TWO WAYS OF THINKING ABOUT FITNESS AND NATURAL SELECTION

Mohan Matthen The University of British Columbia

and

André Ariew The University of Rhode Island

I. VERNACULAR FITNESS AND PREDICTIVE FITNESS

"The concept of fitness is," Philip Kitcher says, "important both to informal presentations of evolutionary theory and to the mathematical formulations of [population genetics]."¹ He is absolutely right. The difficulty is to harmonize these very different ways of understanding its role. In this paper, we examine how natural selection relates to the other explanatory factors invoked by evolutionary theory. We argue that the "informal presentations" to which Kitcher alludes give an incoherent account of the relation. A more appropriate model is drawn from the statistical conceptual framework of population genetics. We argue that this model demands a far-reaching revision of some widely accepted notions of causal relations in evolution.

The two formulations of fitness are well illustrated by expositions given by Kitcher

1

Kitcher, *Vaulting Ambition: Sociobiology and the Quest for Human Nature* (Cambridge Mass: MIT Press, 1984) Bradford Books: 50. Elliott Sober makes a similar distinction: "Fitness is both an ecological descriptor and a mathematical predictor." ("The Two Faces of Fitness," in R. Singh, D. Paul, C. Krimbas, and J. Beatty (eds) *Thinking about Evolution: Historical, Philosophical and Political Perspectives* (Cambridge: Cambridge University Press, forthcoming). Kitcher and (even more so) Sober imply that the underlying logic of these concepts is somehow the same, though Sober acknowledges persistent difficulties drawing them together.

himself. Consider:

The Principle of Variation in Fitness Organisms differ in ways that affect their competitive abilities. Some organisms have characteristics that better enable them to survive and reproduce than others. (38)

An organism's ability to "survive and reproduce" arises from its traits. To the extent that relatively "advantageous" traits can be inherited by an organism's descendants, they will be reproduced and retained in the population at a higher rate than less optimal ones. Consequently, these traits will spread faster than less optimal ones. This much is obvious to common-sense analysis, and for many this notion of an organism's *overall competitive advantage traceable to heritable traits* is at the heart of the theory of natural selection. Recognizing this, we'll call this measure of an organism's evolutionary potential its *vernacular fitness*. According to one standard way of understanding natural selection, vernacular fitness—or rather the variation thereof—is a *cause* of evolutionary change. Note, however, that, as formulated above, vernacular fitness is merely a *comparative* measure, not a *quantity*, and that principles such as the above afford us no way of predicting or explaining the *magnitude* of evolutionary change. This is why it is appropriate to call it "informal."

As Kitcher observes, this is not the only context in which the concept of fitness crops up in evolutionary theory. Fitness occurs also in the equations of population genetics that predict, with some level of probability, the frequency with which a gene occurs in a population in generation n+1 given its frequency in generation n. In population genetics, *predictive fitness* (as we shall call it) is an *expected* rate of increase (normalized relative to others) of a gene, a trait, or an organism's representation in future generations.² Unlike vernacular fitness, predictive fitness is a quantitative measure, not merely a number. Moreover, it is not a cause of selection, or of evolution for that matter. The expected rate of return on an investment is not the *cause* of its growth—it just *is* its growth, numerically estimated and projected into the future. In the same

There are many alternative ways of defining predictive fitness. Sober's "Two Faces of Fitness" gives reasons to be cautious about "expected values" as predictors. We ignore such subtleties in what follows: our preoccupation is with statistical conceptions as a group, as against causal ones.

way, predictive fitness is a measure of growth, not a cause.

Now, what is the relationship between vernacular fitness and predictive fitness? According to conventional wisdom, vernacular fitness is a cause of evolution. Since predictive fitness measures evolution, vernacular fitness might be thought to be a constitutive cause of predictive fitness, a part of its biological *why*, much as the shape, composition, and symmetry of a coin are physical causes of its being "fair," this being (like predictive fitness) a statistical characterization of its behaviour when tossed. On the other hand, vernacular fitness is not, so it is argued, the *only* thing responsible for evolution. Other factors such as the underlying mechanism of inheritance and drift must be taken into account, as well as various constraints on evolutionary change. These supposedly "non-selective" factors need to be combined with vernacular fitness in any adequate account of evolution. How are they to be combined? Any attempt to harmonize Kitcher's "informal" and "mathematical" discourses needs to come to grips with this question.

