a “property of an organism”. Brandon observes that “it is not completely clear what should count as group properties” (ibid.), but adds that “obvious examples include . . . the relative frequency of certain alleles within the group [and] the phenotypic distribution within the group . . .” (ibid.). However, these group properties apparently do not screen-off, since they are determined by the array of properties that the individuals in the group possess.\(^{15}\)

---

\(^{14}\)Brandon focuses on the number of offspring groups, without taking their census size into account, because he feels that this is essential for the idea of group selection. This formulation will be problematic when parental groups all have the same number of groups as offspring, but differ in the census size of the groups they found.

\(^{15}\)Brandon connects his screening-off criterion for the units of selection problem with a more general view concerning what constitutes the best explanation of an effect. He suggests that if \(Y\) screens off \(X\) from \(Z\), then \(Y\) is a better explanation of \(Z\) than \(X\) is. Mitchell
7. On Whether Genic Selectionism Is Both Substantive and Plausible. We have argued on several occasions (e.g., Sober 1984, 1990b, 1993; Wilson and Sober 1989, 1994a) that Dawkins’s (1976, 1982) thesis that the gene is the one and only unit of selection is either false or vacuous. Sterelny and Kitcher (1988) have defended Dawkins, arguing that his claim is both nontrivial and plausible. According to them, Dawkins’s substantive point is that “barring complication, the average ability of the genes in the gene pool to leave copies of themselves increases with time” (ibid., 340). We believe that this claim is not part of Dawkins’s theory, and, in any event, is not biologically plausible.

Dawkins has frequently emphasized that natural selection operating within the confines of a single group will eliminate altruistic characteristics. The same point holds if one talks about genes. A gene for altruism will be displaced by a gene for selfishness in this instance.

Whether the issue concerns phenotypes or genes, the relationship between selfishness and altruism is the one depicted in figure 4.1. The quantity $\bar{w}$ measures the average fitness of the individuals in the population; equivalently, it measures the average fitness of the altruistic and selfish genes in the population. This means that purely within-group selection reduces the value of $\bar{w}$. As selfishness displaces altruism, the individuals in the population decline in their average fitness. Indeed, the very same thing is true of genes. Genes become less fit as a result of the process of “subversion from within” that Dawkins has highlighted. This shows, we believe, that Sterelny and Kitcher’s positive reconstruction of what is supposed to be nontrivial in Dawkins’s theory cannot be sustained. The fitness of genes does not increase when genes for selfishness replace genes for altruism (Sober 1990b). Ironically, the fitness of genes can increase when group selection leads altruism to evolve; but group selection is a process that Dawkins would not touch with a stick.

Sterelny and Kitcher (1988), Kitcher et al. (1990), and Waters (1991) have argued that a point in favor of the selfish gene point of view is that all selection processes can be represented in terms of single genes and their properties. They note that the same cannot be said of the organism or the group as units of selection. We agree that this difference exists, but we see it as a defect, not a strength. The argument that Kitcher, Sterelny, and Waters present here is a variant of the “representation ar-

---

(1987) shows that by switching explananda, Brandon’s criterion can be used to defend the genic point of view. Sober (1992) also develops objections to this proposal. Brandon et al. (1994) reply to this criticism.

16 Alternatively, one could interpret Sterelny and Kitcher as saying, not that $\bar{w}$ increases, but that genes of higher than average fitness tend to leave more copies of themselves. This reading turns Dawkins’s position into a triviality; it is true even when the group is the unit of selection.
argument” advanced by Williams and Dawkins discussed in section 4. If even group selection can be represented as a kind of genic selection, then genic selectionism is not a substantive alternative to anything. The selfish gene theory is vacuous if it is consistent with any and all types of selection process.

If it is not an automatic truism that the group or the organism is the unit of selection, the same should hold for the gene. An adequate clarification of the units of selection problem should treat these three levels on a conceptual par, not because they are equally correct, but because they should be evaluated by the same standards. 17

8. Common Fate. A proper understanding of the units of selection problem must take account of an important symmetry: Just as organisms are parts of groups, so genes are parts of organisms. The parts of a whole can interact cooperatively, enhancing the fitness of the whole at their own expense. Alternatively, the parts can interact competitively, enhancing their own fitness at the expense of the whole in which they reside. In the former case, the parts behave altruistically; in the latter, they behave selfishly.

Dawkins (1976, 1982) has rightly emphasized that evolution does not inevitably produce the highly integrated and well-adapted individual organisms we now observe. If this is to happen, competition among the parts of organisms must be modest. The Mendelian system has largely succeeded in creating this circumstance by making meiosis “fair”. Each of the genes in an organism standardly has the same chance of making it to the next generation. Within any organism, the genes are identical in fitness. 18 Exceptions to this pattern occur, of course, as in meiotic drive. But these exceptions aside, the genes in an organism have a common fate (Sober 1981, Walton 1991, and Wilson and Sober 1994a); this helps explain why organisms were able to evolve into functionally integrated wholes.

