Journal of Philosophy, Inc.

The Return of the Gene

Author(s): Kim Sterelny and Philip Kitcher

Source: The Journal of Philosophy, Vol. 85, No. 7 (Jul., 1988), pp. 339-361

Published by: Journal of Philosophy, Inc.

Stable URL: https://www.jstor.org/stable/2026953

Accessed: 13-12-2019 07:58 UTC

REFERENCES

Linked references are available on JSTOR for this article: https://www.jstor.org/stable/2026953?seq=1&cid=pdf-reference#references_tab_contents You may need to log in to JSTOR to access the linked references.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at https://about.jstor.org/terms



 $\it Journal\ of\ Philosophy,\ Inc.$ is collaborating with JSTOR to digitize, preserve and extend access to $\it The\ Journal\ of\ Philosophy$

THE JOURNAL OF PHILOSOPHY

VOLUME LXXXV, NO. 7, JULY 1988

THE RETURN OF THE GENE*

e have two images of natural selection. The orthodox story is told in terms of individuals. More organisms of any given kind are produced than can survive and reproduce to their full potential. Although these organisms are of a kind, they are not identical. Some of the differences among them make a difference to their prospects for survival or reproduction, and hence, on the average, to their actual reproduction. Some of the differences which are relevant to survival and reproduction are (at least partly) heritable. The result is evolution under natural selection, a process in which, barring complications, the average fitness of the organisms within a kind can be expected to increase with time.

There is an alternative story. Richard Dawkins¹ claims that the "unit of selection" is the gene. By this he means not just that the result of selection is (almost always) an increase in frequency of some gene in the gene pool. That is uncontroversial. On Dawkins's conception, we should think of genes as differing with respect to properties that affect their abilities to leave copies of themselves. More genes appear in each generation than can copy themselves up to their full potential. Some of the differences among them make a

* We are equally responsible for this paper which was written when we discovered that we were writing it independently. We would like to thank those who have offered helpful suggestions to one or both of us, particularly Patrick Bateson, Robert Brandon, Peter Godfrey-Smith, David Hull, Richard Lewontin, Lisa Lloyd, Philip Pettit, David Scheel, and Elliott Sober.

¹ The claim is made in *The Selfish Gene* (New York: Oxford, 1976); and, in a somewhat modified form, in *The Extended Phenotype* (San Francisco: Freeman, 1982). We shall discuss the difference between the two versions in the final section of this paper, and our reconstruction will be primarily concerned with the later version of Dawkins's thesis. We shall henceforth refer to *The Selfish Gene* as SG, and to *The Extended Phenotype* as EP. To forestall any possible confusion, our reconstruction of Dawkins's position does not commit us to the provocative claims about altruism and selfishness on which many early critics of SG fastened.

0022-362X/88/8507/0339\$02.30

© 1988 The Journal of Philosophy, Inc.

difference to their prospects for successful copying and hence to the number of actual copies that appear in the next generation. Evolution under natural selection is thus a process in which, barring complication, the average ability of the genes in the gene pool to leave copies of themselves increases with time.

Dawkins's story can be formulated succinctly by introducing some of his terminology. Genes are *replicators* and selection is the struggle among *active germ-line* replicators. Replicators are entities that can be copied. Active replicators are those whose properties influence their chances of being copied. Germ—line replicators are those which have the potential to leave infinitely many descendants. Early in the history of life, coalitions of replicators began to construct *vehicles* through which they spread copies of themselves. Better replicators build better vehicles, and hence are copied more often. Derivatively, the vehicles associated with them become more common too. The orthodox story focuses on the successes of prominent vehicles—individual organisms. Dawkins claims to expose an underlying struggle among the replicators.

We believe that a lot of unnecessary dust has been kicked up in discussing the merits of the two stories. Philosophers have suggested that there are important connections to certain issues in the philosophy of science: reductionism, views on causation and natural kinds, the role of appeals to parsimony. We are unconvinced. Nor do we think that a willingness to talk about selection in Dawkinspeak brings any commitment to the adaptationist claims which Dawkins also holds. After all, adopting a particular perspective on selection is logically independent from claiming that selection is omnipresent in evolution.

In our judgment, the relative worth of the two images turns on two theoretical claims in evolutionary biology.

- 1. Candidate units of selection must have systematic causal consequences. If Xs are selected for, then X must have a systematic effect on its expected representation in future generations.
- 2. Dawkins's gene selectionism offers a more general theory of evolution. It can also handle those phenomena which are grist to the mill of individual selection, but there are evolutionary phenomena which fit the picture of individual selection ill or not at all, yet which can be accommodated naturally by the gene selection model.

Those skeptical of Dawkins's picture—in particular, Elliott Sober, Richard Lewontin, and Stephen Jay Gould—doubt whether genes can meet the condition demanded in (1). In their view, the phenomena of epigenesis and the extreme sensitivity of the phenotype to

gene combinations and environmental effects undercut genic selectionism. Although we believe that these critics have offered valuable insights into the character of sophisticated evolutionary modeling, we shall try to show that these insights do not conflict with Dawkins's story of the workings of natural selection. We shall endeavor to free the thesis of genic selectionism from some of the troublesome excresences which have attached themselves to an interesting story.

I. GENE SELECTION AND BEAN-BAG GENETICS

Sober and Lewontin² argue against the thesis that all selection is genic selection by contending that many instances of selection do not involve selection for properties of individual alleles. Stated rather loosely, the claim is that, in some populations, properties of individual alleles are not positive causal factors in the survival and reproductive success of the relevant organisms. Instead of simply resting this claim on an appeal to our intuitive ideas about causality, Sober has recently provided an account of causal discourse which is intended to yield the conclusion he favors, thus rebutting the proposals of those (like Dawkins) who think that properties of individual alleles can be causally efficacious.³

The general problem arises because replicators (genes) combine to build vehicles (organisms) and the effect of a gene is critically dependent on the company it keeps. However, recognizing the general problem, Dawkins seeks to disentangle the various contributions of the members of the coalition of replicators (the genome). To this end, he offers an analogy with a process of competition among rowers for seats in a boat. The coach may scrutinize the relative times of different teams but the competition can be analyzed by investigating the contributions of individual rowers in different contexts (SG 40/1 91/2, EP 239).