Some suggest that the contribution that vernacular fitness makes to evolution can be understood by means of an analogy with the standard physical-science conception of multiple forces summed up to produce a resultant total force. This way of addressing the harmonization problem is one of the "two ways of thinking" to which our title alludes. In sections II-IV, we take issue with this analogy. We argue that it makes no sense to *add* natural selection to other evolutionary factors in this way; the combination of factors is better described by the *hierarchical realization* model presented in sections V-VII. We argue further, in section VIII, that natural selection is not a process driven by various evolutionary factors; rather, it is a statistical "trend" with these factors (vernacular fitness excluded) as predictors.³

II. THE THEORY OF EVOLUTION AS A THEORY OF FORCES

Elliott Sober is the author of one influential suggestion concerning how the theory of evolution

The distinction between physical process and statistical trend parallels a distinction between dynamical and statistical analyses in Denis Walsh, Tim Lewens, and André Ariew, "The Trials of Life: Natural Selection and Random Drift" (in preparation).

accommodates natural selection alongside drift, developmental constraints, architectural constraints, and other such determinants of evolutionary change.⁴ He proposes that we should think of the theory of evolution as a theory of forces along the lines suggested by Newton's mechanics (31), and urges that the following structural elements of this theory be incorporated into the theory of evolution.

1. *A zero-force law* Such a law equates the steady-state, or no-change, condition to no net force. In classical mechanics, a body is not changing if it continues in its state of rest or steady motion. Newton's first law of motion stipulates that a body is in this steady state exactly when no force is impressed upon it. Sober claims (32) that the corresponding steady state in evolution is specified by the Hardy-Weinberg law of population genetics. This is not exactly right, how-ever—the Hardy-Weinberg law tells us about genotype frequencies *in sexually reproducing populations* when there is no change of gene frequencies within them. The most general way of capturing Sober's intentions is to stipulate that the steady state occurs when gene frequencies do not change from one generation to the next. Following the Newtonian analogy, he concludes that there is no net force acting on a population in this condition. This would be the zero force law for evolutionary theory.

2. *Decomposition of force* Newtonian physics envisages several sources of forceCgravitation, electrostatic force, magnetic force, impact, etc. In classical physics, "source laws," as Sober calls them, describe how these forces emerge, one by one, from various physical situations. The

Sober, *The Nature of Selection: Evolutionary Theory in Philosophical Focus* (Cambridge MA: Bradford Books, MIT Press, 1984), in chapter 1. Since a large part of this paper is devoted to taking issue with Sober, we want to emphasize that we, like many other philosophers, learned much of what we know about natural selection from Sober's classic work. We should also like to acknowledge Sober's patient discussion of the issues dealt with in this paper, and our own personal debt to his influence.

law of gravitation, Coulomb's law, the law of elastic bodies, etc., are examples of such source laws. The total force acting on a body (the force that feeds into the consequence laws described under 4, below) can always be decomposed into these components. Sober suggests (38) that selection, mutation, migration, random genetic drift, linkage, and in-breeding are the individual forces that act on a population and cause it to evolve. Laws governing these are source laws: for example, the principles of optimality analysis that predict the selective value of various traits and the laws that govern how genes located on the same chromosome assort during recombination. Causal responsibility for total evolutionary change can apportioned among these factors.

3. *Resolution of forces* Given two or more forces acting on a body, Newtonian physics uses the parallelogram of forces (that is, vector addition) to compute the total resultant force acting on the body. Conversely, every force has a unique correct decomposition into *real* component forces: such a decomposition may not always be epistemically feasible, but it does have ontological reality. Evolutionary forces combine in this way, according to Sober. He offers the case of *heterosis* to illustrate how selection combines with genetic recombination (47). (Our own discussion of this phenomenon will be found in section V.)