When the genes inside an organism sink or swim together, competition can occur between organisms, but not within them. In such cases, the unit of selection is the organism, not the gene. Meiotic drive, on the other hand, is a genuine case of the gene as unit of selection (as is the dynamics

17Dawkins treats genes that cooperate with each other and genes that compete with each other as both exemplifying the gene as unit of selection. This contributes to the vacuity of his version of genic selectionism; no matter what a gene does, it is “acting in its own interest”. No such confusion could arise in connection with the organism/group relationship. An organism that sacrifices its own welfare for the sake of the group differs from an organism that sacrifices group welfare for its own selfish advantage.

18Of course, germ line and somatic copies of the same gene have different probabilities of making it to the next generation. The point is that different germ line genes in the same organism have the same probability. Buss (1987) discusses how this arrangement evolved.
of junk DNA). The gene is sometimes the unit of selection, but very often it is not.\textsuperscript{19}

Empirically detectable cases of meiotic drive involve both genic and organismic selection, acting in opposite directions.\textsuperscript{20} Within an organism, the driving allele $D$ is fitter than the normal allele $N$ against which it competes. However, organisms with two copies of $D$ do worse than organisms with one or zero. Two causal processes are at work here. In one, genes in the same organism compete against each other; in the other, all the genes in the same organism are in the same boat.

These ideas may be frameshifted upward one level to provide a perspective on how groups and organisms are related to each other as candidate units of selection. We may begin by asking why groups are often less integrated and adapted than the organisms that are their parts. The answer is that within-group competition is often substantial. Organisms in the same group often compete with each other and have unequal chances of surviving and reproducing.

When group and organismic selection occur together, two types of causal process occur—one within groups, the other between them. Let us consider the example of how altruism ($A$) and selfishness ($S$) evolve. Within any group, $S$ individuals are fitter than $A$ individuals. But groups that contain more $S$ individuals do worse than groups that contain fewer. The evolution of the $D$ and $N$ alleles is isomorphic with the evolution of the $S$ and $A$ phenotypes.

Unifying these two examples is the concept of common fate. In each case, a process of competition obtains between parts, but a second process binds together the parts in a single whole by common fate. One causal process affects the parts in the same whole differentially, but another lumps together the parts in the same whole and treats them similarly.

We so far have addressed the complicated type of case in which selection occurs at the level of parts and also at the level of wholes. However, simpler scenarios in which there is just one unit of selection can be described merely by suppressing variation in fitness at all levels but one. Cases in which group and genic selection do not occur, and organismic selection is the only process at work, are easily described. Similarly monolithic dynamics can be described for the other candidate levels. In all such cases, identifying the unit(s) of selection involves discovering how parts and wholes vary in fitness and why they do so.

\textsuperscript{19}This proposal is orthogonal to the additivity criterion discussed in section 4. If the genes inside an organism are bound together by common fate, the gene fails to be a unit of selection, regardless of whether genotypic fitnesses are additive.

\textsuperscript{20}Without selection against the driving gene, it will go to fixation. If so, the population will contain no heterozygotes, and the scientist will be unable to see that the gene is, in fact, a driving gene.
Our criterion does not have the consequence that frequency dependent selection is automatically an instance of group selection, a problem discussed in section 5. It is not enough that groups vary in fitness because of their different internal compositions; for the members of the group to have “common fate”, some property of the group must have the effect of putting them “in the same boat”.

When wholes compete against wholes and parts also compete against parts, two units of selection exist. The traits that then evolve will often represent compromises between what is good for the whole and what is good for the parts. For example, when selection takes place purely within the confines of a single population, individual selection will lead to a sex ratio in which parents invest equally in the two sexes (Fisher 1958); when the sexes are equally costly, the sex ratio will be 1:1. Alternatively, when selection is purely between groups, they should evolve a heavily female-biased sex ratio in which males are produced only to the extent that they are needed to fertilize the females. The female-biased sex ratios so often observed in nature typically are compromises between these two “pure” cases. The sex ratio is not a purely individual adaptation, nor did it evolve solely because it benefits the group. It evolved for two, conflicting, reasons, and so is imperfect when judged by either monolithic criterion.

A similar compromise solution is evident in observed cases of meiotic drive. The driving gene does not go to fixation (as it would if the gene were the sole unit of selection), nor does the driving gene go extinct (which would happen if the organism were the sole unit of selection).