Sober's Case. At the general level, we are left trading general intuitions and persuasive analogies. But Sober (and, earlier, Sober and Lewontin) attempted to clarify the case through a particular example. Sober argues that heterozygote superiority is a phenomenon that cannot be understood from Dawkins's standpoint. We shall discuss Sober's example in detail; our strategy is as follows. We first set out Sober's case: heterozygote superiority cannot be understood as a gene-level phenomenon, because only pairs of genes can be, or fail to be, heterozygous. Yet being heterozygous can be causally

² "Artifact, Cause and Genic Selection," *Philosophy of Science*, XLIX (1982): 157-180.

⁵ See Sober, *The Nature of Selection* (Cambridge: MIT, 1984), chs. 7–9, especially 302–314. We shall henceforth refer to this book as NS.

salient in the selective process. Against Sober, we first offer an analogy to show that there must be something wrong with his line of thought: from the gene's eye view, heterozygote superiority is an instance of a standard selective phenomenon, namely *frequency-dependent* selection. The advantage (or disadvantage) of a trait can depend on the frequency of that trait in other members of the relevant population.

Having claimed that there is something wrong with Sober's argument, we then try to say what is wrong. We identify two principles on which the reasoning depends. First is a general claim about causal uniformity. Sober thinks that there can be selection for a property only if that property has a positive uniform effect on reproductive success. Second, and more specifically, in cases where the heterozygote is fitter, the individuals have no uniform causal effect. We shall try to undermine both principles, but the bulk of our criticism will be directed against the first.

Heterozygote superiority occurs when a heterozygote (with genotype Aa, say) is fitter than either homozygote (AA or aa). The classic example is human sickle-cell anemia: homozygotes for the normal allele in African populations produce functional hemoglobin but are vulnerable to malaria, homozygotes for the mutant ("sickling") allele suffer anemia (usually fatal), and heterozygotes avoid anemia while also having resistance to malaria. The effect of each allele varies with context, and the contexts across which variation occurs are causally relevant. Sober writes:

In this case, the a allele does not have a unique causal role. Whether the gene a will be a positive or a negative causal factor in the survival and reproductive success of an organism depends on the genetic context. If it is placed next to a copy of A, a will mean an increase in fitness. If it is placed next to a copy of itself, the gene will mean a decrement in fitness (NS 303).

The argument against Dawkins expressed here seems to come in two parts. Sober relies on the principle

(A) There is selection for property P only if in all causally relevant background conditions P has a positive effect on survival and reproduction.

He also adduces a claim about the particular case of heterozygote superiority.

(B) Although we can understand the situation by noting that the heterozygote has a uniform effect on survival and reproduction, the property

of having the A allele and the property of having the a allele cannot be seen as having uniform effects on survival and reproduction.

We shall argue that both (A) and (B) are problematic.

Let us start with the obvious reply to Sober's argument. It seems that the heterozygote superiority case is akin to a familiar type of frequency-dependent selection. If the population consists just of AAs and a mutation arises, the a-allele, then, initially a is favored by selection. Even though it is very bad to be aa, a alleles are initially likely to turn up in the company of A alleles. So they are likely to spread, and, as they spread, they find themselves alongside other a alleles, with the consequence that selection tells against them. The scenario is very similar to a story we might tell about interactions among individual organisms. If some animals resolve conflicts by playing hawk and others play dove, then, if a population is initially composed of hawks (and if the costs of bloody battle outweigh the benefits of gaining a single resource), doves will initially be favored by selection. 4 For they will typically interact with hawks, and, despite the fact that their expected gains from these interactions are zero, they will still fare better than their rivals whose expected gains from interactions are negative. But, as doves spread in the population, hawks will meet them more frequently, with the result that the expected payoffs to hawks from interactions will increase. Because they increase more rapidly than the expected payoffs to the doves, there will be a point at which hawks become favored by selection, so that the incursion of doves into the population is halted.

We believe that the analogy between the case of heterozygote superiority and the hawk-dove case reveals that there is something troublesome about Sober's argument. The challenge is to say exactly what has gone wrong.

Causal Uniformity. Start with principle (A). Sober conceives of selection as a force, and he is concerned to make plain the effects of component forces in situations where different forces combine. Thus, he invites us to think of the heterozygote superiority case by analogy with situations in which a physical object remains at rest because equal and opposite forces are exerted on it. Considering the situation only in terms of net forces will conceal the causal structure of the situation. Hence, Sober concludes, our ideas about units of selection should penetrate beyond what occurs on the average, and

⁴ For details, see John Maynard Smith, Evolution and the Theory of Games (New York: Cambridge, 1982); and, for a capsule presentation, Philip Kitcher, Vaulting Ambition: Sociobiology and the Quest for Human Nature (Cambridge: MIT, 1985), pp. 88-97.

we should attempt to isolate those properties which positively affect survival and reproduction in every causally relevant context.

Although Sober rejects determinism, principle (A) seems to hanker after something like the uniform association of effects with causes that deterministic accounts of causality provide. We believe that the principle cannot be satisfied without doing violence to ordinary ways of thinking about natural selection, and, once the violence has been exposed, it is not obvious that there is any way to reconstruct ideas about selection that will fit Sober's requirement.

Consider the example of natural selection, the case of industrial melanism. We are inclined to say that the moths in a Cheshire wood. where lichens on many trees have been destroyed by industrial pollutants, have been subjected to selection pressure and that there has been selection for the property of being melanic. But a moment's reflection should reveal that this description is at odds with Sober's principle. For the wood is divisible into patches, among which are clumps of trees that have been shielded from the effects of industrialization. Moths who spend most of their lives in these areas are at a disadvantage if they are melanic. Hence, in the population comprising all the moths in the wood, there is no uniform effect on survival and reproduction: in some causally relevant contexts (for moths who have the property of living in regions where most of the trees are contaminated), the trait of being melanic has a positive effect on survival and reproduction, but there are other contexts in which the effect of the trait is negative.

The obvious way to defend principle (A) is to split the population into subpopulations and identify different selection processes as operative in different subgroups. This is a revisionary proposal, for our usual approach to examples of industrial melanism is to take a coarse-grained perspective on the environments, regarding the existence of isolated clumps of uncontaminated trees as a perturbation of the overall selective process. Nonetheless, we might be led to make the revision, not in the interest of honoring a philosophical prejudice, but simply because our general views about selection are consonant with principle (A), so that the reform would bring our treatment of examples into line with our most fundamental beliefs about selection.