4. *A law describing change under resultant force* Newton's second law of motion tells us that the change of a body's momentum is proportionate to the total resultant force impressed on it. Sober refers to this as a "consequence law." The laws of population genetics are supposed to play this role in evolutionary theory.

In Newton's theory, force is the common currency of the different influences on a body: gravitational, electrostatic, and other influences on a body are expressed in terms of force, and any number of forces can be combined into a single resultant force, regardless of their source. The second law of motion (F = ma) predicts change given this net force. In Sober's account of evolutionary theory, source laws tell us where fitness differences, constraints, etc. come from, consequence laws tell us how populations respond to these fitness differences, constraints, etc.

acting together. Thus, fitness values are the outputs of source laws and the input of consequence laws. If Sober is right, then, fitness plays a role in evolutionary theory analogous to that of force in Newtonian mechanics. But the fitness that emerges from the source laws is vernacular fitness, since these laws are concerned with the competitive advantage that organisms get from their traits in different environments. On the other hand, the fitness that feeds into consequence laws is predictive fitness, for as Sober says, "the consequence laws for natural selection are preeminently the province of population genetics" (59). So, in order to maintain the analogy with a theory of forces, the two conceptions need to be made commensurate. But, as we shall argue in the next two section, no such harmonization is possible. Consequently, the first three structural elements mentioned above find no counterpart in evolutionary theory, so we shall argue, and the fourth, the notion of a consequence law, makes sense only in terms of predictive fitness.

III. DRIFT AND THE ZERO FORCE LAW

The trouble starts with the zero force law. In Newton's theory, inertiaCthe property of each massive body to resist changes to its "state of rest or uniform motion along a straight line" (*Principia*, Definition III)Censures that if there is no force acting on a body, then there is no change in its momentum. But *prima facie* at least, this is not the case in evolution. For there are chance variations in the genetic composition of a population from generation to generation. These changes can have a dramatic cumulative effect. According to the theory of the "random walk," an entity that is varying under chance influences will drift further and further away from its starting condition, the cumulative deviation being proportionate to elapsed time. It follows that the frequency of a given gene will, with time, drift further and further away from its starting frequency, and given enough time, it will approach a frequency of either zero or one as chance has it. But once it reaches zero or one, the change is (barring regeneration by mutation) irreversible: either the trait has gone extinct or it has reached "fixation." Thus, given a set of alleles, one will eventually go to fixation even without the influence of fitness differences. But this creates a problem for the zero-force law. In order to save it, Sober claims that this kind of

change is caused by "drift," which he therefore counts among the "forces" of evolution.

In evolutionary theory as in Newtonian physics, the principal use of a zero-force law is to discover when evolutionary forces *have* played a role. If genotype frequencies depart from . . . equilibrium, some force must have been at work. . . . "[S]ampling error" (*random genetic drift*) may lead the frequency of a gene among the zygotes to differ from the frequency of that gene among the gamete pool. (34)

But does it really makes sense to say that drift is a *force* or, more generally, a cause of change that acts independently of selection? Consider this analogy. You toss a coin four times. What would explain the outcome *two heads*? Answer: the physical setup of the coin-tossing trials. What would explain the outcome *four heads*? The same thing: that is, the same physical setup. Though the first result is less probable, the same setup explains both outcomes equally. The set of possible worlds in which fair coins are tossed four times form what Wesley Salmon⁵ calls a *homogeneous reference class*; there is no way to partition such a reference class by an independent relevant factor such that the probability of four heads is greater in one of the cells of the partition.⁶ Perhaps one might say that the four-head outcome was less predictable, therefore less well explained. Nevertheless, there is nothing that one can invoke to strengthen the explanation of the four-head outcome. If the physical setup is the only thing relevant to the two-head outcome, then nothing else is available to explain the four-head outcome. In particular, you cannot partition the homogeneous reference class to which these trials belong by *improbability* or *chance*, and so you cannot tag these things as "forces" which occasionally favour the less likely outcome.