In section 2, $X$ was said to be a unit of selection in the evolution of trait $T$ precisely when one of the factors that influenced $T$’s evolution was that $T$ benefitted Xs. We now can use the concept of “common fate” to clarify this schematic idea. The crucial notion is differential benefit. When a trait is an organismic adaptation (and so the organism was a unit of selection in its evolution), the trait benefitted organisms in the sense that organisms who had the trait did better than organisms in the same group who did not. Our initial formulation raised the question of whether genes are always units of selection, since any trait that evolves seems to “benefit” the genes that code for it. We now can see that this impression is mistaken. If all the genes in an organism are “in the same boat”, one gene cannot do better than other genes in the same organism.

We are not suggesting that the idea of common fate is original with us. Dawkins and Hull recognize that if the units of selection problem concerns what types of adaptations have evolved, then it is an issue about vehicles/interactors, not replicators. We believe that the idea of common fate helps clarify what it takes for genes, organisms, and groups to be vehicles/interactors.
9. Realism, Pluralism, and Conventionalism. Some have questioned why it is necessary to choose a single view concerning which units of selection exist in nature. For example, Dawkins (1982) has argued that choice of the gene as the single unit of selection is a matter of convenience, not a matter of fact. Several philosophers (e.g., Cassidy 1978, Sterelny and Kitcher 1988, Kitcher et al. 1990, Waters 1991) have elaborated positions of this sort.

We believe that the units of selection problem is factual, not conventional, because different hypotheses about the units of selection typically make contrary predictions about which traits will evolve. For example, pure group selection will lead altruism to evolve, whereas pure individual selection will lead selfishness to evolve instead. Since the mix of characters found in a population is an observable matter, we have here an uncontroversial reason to regard the units of selection problem as nonconventional in character.\footnote{We take it to be \textit{sufficient} for the dispute between two hypotheses to be nonconventional that they make contrary predictions about observables. We will not discuss here whether this is a necessary condition. We note, however, that competing hypotheses about the units of selection sometimes predict the same equilibrium configuration. This happens, for example, when group and individual selection both favor the evolution of a particular trait. However, even in this kind of case, it usually is possible to design a test to discriminate between the competing hypotheses.}

As noted in section 7, Sterelny and Kitcher (1988) and Kitcher et al. (1990) argued that the causal processes at work in natural selection can always be described in terms of what happens to genes. We grant this, but the sense of “causal description” they discuss is not relevant to the empirical problem of deciding what types of adaptation are found in nature. As noted earlier, their argument is simply a version of the representation argument advanced by Williams (1966) and Dawkins (1976) on which we have already commented.

Our “realist” position with respect to the units of selection problem does not force us to choose between every pair of causal descriptions. When zebras evolve a faster running speed, they also evolve a suite of genes that codes for this phenotype. When genotype causes phenotype, both types of trait are causes of survival and reproductive success, and both “benefit” in the sense that both are made to increase in frequency. There is no need to choose (Mitchell 1987). However, in this instance, the organism, not the gene or the group, is the unit of selection. The genes in a zebra are bound together by a common fate; this is quite consistent with the fact that zebras with one suite of genes run faster than zebras with another.

Although we believe that the units of selection problem is substantive and nonconventional, we recognize that the genic point of view often has heuristic value, apart from whether the gene is in fact a unit of selection.
Even when group selection occurs (and so the group is a unit of selection), thinking about evolution from the point of view of a single gene (e.g., a gene for altruism) can be useful. We are inclined to be pluralists at the level of heuristic approaches and rather more monistic at the level of factual statements about nature.  

10. Concluding Remarks. In this article, we have tried to canvas some of the themes that have occupied philosophical reflection on the units of selection problem during the last ten to fifteen years. However, space limitations prevent us from considering other important issues. For example, Williams (1966) and Dawkins (1976) appealed to a principle of parsimony to justify the single gene as unit of selection, and we have not attended to the role of parsimony considerations in the units of selection problem (but see Sober 1990a). And we have not discussed the pragmatic advantages of different frameworks for thinking about selection, or empirical problems of measurement and testing (Lloyd 1986, 1988). We also have had to pass over questions concerning the interpretation of Hamilton’s (1964) notion of inclusive fitness and whether kin selection is a type of group selection (Sober 1993; Wilson and Sober 1989, 1994a), the status of species selection (Eldredge and Gould 1972, Stanley 1979, Sober 1984, Lloyd 1988, Williams 1992, Lloyd and Gould 1993), and the connection of units of selection issues with the concept of individuality (Hull 1980; Sober 1991, 1993). And we have had to shy away from discussing more general philosophical questions concerning causality, explanation, realism, and conventionalism. We regard all of these problems as interesting and important. It is not for nothing that this problem has excited so much philosophical discussion.

REFERENCES
———. (1982), The Extended Phenotype: The Gene as the Unit of Selection. San Francisco: Freeman.

22 The relationship between what happens in nature and which approaches are heuristic exhibits an interesting asymmetry. When group selection is not operating, it is hard to see the utility of representing processes at the group level. However, when group selection is operating, representing processes at the level of genes can be useful.