In our judgment, a defense of this kind fails for two connected reasons. First, the process of splitting populations may have to continue much further—perhaps even to the extent that we ultimately

⁵ The *locus classicus* for discussion of this example is H.B.D. Kettlewell, *The Evolution of Melanism* (New York: Oxford, 1973).

conceive of individual organisms as making up populations in which a particular type of selection occurs. For, even in contaminated patches, there may be variations in the camouflaging properties of the tree trunks and these variations may combine with propensities of the moths to cause local disadvantages for melanic moths. Second, as many writers have emphasized, evolutionary theory is a statistical theory, not only in its recognition of drift as a factor in evolution but also in its use of fitness coefficients to represent the expected survivorship and reproductive success of organisms. The envisaged splitting of populations to discover some partition in which principle (A) can be maintained is at odds with the strategy of abstracting from the thousand natural shocks that organisms in natural populations are heir to. In principle, we could relate the biography of each organism in the population, explaining in full detail how it developed, reproduced, and survived, just as we could track the motion of each molecule of a sample of gas. But evolutionary theory, like statistical mechanics, has no use for such a fine grain of description: the aim is to make clear the central tendencies in the history of evolving populations, and, to this end, the strategy of averaging, which Sober decries, is entirely appropriate. We conclude that there is no basis for any revision that would eliminate those descriptions which run counter to principle (A).

At this point, we can respond to the complaints about the gene's eye view representation of cases of heterozygote superiority. Just as we can give sense to the idea that the trait of being melanic has a unique environment-dependent effect on survival and reproduction, so too we can explicate the view that a property of alleles, to wit, the property of directing the formation of a particular kind of hemoglobin, has a unique environment-dependent effect on survival and reproduction. The alleles form parts of one another's environments. and, in an environment in which a copy of the A allele is present, the typical trait of the S allele (namely, directing the formation of deviant hemoglobin) will usually have a positive effect on the chances that copies of that allele will be left in the next generation. (Notice that the effect will not be invariable, for there are other parts of the genomic environment which could wreak havoc with it). If someone protests that the incorporation of alleles as themselves part of the environment is suspect, then the immediate rejoinder is that, in cases of behavioral interactions, we are compelled to treat organisms as parts of one another's environments.⁶ The effects of playing hawk

⁶ In the spirit of Sober's original argument, one might press further. Genic selectionists contend that an A allele can find itself in two different environments, one in which the effect of directing the formation of a normal globin chain is

depend on the nature of the environment, specifically on the frequency of doves in the vicinity.⁷

The Causal Powers of Alleles. We have tried to develop our complaints about principle (A) into a positive account of how cases of heterozygote superiority might look from the gene's eye view. We now want to focus more briefly on (B). Is it impossible to reinterpret the examples of heterozygote superiority so as to ascribe uniform effects on survival and reproduction to allelic properties? The first point to note is that Sober's approach formulates the Dawkinsian point of view in the wrong way: the emphasis should be on the effects of properties of alleles, not on allelic properties of organisms (like the property of having an A allele) and the accounting ought to be done in terms of allele copies. Second, although we argued above that the strategy of splitting populations was at odds with the character of evolutionary theory, it is worth noting that the same strategy will be available in the heterozygote superiority case.

positive and one in which that effect is negative. Should we not be alarmed by the fact that the distribution of environments in which alleles are selected is itself a function of the frequency of the alleles whose selection we are following? No. The phenomenon is thoroughly familiar from studies of behavioral interactions—in the hawk-dove case we treat the frequency of hawks both as the variable we are tracking and as a facet of the environment in which selection occurs. Maynard Smith makes the parallel fully explicit in his paper "How To Model Evolution," in John Dupre, ed., The Latest on the Best: Essays on Optimality and Evolution (Cambridge: MIT, 1987), pp. 119–131, especially pp. 125/6.

⁷ Moreover, we can explicitly recognize the co-evolution of alleles with allelic environments. A fully detailed general approach to population genetics from the Dawkinsian point of view will involve equations that represent the functional dependence of the distribution of environments on the frequency of alleles, and equations that represent the fitnesses of individual alleles in different environments. In fact, this is just another way of looking at the standard population genetics equations. Instead of thinking of W_{AA} as the expected contribution to survival and reproduction of (an organism with) an allelic pair, we think of it as the expected contribution of copies of itself of the allele A in environment A. We now see W_{AS} as the expected contribution of S in environment A. The frequencies p, q are not only the frequencies of the alleles, but also the frequencies with which certain environments occur. The standard definitions of the overall (net) fitnesses of the alleles are obtained by weighting the fitnesses in the different environments by the frequencies with which the environments occur.

Lewontin has suggested to us that problems may arise with this scheme of interpretation if the population should suddenly start to reproduce asexually. But this hypothetical change could be handled from the genic point of view by recognizing an alteration of the coevolutionary process between alleles and their environments: whereas certain alleles used to have descendants that would encounter a variety of environments, their descendants are now found only in one allelic environment. Once the algebra has been formulated, it is relatively straightforward to extend the reinterpretation to this case.

Consider the following division of the original population: let P_1 be the collection of all those allele copies which occur next to an S allele, and let P_2 consist of all those allele copies which occur next to an A allele. Then the property of being A (or of directing the production of normal hemoglobin) has a positive effect on the production of copies in the next generation in P_1 , and conversely in P_2 . In this way, we are able to partition the population and to achieve a Dawkinsian redescription that meets Sober's principle (A)—just in the way that we might try to do so if we wanted to satisfy (A) in understanding the operation of selection on melanism in a Cheshire wood or on fighting strategies in a population containing a mixture of hawks and doves.

Objection: the "populations" just defined are highly unnatural, and this can be seen once we recognize that, in some cases, allele copies in the same organisms (the heterozygotes) belong to different "populations." Reply: so what? From the allele's point of view, the copy next door is just a critical part of the environment. The populations P_1 and P_2 simply pick out the alleles that share the same environment. There would be an analogous partition of a population of competing organisms which occurred locally in pairs such that some organisms played dove and some hawk. (Here, mixed pairs would correspond to heterozygotes).