5

Salmon, *Scientific Explanation and the Causal Structure of the World* (Princeton: Princeton University Press, 1984): 36-37.

6

That is: if *C* is the condition that defines the homogeneous reference class to which this trial belongs, then the probability of the target explanandum *four heads* is the same (1/16) given *C* as it is given *C* and any other putative factor, *F*.

For closely related reasons, "drift" should not be regarded as a force that can be added to others acting on a population. (In any case, drift isn't the sort of thing that can play the role of a forceCit doesn't have predictable and constant direction.) Imagine allelic alternative traits T and T', such that T confers more causal fitness on its bearer. Suppose that T and T' compete in two similar populations of conspecific organisms, subject to the same pressures of selection. In one population, T goes to fixation, in the other, T' does. What explains this difference? The answer, just as in the case of the coin, might well be: *exactly the same thing*, the same positioning of the traits in the adaptive landscape. Maybe the explanation is weaker in the case of T' going to fixation, particularly if the fitness difference is large. But it does not follow that any other explanatory factor is available to help out. So one cannot say that "if genotype frequencies depart from equilibrium, some force must have been at work." Because the causes here are probabilistic, change might have the same cause as equilibrium.

Does this miss the point? Sober says (see above): "In evolutionary theory as in Newtonian physics, the principal use of a zero-force law is to discover when evolutionary forces *have* played a role." And this might be taken to mean that one can retrospectively identify drift in particular evolutionary histories, not simply in stochastic aggregates. Histories of natural selection consist, after all, of collections of concrete individual events: births, deaths, matings, mutations, etc. Some of these events are predictable on the basis of advantageous traits (that is, vernacular fitness), others are not. In the first kind of case, we have, as some say, "discriminate" sampling, in the second, "indiscriminate." And so one might think that one can fix the exact role of vernacular fitness and of drift by looking more minutely at individual events and determining when discriminate sampling has been at work, and when indiscriminate. In this way, one apportions the outcome respectively to fitness and drift.

Consider, then, a case like this: two organisms, *O1* and *O2*, otherwise very similar, differ in (vernacular) fitness because *O1* has better eye-sight than *O2*. Now, contrast the following possible events.

C1 O2's bad eye-sight leads to its falling off a cliff. It dies and O1 survives.

C2 O1 is killed by a lightning strike—the difference of visual acuity was irrelevant to this event.

There is evolutionary change in both these cases, because both lead to a change in the genefrequency of the population: in C1, O1's visual acuity genes get a slight boost, in C2, they suffer a small setback. C1, however, seems to be a case in which the difference of vernacular fitness (the difference in eyesight) contributed to evolution, and C2 one in which a chance event thwarted the fitness difference that drives natural selection. So it may be tempting to say that natural selection is the cause of evolution in C1, and that it consists, over a longer period of time, of "predictable" (or fitness-biassed) cases like C1, but that it excludes anomalous (or fitnessindiscriminate) cases like C2. Then it might be thought plausible to say that something else—drift? neutral selection?—is operating in C2.

But this violates sound probabilistic thinking. Probability enters into the picture because the theory of evolution abstracts away from individual concrete events like C1 and C2.⁷ This process of abstraction is an attempt to isolate the causal factors that make a predictable (but probabilistic) *difference* to evolution. Factors such as lightning strikes do not appear in evolutionary theory because they are neutral—*ex hypothesi* your chances of being killed by a lightning strike are the same whether you have good eyesight or bad. (See note 6.) Because it is

7

There are several kinds of situations in which the notion of probabilistic causation is invoked. One is in quantum mechanics, where it is claimed that the interaction of certain variables is irreducibly indeterministic. Thus if QM variable X brings about result Y with probability p%(less than unity), it is unclear whether one should say that X is 100% involved in p% of the cases, thus bringing about Y, and inert in the rest, or rather that X is p% involved in 100% of the cases, and brings about Y in p% of the cases as a result. But let us leave this example aside, along with the complications it brings. Indeterminism is not the issue in the cases that we are discussing. As with the fair coin, one would not be justified in claiming that the individual events above contained elements of indeterminacy.