So the genic picture survives an important initial challenge. The moral of our story so far is that the picture must be applied consistently. Just as paradoxical conclusions will result if one offers a partial translation of geometry into arithmetic, it is possible to generate perplexities by failing to recognize that the Dawkinsian *Weltanschauung* leads to new conceptions of environment and of population. We now turn to a different worry, the objection that genes are not "visible" to selection.

II. EPIGENESIS AND VISIBILITY

In a lucid discussion of Dawkins's early views, Gould claims to find a "fatal flaw" in the genic approach to selection. According to Gould, Dawkins is unable to give genes "direct visibility to natural selection." Bodies must play intermediary roles in the process of selection, and, since the properties of genes do not map in one-one fashion onto the properties of bodies, we cannot attribute selective advantages to individual alleles. We believe that Gould's concerns raise two important kinds of issues for the genic picture: (i) Can

⁸ "Caring Groups and Selfish Genes," in *The Panda's Thumb* (New York: Norton, 1980), p. 90. There is a valuable discussion of Gould's claims in Sober, NS 227 ff.

Dawkins sensibly talk of the effect of an individual allele on its expected copying frequency? (ii) Can Dawkins meet the charge that it is the phenotype that makes the difference to the copying of the underlying alleles, so that, whatever the causal basis of an advantageous trait, the associated allele copies will have enhanced chances of being replicated? We shall take up these questions in order.

Do Alleles Have Effects? Dawkins and Gould agree on the facts of embryology which subvert the simple Mendelian association of one gene with one character. But the salience of these facts to the debate is up for grabs. Dawkins regards Gould as conflating the demands of embryology with the demands of the theory of evolution. While genes' effects blend in embryological development, and while they have phenotypic effects only in concert with their gene-mates, genes "do not blend as they replicate and recombine down the generations. It is this that matters for the geneticist, and it is also this that matters for the student of units of selection" (EP 117).

Is Dawkins right? Chapter 2 of EP is an explicit defense of the meaningfulness of talk of "genes for" indefinitely complex morphological and behavioral traits. In this, we believe, Dawkins is faithful to the practice of classical geneticists. Consider the vast number of loci in *Drosophila melanogaster* which are labeled for eye-color traits—white, eosin, vermilion, raspberry, and so forth. Nobody who subscribes to this practice of labeling believes that a pair of appropriately chosen stretches of DNA, cultured in splendid isolation, would produce a detached eye of the pertinent color. Rather, the intent is to indicate the effect that certain changes at a locus would make against the background of the rest of the genome.

Dawkins's project here is important not just in conforming to traditions of nomenclature. Remember: Dawkins needs to show that we can sensibly speak of alleles having (environment-sensitive) effects, effects in virtue of which they are selected for or selected against. If we can talk of a gene for X, where X is a selectively important phenotypic characteristic, we can sensibly talk of the effect of an allele on its expected copying frequency, even if the effects are always indirect, via the characteristics of some vehicle.

What follows is a rather technical reconstruction of the relevant notion. The precision is needed to allow for the extreme environmental sensitivity of allelic causation. But the intuitive idea is simple: we can speak of genes for X if substitutions on a chromosome would lead, in the relevant environments, to a difference in the X-ishness of the phenotype.

Consider a species S and an arbitrary locus L in the genome of members of S. We want to give sense to the locution L is a locus

affecting P' and derivatively to the phrase 'G is a gene for P^{*} ' (where, typically, P will be a determinable and P^* a determinate form of P). Start by taking an environment for a locus to be an aggregate of DNA segments that would complement L to form the genome of a member of S together with a set of extra-organismic factors (those aspects of the world external to the organism which we would normally count as part of the organism's environment). Let a set of variants for L be any collection of DNA segments, none of which is debarred, on physico-chemical grounds, from occupying L. (This is obviously a very weak constraint, intended only to rule out those segments which are too long or which have peculiar physico-chemical properties). Now, we say that L is a locus affecting P in S relative to an environment E and a set of variants V just in case there are segments s, s^* , and s^{**} in V such that the substitution of s^{**} for s^* in an organism having s and s* at L would cause a difference in the form of P. against the background of E. In other words, given the environment E, organisms who are ss* at L differ in the form of P from organisms who are ss** at L and the cause of the difference is the presence of s* rather than s**. (A minor clarification: while s* and s** are distinct, we do not assume that they are both different from s.)

L is a locus affecting P in S just in case L is a locus affecting P in Srelative to any standard environment and a feasible set of variants. Intuitively, the geneticist's practice of labeling loci focuses on the "typical" character of the complementary part of the genome in the species, the "usual" extra-organismic environment, and the variant DNA segments which have arisen in the past by mutation or which "are likely to arise" by mutation. Can these vague ideas about standard conditions be made more precise? We think so. Consider first the genomic part of the environment. There will be numerous alternative combinations of genes at the loci other than L present in the species S. Given most of these gene combinations, we expect modifications at L to produce modifications in the form of P. But there are likely to be some exceptions, cases in which the presence of a rare allele at another locus or a rare combination of alleles produces a phenotypic effect that dominates any effect on P. We can either dismiss the exceptional cases as nonstandard because they are infrequent or we can give a more refined analysis, proposing that each of the nonstandard cases involves either (a) a rare allele at a locus L' or (b) a rare combination of alleles at loci L', L''... such that that locus (a) or those loci jointly (b) affect some phenotypic trait Q that dominates P in the sense that there are modifications of Q which prevent the expression of any modifications of P. As a concrete example, consider the fact that there are modifications at some loci in Drosophila which produce embryos that fail to develop heads; given such modifications elsewhere in the genome, alleles affecting eye color do not produce their standard effects!

We can approach standard extra-genomic environments in the same way. If L affects the form of P in organisms with a typical gene complement, except for those organisms which encounter certain rare combinations of external factors, then we may count those combinations as nonstandard simply because of their infrequency. Alternatively, we may allow rare combinations of external factors to count provided that they do not produce some gross interference with the organism's development, and we can render the last notion more precise by taking nonstandard environments to be those in which the population mean fitness of organisms in S would be reduced by some arbitrarily chosen factor (say, $\frac{1}{2}$).

Finally, the feasible variants are those which actually occur at L in members of S, together with those which have occurred at L in past members of S and those which are easily attainable from segments that actually occur at L in members of S by means of insertion, deletion, substitution, or transposition. Here the criteria for ease of attainment are given by the details of molecular biology. If an allele is prevalent at L in S, then modifications at sites where the molecular structure favors insertions, deletions, substitutions, or transpositions (so-called "hot spots") should count as easily attainable even if some of these modifications do not actually occur.