neutral, the lightning strike factor cannot be conjoined with others to sub-divide reference classes, and thus yield finer predictions. This does not mean, however, that evolutionary theory *ignores* lightning strikes, for they are taken into account when the set of possibilities is partitioned by reference to the factors that make a difference. Proneness to be killed by lightning does not appear in the formulae that define the partition because they occur with equal frequency in each and every cell of the partition. In other words, there is no variance among individuals in the population with respect to proneness to being killed by lightning. Thus, cases like *C2* have already been taken into account when we say that bad eyesight etc. causes a greater rate of death *with some level of probability*.

As long as we are dealing with the causal factors that make a probabilistic difference to evolutionary outcomes, then, we have to resist the temptation to say that the two cases above are distinguishable in terms of relevant factors. For it is only by a kind of equivocation that we can say that lightning strikes are not a factor in natural selection, but are the cause of evolutionary change in cases like C2.⁸ Lightning strikes do not, over large populations, over the long haul, result in the differential retention of heritable traits; in this sense, they are not factors in natural selection. On the other hand, any death results in evolutionary change since it results in a change of gene frequencies; in this sense the lightning strike caused evolutionary change. This pair of statements seem contradictory when the equivocation in 'cause' goes unnoticed. It is precisely because it goes unnoticed that some say that in cases like the one described above evolution has to be explained by a factor operating outside natural selection.

This equivocation highlights a confusion of levels that can occur when we speak of probabilistic causes. Suppose that good eyesight causes individuals like X to live n years longer with probability m. This statement has something to do with proportions of X-like individuals who live n years longer in various possible worlds. That is, it is an *ensemble-level* or "e-fact." Turning now to concrete individual or "i-facts," suppose that a particular individual X has good eyesight and lives a long life. To what degree did eyesight contribute to *her* long life, and how

We are indebted here to Andrew Irvine and Joel Pust for helpful discussion.

much should we attribute to good luck? The question is ill-posed. When one is talking about *types* of events, it makes sense to quantify the role of chance. For once a reference class has been partitioned in terms of all the factors that make a difference, the residual variation within the cells of the partition—the unassigned variation—is uncaused. It is due to chance if you like. However, unassigned variation has no **i**-counterpart. Though it is six times more probable that two heads will turn up in a run of four tosses of a coin than that four will, chance does not play any more of a role in a particular run of four heads than in a particular run of two and two. Thus, one cannot differentiate between individual events on the basis of how much they are attributable to chance.

For exactly the same reason, one cannot, even retrospectively, determine whether discriminate or indiscriminate sampling was responsible for an actual sequence of historical events. In an example discussed above, T and T' were traits which suffered opposite fates in two closely comparable populations. We can certainly estimate the probability that T will win or that T' will, counting these as event types. We know, for instance, that the chances of the less viable trait going to fixation are greater if the population is smaller-this application of the Law of Large Numbers is the stochastic basis for the Founder Effect, that is, for novelty arising when small populations are isolated. These statements are based on the assignment of unassigned variation to event-types. There is, however, no such thing as unassigned variation with respect to *concrete* sequences of events. Suppose then that one is trying to explain a particular biological characteristic, for instance, the absence of body hair in humans. It would be correct to say that this characteristic probably arose in a small population, where novelty had a better chance of taking hold-but only if this is meant as a statement about types to which we revert because we lack certainty about evolutionary history. It would be a mistake, however, to ask how much chance contributed to the actual history of human hairlessness, just as it would be a mistake to try to apportion the role of chance in a particular series of coin-tosses.⁹ The actual history of a

Alexander Rosenberg, *Instrumental Biology or The Disunity of Science* (Chicago: University of Chicago Press, 1994) has recently argued that drift enters into evolutionary theory only because we are ignorant of particular causes. He is right if he is talking about historical reconstructions of

trait is an *i*-history and probability does not translate downwards from the *e*-level.

So the theory of evolutionary forces has a problem. In the passage quoted above, Sober rightly demands that one ought to be able to tell when a force is acting on a system. He puts the point in epistemological terms, but one can rephrase it ontologically. *Whether there is a force acting on a system should be a matter of fact.* This, however, is precisely the question that is not well-defined in this instance. Suppose that over a period of time a population stays exactly the same, or changes in some determinate way. It makes no sense to ask whether or not drift was implicated in these individual-level outcomes.

IV. SEPARATING EVOLUTIONARY FORCES

Continuing our examination of evolutionary theory as a theory of forces, we now ask how one should go about decomposing evolutionary force. Here it is useful to introduce another context in which the notion of component force has sometimes been used. Stephen Jay Gould and Richard Lewontin are associated with the claim that to appeal to natural selection alone as an explanation of evolutionary change, and to underestimate the other factors, is a methodological error to which they gave the name "adaptationism."¹⁰ They define this error as the belief that "natural selection [is] so powerful and the constraints upon it so few that the direct production of adaptation through its operation becomes the primary cause of nearly all organic form, function, and behaviour" (256).

evolutionary scenarios. But he is wrong if he is talking about evolutionary *theory*, which abstracts away from individual causes.

10

Stephen Jay Gould and Richard Lewontin, "The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme," *Proceedings of the Royal Society of London* 205 (1978): 581-98. Reprinted in Sober *Conceptual Issues in Evolutionary Theory,* (Cambridge MA: Bradford Books, MIT Press, 1984). Page references are to the latter version.

For the sake of the present argument, let us restate their critique in terms appropriate to a theory of force. (At the end of section VII, we'll consider another way of stating it.)

*Adaptationism*_{force} is the view that natural selection is unopposed, or only weakly opposed, by other forces of evolution, with the consequence that it always achieves the optimal result. This view is mistaken because in fact natural selection is strongly opposed, indeed often thwarted, by these other forces, in particular by genetic, architectural, and developmental constraints.

This invites the following question in light of item 2 in the characterization of theories of force in section II: what would happen if selection were to act *by itself*? For in order to estimate how much of an evolutionary outcome should be ascribed to non-selective constraints, and how much to natural selection,¹¹ we need first to appreciate the action of the latter acting alone.¹² In the context of the debate about adaptationism, it is clear that the supposed action of natural selection as a single factor is construed in terms of optima. The idea is that if natural selection were to act on its own, it would achieve optima. Evolution does not always produce optima, however, because natural selection is opposed by constraints.

But it is completely inappropriate to conceptualize the "power" of natural selection by specifying an optimum and asking how many obstacles it is able to overcome in achieving it, or

11

12

Here, we follow Gould and Lewontin in making *natural selection* the focus of the debate about adaptationism; we could, however, have dealt with it in terms of differences in vernacular fitness as driving natural selection and being opposed by constraints. There is no discontinuity, then, between this section and the previous discussion.

Note that unlike drift, which has no fixed "direction," these constraints act consistently and predictably. Thus the argument of the last section, to the conclusion that drift cannot be isolated as a separate component of evolutionary change, has no application to constraints.

attempting to calculate the *speed* with which natural selection would achieve an optimum acting "by itself." In population genetics, issues of speed and of overcoming constraint are addressed in the context of an underlying chromosomal structure; in evolutionary systematics, they are considered in the context of pre-existing body-plans; in studies of development, they are posed as questions of heterochrony (changes of developmental sequences). In short, it is clear that natural selection acts in certain underlying causal media, and the the so-called constraints are features of these media. Since natural selection cannot act without such a medium, it is not at all clear what sense can be made of the idea of natural selection acting "without the intrusion of constraints." (We'll return to this point in the next section.) If this is right, we have no way of making sense in this context of the Newtonian idea that if an effect is to be analysed in terms of two forces acting together, then a vector value has to be assigned to each acting independently.