Obviously, these concepts of "standard conditions" could be articulated in more detail, and we believe that it is possible to generate a variety of explications, agreeing on the core of central cases but adjusting the boundaries of the concepts in different ways. If we now assess the labeling practices of geneticists, we expect to find that virtually all of their claims about loci affecting a phenotypic trait are sanctioned by all of the explications. Thus, the challenge that there is no way to honor the facts of epigenesis while speaking of loci that affect certain traits would be turned back.

Once we have come this far, it is easy to take the final step. An allele A at a locus L in a species S is for the trait P^* (assumed to be a determinate form of the determinable characteristic P) relative to a local allele B and an environment E just in case (a) L affects the form of P in S, (b) E is a standard environment, and (c) in E organisms that are AB have phenotype P^* . The relativization to a local allele is necessary, of course, because, when we focus on a target allele rather than a locus, we have to extend the notion of the environment—as we saw in the last section, corresponding alleles are potentially important parts of one another's environments. If we say that A is for

 P^* (period), we are claiming that A is for P^* relative to standard environments and common local alleles or that A is for P^* relative to standard environments and itself.

Now, let us return to Dawkins and to the apparently outré claim that we can talk about genes for reading. Reading is an extraordinarily complex behavior pattern and surely no adaptation. Further, many genes must be present and the extra-organismic environment must be right for a human being to be able to acquire the ability to read. Dyslexia might result from the substitution of an unusual mutant allele at one of the loci, however. Given our account, it will be correct to say that the mutant allele is a gene for dyslexia and also that the more typical alleles at the locus are alleles for reading. Moreover, if the locus also affects some other (determinable) trait, say, the capacity to factor numbers into primes, then it may turn out that the mutant allele is also an allele for rapid factorization skill and that the typical allele is an allele for factorization disability. To say that A is an allele for P^* does not preclude saying that A is an allele for Q^* , nor does it commit us to supposing that the phenotypic properties in question are either both skills or both disabilities. Finally, because substitutions at many loci may produce (possibly different types of) dyslexia, there may be many genes for dyslexia and many genes for reading. Our reconstruction of the geneticists' idiom, the idiom which Dawkins wants to use, is innocent of any Mendelian theses about one-one mappings between genes and phenotypic traits.

Visibility. So we can defend Dawkins's thesis that alleles have properties that influence their chances of leaving copies in later generations by suggesting that, in concert with their environments (including their genetic environments), those alleles cause the presence of certain properties in vehicles (such as organisms) and that the properties of the vehicles are causally relevant to the spreading of copies of the alleles. But our answer to question (i) leads naturally to concerns about question (ii). Granting that an allele is for a phenotypic trait P^* and that the presence of P^* rather than alternative forms of the determinable trait P enhances the chances that an organism will survive and reproduce and thus transmit copies of the underlying allele, is it not P* and its competition which are directly involved in the selection process? What selection "sees" are the phenotypic properties. When this vague, but suggestive, line of thought has been made precise, we think that there is an adequate Dawkinsian reply to it.

The idea that selection acts directly on phenotypes, expressed in metaphorical terms by Gould (and earlier by Ernst Mayr), has been explored in an interesting essay by Robert Brandon.⁹ Brandon proposes that phenotypic traits screen off genotypic traits (in the sense of Wesley Salmon¹⁰):

$$Pr(O_n/G\&P) = Pr(O_n/P) \neq Pr(O_n/G)$$

where $\Pr(O_n/G\&P)$ is the probability that an organism will produce n offspring given that it has both a phenotypic trait and the usual genetic basis for that trait, $\Pr(O_n/P)$ is the probability that an organism will produce n offspring given that it has the phenotypic trait, and $\Pr(O_n/G)$ is the probability that it will produce n offspring given that it has the usual genetic basis. So fitness seems to vary more directly with the phenotype and less directly with the underlying genotype.

Why is this? The root idea is that the successful phenotype may occur in the presence of the wrong allele as a result of judicious tampering, and, conversely, the typical effect of a "good" allele may be subverted. If we treat moth larvae with appropriate injections, we can produce pseudomelanics that have the allele which normally gives rise to the speckled form and we can produce moths, foiled melanics, that carry the allele for melanin in which the developmental pathway to the emergence of black wings is blocked. The pseudomelanics will enjoy enhanced reproductive success in polluted woods and the foiled melanics will be at a disadvantage. Recognizing this type of possibility, Brandon concludes that selection acts at the level of the phenotype.¹¹

⁹ Gould, op.cit.; Mayr, Animal Species and Evolution (Cambridge: Harvard, 1963), p. 184; and Brandon, "The Levels of Selection," in Brandon and Richard Burian, eds., Genes, Organisms, Populations (Cambridge: MIT, 1984), pp. 133-141.

¹⁰ Brandon refers to Salmon's "Statistical Explanation," in Salmon, ed., Statistical Explanation and Statistical Relevance (Pittsburgh: University Press, 1971). It is now widely agreed that statistical relevance misses some distinctions which are important in explicating causal relevance. See, for example, Nancy Cartwright, "Causal Laws and Effective Strategies," Noûs, XIII (1979): 419–437; Sober, NS ch. 8; and Salmon, Scientific Explanation and the Causal Structure of the World (Princeton: University Press, 1984).

¹¹ Unless the treatments are repeated in each generation, the presence of the genetic basis for melanic coloration will be correlated with an increased frequency of grandoffspring, or of great-grandoffspring, or of descendants in some further generation. Thus, analogs of Brandon's probabilistic relations will hold only if the progeny of foiled melanics are treated so as to become foiled melanics, and the progeny of pseudomelanics are treated so as to become pseudomelanics. This point reinforces the claims about the relativization to the environment that we make below. Brandon has suggested to us in correspondence that now his preferred strategy for tackling issues of the units of selection would be to formulate a principle for identifying genuine environments.