The *resolution of forces* condition also fails in general in evolutionary theory. In Newton's theory, force is specifiable in absolute terms. Thus any two Newtonian forces are comparable. For example, the force produced by gravitation is comparable to that produced by electrostatic attraction: they are expressible in the same units as each other and as the force mentioned in Newton's second law of motion. It is this commensurability that allows us to estimate, for example, how an oil-drop behaves when it is subject to both gravitational and electrostatic force. Coulomb's law gives us a value for electrostatic force, the law of gravitation gives us a value for gravitational force; each of these forces would act on its own if the other were absent. The law of addition of forces gives us a value for *total* force, and the latter feeds into the consequence law.

By contrast, vernacular fitness is specified in comparative terms and in restricted contexts. For example, you might learn that the optimal reproductive strategy with respect to sexdetermination is to produce male offspring when there are fewer males in the population, and females when there are fewer females. But this only tells you about the relative merits of strategies within a circumscribed set, with other factors held constant. The analysis does not tell you whether producing offspring of the minority sex is more or less advantageous than other fitness-relevant things you can do; there is, generally speaking, no way of combining the effects of a good strategy in this game, with good or bad strategies in other games. For example, we

have no way of calculating whether a given sex-selection strategy interacts with a given parental care strategy, and how the fitness produced by variants of these strategies combine. This inability to add the "forces" of fitness is even more pronounced when the source laws are in unrelated domains. Suppose a certain species undertakes parental care, is resistant to malaria, and is somewhat weak but very quick. How do these fitness factors add up? We have no idea at all. The theory of probability has no general way to deal with such questions.¹³

Since we cannot add up the values produced by the source laws, the occurrences of fitness in the laws of population genetics are conceptually independent of their occurrence in source laws in a way not parallelled in the Newtonian treatment of force. The overall fitness values demanded by consequence laws must be estimated statistically, i.e., by looking at actual values for number of offspring, and using these actual values to estimate expected values and other statistical quantities. This independence has been remarked on by more than one philosopher of biology. "Fitness must be measured by its effects," says Alexander Rosenberg.¹⁴ However, the independence of consequence laws has often been misdiagnosed. Rosenberg thinks that it arises because there are multiple physical realizations of fitness. Sober takes a related line: "The supervenience of fitness—the fact that fitness is not a single physical property—helps explain why *general* source laws are hard to come by."¹⁵ But the multiplicity of realizations, whether

13

Donald Davidson makes structurally the same point in "How is the Weakness of the Will Possible?" (In Joel Feinberg (ed.) *Moral Concepts* (Oxford: Clarendon Press, *Oxford Readings in Philosophy*, 1971): 93-113. He argues (108-111) that there is no systematic way of comparing the negative value of brushing one's teeth given that one is sleepy with the positive value of doing so given that one values healthy teeth. And it follows that there is no systematic way of constructing the "all things considered" judgement on whether one should brush one's teeth.

14

Rosenberg, "Fitness," Journal of Philosophy 80 (1983): 457-73, at 459.

Sober, The Nature of Selection, 51.

with regard to source or consequence laws, does not mark a disanalogy between the physics and evolutionary biology: Newtonian force has multiple physical bases too—charge, mass, collisions, etc. The disanalogy is that while force affords Newtonian mechanics the means to compare and add up the consequences of these diverse causes, fitness does not add up or resolve. This is why population genetics is forced to estimate fitness by measuring population change.

V. SELECTION AND ITS SUBSTRATES

We have now argued in support of three propositions. First, there can be no objective, as opposed to epistemic, apportioning of causal responsibility to selection as against drift in a concrete evolutionary history. Consequently, evolutionary theory has no zero-force law. Second, it makes no sense to think of natural selection acting "on its own." So, the decomposition-of-forces condition fails. Third, there is no common currency in which to compare the contributions of different "forces." Thus, the resolution of forces does not have any counterpart in evolutionary theory. We conclude that it is inappropriate to think of evolution in terms of separate additive forces. How then shall we think of drift in relation to fitness? How shall we frame the role of constraints?