Once again, there is no dispute about the facts. But our earlier discussion of epigenesis should reveal how genic selectionists will want to tell a different story. The interfering conditions that affect the phenotype of the vehicle are understood as parts of the allelic environment. In effect, Brandon, Gould, and Mayr contend that, in a polluted wood, there is selection for being dark colored rather than for the allelic property of directing the production of melanin, because it would be possible to have the reproductive advantage associated with the phenotype without having the allele (and conversely it would be possible to lack the advantage while possessing the allele). Champions of the gene's eye view will maintain that tampering with the phenotype reverses the typical effect of an allele by changing the environment. For these cases involve modification of the allelic environment and give rise to new selection processes in which allelic properties currently in favor prove detrimental. The fact that selection goes differently in the two environments is no more relevant than the fact that selection for melanic coloration may go differently in Cheshire and in Dorset.

If we do not relativize to a fixed environment, then Brandon's claims about screening off will not generally be true. 12 We suppose

¹² Intuitively, this will be because Brandon's identities depend on there being no correlation between O_n and G in any environment, except through the property P. Thus, ironically, the screening-off relations only obtain under the assumptions of simple bean-bag genetics! Sober seems to appreciate this point in a cryptic footnote (NS 229-230).

To see how it applies in detail, imagine that we have more than one environment and that the reproductive advantages of melanic coloration differ in the different environments. Specifically, suppose that E_1 contains m_1 organisms that have P (melanic coloration) and G (the normal genetic basis of melanic coloration), that E_2 contains m_2 organisms that have P and G, and that the probabilities $\Pr(O_n/G\&P\&E_1)$ and $\Pr(O_n/G\&P\&E_2)$ are different. Then, if we do not relativize to environments, we shall compute $\Pr(O_n/G\&P)$ as a weighted average of the probabilities relative to the two environments.

$$Pr(O_n/G\&P) = Pr(E_1/G\&P) \cdot Pr(O_n/G\&P\&E_1) + Pr(E_2/G\&P) \cdot Pr(O_n/G\&P\&E_2)$$
$$= m_1/(m_1 + m_2) \cdot Pr(O_n/G\&P\&E_1) + m_2/(m_1 + m_2) \cdot Pr(O_n/G\&P\&E_2)$$

Now, suppose that tampering occurs in E_2 so that there are m_3 pseudomelanics in E_2 . We can write $\Pr(O_n/P)$ as a weighted average of the probabilities relative to the two environments.

$$Pr(O_n/P) = Pr(E_1/P) \cdot Pr(O_n/P\&E_1) + Pr(E_2/P) \cdot Pr(O_n/P\&E_2).$$

By the argument that Brandon uses to motivate his claims about screening off, we can take $\Pr(O_n/G\&P\&E_i) = \Pr(O_n/P\&E_i)$ for i=1,2. However, $\Pr(E_1/P) = m_1/(m_1+m_2+m_3)$ and $\Pr(E_2/P) = (m_2+m_3)/(m_1+m_2+m_3)$, so that $\Pr(E_i/P) \neq \Pr(E_i/G\&P)$. Thus, $\Pr(O_n/G\&P) \neq \Pr(O_n/P)$, and the claim about screening off fails.

that Brandon intends to relativize to a fixed environment. But now he has effectively begged the question against the genic selectionist by deploying the orthodox conception of environment. Genic selectionists will also want to relativize to the environment, but they should resist the orthodox conception of it. On their view, the probability relations derived by Brandon involve an illicit averaging over environments (see fn. 12). Instead, genic selectionists should propose that the probability of an allele's leaving n copies of itself should be understood relative to the total allelic environment, and that the specification of the total environment ensures that there is no screening off of allelic properties by phenotypic properties. The probability of producing n copies of the allele for melanin in a total allelic environment is invariant under conditionalization on phenotype.

Here too the moral of our story is that Dawkinspeak must be undertaken consistently. Mixing orthodox concepts of the environment with ideas about genic selection is a recipe for trouble, but we have tried to show how the genic approach can be thoroughly articulated so as to meet major objections. But what is the point of doing so? We shall close with a survey of some advantages and potential drawbacks.

III. GENES AND GENERALITY

Relatively little fossicking is needed to uncover an extended defense of the view that gene selectionism offers a more general and unified picture of selective processes than can be had from its alternatives. Phenomena anomalous for the orthodox story of evolution by individual selection fall naturally into place from Dawkins' viewpoint. He offers a revision of the "central theorem" of Darwinism. Instead of expecting individuals to act in their best interests, we should expect an animal's behavior "to maximize the survival of genes 'for' that behavior, whether or not those genes happen to be in the body of that particular animal performing it" (EP 223).

The cases that Dawkins uses to illustrate the superiority of his own approach are a somewhat motley collection. They seem to fall into two general categories. First are outlaw and quasi-outlaw examples. Here there is competition among genes which cannot be translated into talk of vehicle fitness because the competition is among co-

Notice that, if environments are lumped in this way, then it will only be under fortuitous circumstances that the tampering makes the probabilistic relations come out as Brandon claims. Pseudomelanics would have to be added in both environments so that the weights remain exactly the same.

builders of a single vehicle. The second group comprises "extended phenotype" cases, instances in which a gene (or combination of genes) has selectively relevant phenotypic consequences which are not traits of the vehicle that it has helped build. Again the replication potential of the gene cannot be translated into talk of the adaptedness of its vehicle.

We shall begin with outlaws and quasi outlaws. From the perspective of the orthodox story of individual selection, "replicators at different loci within the same body can be expected to 'cooperate'." The allele surviving at any given locus tends to be one best (subject to all the constraints) for the whole genome. By and large this is a reasonable assumption. Whereas individual outlaw organisms are perfectly possible in groups and subvert the chances for groups to act as vehicles, outlaw genes seem problematic. Replication of any gene in the genome requires the organism to survive and reproduce, so genes share a substantial common interest. This is true of asexual reproduction, and, granting the fairness of meiosis, of sexual reproduction too.

But there is the rub. Outlaw genes are genes which subvert meiosis to give them a better than even chance of making it to the gamete, typically by sabotaging their corresponding allele (EP 136). Such genes are segregation distorters or meiotic drive genes. Usually, they are enemies not only of their alleles but of other parts of the genome, because they reduce the individual fitness of the organism they inhabit. Segregation distorters thrive, when they do, because they exercise their phenotypic power to beat the meiotic lottery. Selection for such genes cannot be selection for traits that make organisms more likely to survive and reproduce. They provide uncontroversial cases of selective processes in which the individualistic story cannot be told.

There are also related examples. Altruistic genes can be outlaw-like, discriminating against their genome mates in favor of the inhabitants of other vehicles, vehicles that contain copies of themselves. Start with a hypothetical case, the so-called "green beard" effect. Consider a gene Q with two phenotypic effects. Q causes its vehicle to grow a green beard and to behave altruistically toward green-bearded conspecifics. Q's replication prospects thus improve, but the particular vehicle that Q helped build does not have its prospects for survival and reproduction enhanced. Is Q an outlaw not just with respect to the vehicle but with respect to the vehicle builders? Will there be selection for alleles that suppress Q's effect? How the selection process goes will depend on the probability that Q's cobuilders

are beneficiaries as well. If Q is reliably associated with other gene kinds, those kinds will reap a net benefit from Q's outlawry.

So altruistic genes are sometimes outlaws. Whether coalitions of other genes act to suppress them depends on the degree to which they benefit only themselves. Let us now move from a hypothetical example to the parade case.

Classical fitness, an organism's propensity to leave descendants in the next generation, seems a relatively straightforward notion. Once it was recognized that Darwinian processes do not necessarily favor organisms with high classical fitness, because classical fitness ignores indirect effects of costs and benefits to relatives, a variety of alternative measures entered the literature. The simplest of these would be to add to the classical fitness of an organism contributions from the classical fitness of relatives (weighted in each case by the coefficient of relatedness). Although accounting of this sort is prevalent, Dawkins (rightly) regards it as just wrong, for it involves double bookkeeping and, in consequence, there is no guarantee that populations will move to local maxima of the defined quantity. This measure and measures akin to it, however, are prompted by Hamilton's rigorous development of the theory of inclusive fitness (in which it is shown that populations will tend toward local maxima of inclusive fitness). 13 In the misunderstanding and misformulation of Hamilton's ideas, Dawkins sees an important moral.

Hamilton, he suggests, appreciated the gene selectionist insight that natural selection will favor "organs and behavior that cause the individual's genes to be passed on, whether or not the individual is an ancestor" (EP 185). But Hamilton's own complex (and much misunderstood) notion of inclusive fitness was, for all its theoretical importance, a dodge, a "brilliant last-ditch rescue attempt to save the individual organism as the level at which we think about natural selection" (EP 187). More concretely, Dawkins is urging two claims: first, that the uses of the concept of inclusive fitness in practice are difficult, so that scientists often make mistakes; second, that such uses are conceptually misleading. The first point is defended by identifying examples from the literature in which good researchers have made errors, errors which become obvious once we adopt the gene selectionist perspective. Moreover, even when the inclusive fit-

¹⁸ For Hamilton's original demonstration, see "The Genetical Evolution of Social Behavior I," in G.C. Williams, ed., *Group Selection* (Chicago: Aldine, 1971), pp. 23–43. For a brief presentation of Hamilton's ideas, see Kitcher, *op.cit.*, pp. 77–87; and for penetrating diagnoses of misunderstandings, see A. Grafen, "How Not to Measure Inclusive Fitness," *Nature*, CCXCVIII (1982): 425/6; and R. Michod, "The Theory of Kin Selection," in Brandon and Burian, *op.cit.*, pp. 203–237.

ness calculations make the right predictions, they often seem to mystify the selective process involved (thus buttressing Dawkins's second thesis). Even those who are not convinced of the virtues of gene selectionism should admit that it is very hard to see the reproductive output of an organism's relatives as a property of that organism.

Let us now turn to the other family of examples, the "extended phenotype" cases. Dawkins gives three sorts of "extended" phenotypic effects: effects of genes—indeed key weapons in the competitive struggle to replicate—which are not traits of the vehicle the genes inhabit. The examples are of artifacts, of parasitic effects on host bodies and behaviors, and of "manipulation" (the subversion of an organism's normal patterns of behavior by the genes of another organism via the manipulated organism's nervous system).

Among many vivid, even haunting, examples of parasitic behavior, Dawkins describes cases in which parasites synthesize special hormones with the consequence that their hosts take on phenotypic traits that decrease their own prospects for reproduction but enhance those of the parasites (see, for a striking instance, EP 215). There are equally forceful cases of manipulation: cuckoo fledglings subverting their host's parental program, parasitic queens taking over a hive and having its members work for her. Dawkins suggests that the traits in question should be viewed as adaptations—properties for which selection has occurred—even though they cannot be seen as adaptations of the individuals whose reproductive success they promote, for those individuals do not possess the relevant traits. Instead, we are to think in terms of selectively advantageous characteristics of alleles which orchestrate the behavior of several different vehicles, some of which do not include them.

At this point there is an obvious objection. Can we not understand the selective processes that are at work by focusing not on the traits that are external to the vehicle that carries the genes, but on the behavior that the vehicle performs which brings those traits about? Consider a spider's web. Dawkins wants to talk of a gene for a web. A web, of course, is not a characteristic of a spider. Apparently, however, we could talk of a gene for web building. Web building is a trait of spiders, and, if we choose to redescribe the phenomena in these terms, the extended phenotype is brought closer to home. We now have a trait of the vehicle in which the genes reside, and we can tell an orthodox story about natural selection for this trait.

It would be tempting to reply to this objection by stressing that the selective force acts through the artifact. The causal chain from the gene to the web is complex and indirect; the behavior is only a part of

it. Only one element of the chain is distinguished, the endpoint, the web itself, and that is because, independently of what has gone on earlier, provided that the web is in place, the enhancement of the replication chances of the underlying allele will ensue. But this reply is exactly parallel to the Mayr-Gould-Brandon argument discussed in the last section, and it should be rejected for exactly parallel reasons.

The correct response, we believe, is to take Dawkins at his word when he insists on the possibility of a number of different ways of looking at the same selective processes. Dawkins's two main treatments of natural selection, SG and EP, offer distinct versions of the thesis of genic selectionism. In the earlier discussion (and occasionally in the later) the thesis is that, for any selection process, there is a uniquely correct representation of that process, a representation which captures the causal structure of the process, and this representation attributes causal efficacy to genic properties. In EP, especially in chapters 1 and 13, Dawkins proposes a weaker version of the thesis, to the effect that there are often alternative, equally adequate representations of selection processes and that, for any selection process, there is a maximally adequate representation which attributes causal efficacy to genic properties. We shall call the strong (early) version monist genic selectionism and the weak (later) version pluralist genic selectionism. We believe that the monist version is faulty but that the pluralist thesis is defensible.

In presenting the "extended phenotype" cases, Dawkins is offering an alternative representation of processes that individualists can redescribe in their own preferred terms by adopting the strategy illustrated in our discussion of spider webs. Instead of talking of genes for webs and their selective advantages, it is possible to discuss the case in terms of the benefits that accrue to spiders who have a disposition to engage in web building. There is no privileged way to segment the causal chain and isolate the (really) real causal story. As we noted two paragraphs back, the analog of the Mayr-Gould-Brandon argument for the priority of those properties which are most directly connected with survival and reproduction—here the webs themselves—is fallacious. Equally, it is fallacious to insist that the causal story must be told by focusing on traits of individuals which contribute to the reproductive success of those individuals. We are left with the general thesis of pluralism: there are alternative, maximally adequate representations of the causal structure of the selection process. Add to this Dawkins's claim that one can always find a way to achieve a representation in terms of the causal efficacy of genic properties, and we have pluralist genic selectionism.

Pluralism of the kind we espouse has affinities with some traditional views in the philosophy of science. Specifically, our approach is instrumentalist, not of course in denying the existence of entities like genes, but in opposing the idea that natural selection is a force that acts on some determinate target, such as the genotype or the phenotype. Monists err, we believe, in claiming that selection processes must be described in a particular way, and their error involves them in positing entities, "targets of selection," that do not exist.

Another way to understand our pluralism is to connect it with conventionalist approaches to space-time theories. Just as conventionalists have insisted that there are alternative accounts of the phenomena which meet all our methodological desiderata, so too we maintain that selection processes can usually be treated, equally adequately, from more than one point of view. The virtue of the genic point of view, on the pluralist account, is not that it alone gets the causal structure right but that it is always available.

What is the rival position? Well, it cannot be the thesis that the only adequate representations are those in terms of individual traits which promote the reproductive success of their bearers, because there are instances in which no such representation is available (outlaws) and instances in which the representation is (at best) heuristically misleading (quasi-outlaws, altruism). The sensible rival position is that there is a hierarchy of selection processes: some cases are aptly represented in terms of genic selection, some in terms of individual selection, some in terms of group selection, and some (maybe) in terms of species selection. Hierarchical monism claims that, for any selection process, there is a unique level of the hierarchy such that only representations that depict selection as acting at that level are maximally adequate. (Intuitively, representations that see selection as acting at other levels get the causal structure wrong.) Hierarchical monism differs from pluralist genic selectionism in an interesting way: whereas the pluralist insists that, for any process, there are many adequate representations, one of which will always be a genic representation, the hierarchical monist maintains that for each process there is just one kind of adequate representation, but that processes are diverse in the kinds of representation they demand. 14

¹⁴ In defending pluralism, we are very close to the views expressed by Maynard Smith in "How To Model Evolution." Indeed, we would like to think that Maynard Smith's article and the present essay complement one another in a number of respects. In particular, as Maynard Smith explicitly notes, "recommending a plurality of models of the same process" contrasts with the view (defended by Gould

Just as the simple orthodoxy of individualism is ambushed by outlaws and their kin, so too hierarchical monism is entangled in spider webs. In the "extended phenotype" cases, Dawkins shows that there are genic representations of selection processes which can be no more adequately illuminated from alternative perspectives. Since we believe that there is no compelling reason to deny the legitimacy of the individualist redescription in terms of web-building behavior (or dispositions to such behavior), we conclude that Dawkins should be taken at face value: just as we can adopt different perspectives on a Necker cube, so too we can look at the workings of selection in different ways (EP ch. 1).

In previous sections, we have tried to show how genic representations are available in cases that have previously been viewed as troublesome. To complete the defense of genic selectionism, we would need to extend our survey of problematic examples. But the general strategy should be evident. Faced with processes that others see in terms of group selection or species selection, genic selectionists will first try to achieve an individualist representation and then apply the ideas we have developed from Dawkins to make the translation to genic terms.

Pluralist genic selectionists recommend that practicing biologists take advantage of the full range of strategies for representing the workings of selection. The chief merit of Dawkinspeak is its generality. Whereas the individualist perspective may sometimes break down, the gene's eye view is apparently always available. Moreover, as illustrated by the treatment of inclusive fitness, adopting it may sometimes help us to avoid errors and confusions. Thinking of selection in terms of the devices, sometimes highly indirect, through which genes lever themselves into future generations may also suggest new approaches to familiar problems.

But are there drawbacks? Yes. The principal purpose of the early sections of this paper was to extend some of the ideas of genic selectionism to respond to concerns that are deep and important. Without an adequate rethinking of the concepts of population and of environment, genic representations will fail to capture processes that involve genic interactions or epigenetic constraints. Genic selectionism can easily slide into naive adaptationism as one comes to

and by Sober) of "emphasizing a plurality of processes." Gould's views are clearly expressed in "Is A New and General Theory of Evolution Emerging?" *Paleobiology*, vi (1980): 119–130; and Sober's ideas are presented in NS ch. 9.

credit the individual alleles with powers that enable them to operate independently of one another. The move from the "genes for P" locution to the claim that selection can fashion P independently of other traits of the organism is perennially tempting. But, in our version, genic representations must be constructed in full recognition of the possibilities for constraints in gene-environment coevolution. The dangers of genic selectionism, illustrated in some of Dawkins's own writings, are that the commitment to the complexity of the allelic environment is forgotten in practice. In defending the genic approach against important objections, we have been trying to make this commitment explicit, and thus to exhibit both the potential and the demands of correct Dawkinspeak. The return of the gene should not mean the exile of the organism. 16

KIM STERELNY

Victoria University/Wellington, New Zealand

PHILIP KITCHER

University of California/San Diego

¹¹⁶ As, we believe, Dawkins himself appreciates. See the last chapter of EP, especially his reaction to the claim that "Richard Dawkins has rediscovered the organism" (251).

¹⁵ At least one of us believes that the claims of the present paper are perfectly compatible with the critique of adaptationism developed in Gould and Lewontin, "The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme," in Sober, ed., *Conceptual Problems in Evolutionary Biology* (Cambridge: MIT, 1984). For discussion of the difficulties with adaptationism, see Kitcher, *Vaulting Ambition*, ch. 7; and "Why Not The Best?" in Dupre, *op.cit*.