We remarked earlier that it makes no sense to try to determine the effects of natural selection acting independently of underlying causal determinants—chromosomal structures, mode of reproduction, pre-existing body plans, developmental sequences, and so on. Instead of thinking of these selection as acting independently of these factors, either competing with them or being reinforced by them, we propose to regard selection as a global characteristic of various "substrates." In the hierarchical realization model that we shall now begin to construct, selection is not, physically or biologically, the same force or cause wherever it occurs, but a formally characterized phenomenon, the details of which varies with the substrate. The above mentioned determinants of evolution are substrate characteristics, and so selection varies as they do.

We begin by recalling that Darwin and Wallace first hit on the principle of natural selection in an extremely general form: a heritable trait that enables an organism to reproduce more than those that lack the trait will increase its proportion in the population.¹⁶ Now, it turns

¹⁶

Actually, this idea had occurred already to Empedocles in fifth century B.C.E. Sicily; see

out that, as R.A. Fisher first argued in detail, the principle is insufficient in this general form to explain the facts of evolution.¹⁷ Fisher showed that under the system of inheritance envisaged by Darwin—"blending inheritance"Cthe variation in a population would die out too quickly to sustain large evolutionary changes, except under an implausibly high mutation rate, directed mutation of one sort or another, or Lamarckian inheritance. Mendelian inheritance, in which discrete genes are handed down unchanged from parent to off-spring, is required to maintain variation in the face of ecologically imposed homogenization. Now, think of selection not simply as an ecological phenomenon, but as a series of events which ultimately results in changes in the genetic composition of a population. Fisher's analysis demonstrates that selection so understood depends on the mode of inheritance. This is an illustration of how a supposedly "non-selective force" is *required* in order that differential ecological influences can be written into the genetic composition of a population: Mendelian inheritance is the medium through which the effects of mating, deaths, and other such events affect the genetic make-up of future generations.¹⁸

Aristotle, Physics II 8, 199a5-8.

17

Fisher, "The Genetical Theory of Natural Selection," 2nd edition, (New York: Dover, 1958), chapter 1. (The first edition was published by the Clarendon Press in Oxford, 1928.) "No reasonable assumptions could be made by which the diminution of variance due to selection, in the total absence of mutations, would be much more than a ten-thousandth of that ascribable to blending inheritance" (10).

18

There is a significant codicil here to the so-called "tautology problem." The Darwin-Wallace proposition is a mathematical truth; consequently, some have been puzzled about how it can explain evolution. The usual response is that the antecedents of this proposition are contingent: it is an empirical fact that some heritable traits give organisms a competitive advantage. Fisher's argument shows that the Darwin-Wallace principle is insufficient to explain evolution *even under the contingent stipulation that its antecedents are satisfied*. The theory of natural selection

Now, natural selection is manifested not in one, but a great variety of substrates. Biologically, it occurs in conjunction with both asexual and sexual reproduction and in the presence of varying degrees of cytoplasmic inheritance. Selection occurs at the phenotypic level, in which case effects on gene frequencies are indirect, and it also occurs directly on genes. It can be more or less determinate in its effects, that is, the probabilities involved in natural selection could approach unity, or be quite low. In possible worlds other than this one, selection may be manifested on a substrate of blending inheritance, with Lamarckian feedback, without Weismannian inviolacy of the germ line, with three or more sexes, etc. As we shall see in the next section, selection also occurs in non-biological realms-in the economic domain, for example, as well as in the non-standard domains of "clonal selection" in the mammalian immune system, in classical conditioning,¹⁹ and, according to some, in the propagation of theories²⁰ and other cultural artifacts. While the Darwin-Wallace principle continues to hold true under all of these conditions, it is silent on how events in one generation leave their mark on subsequent generations. The substrates incorporate the concrete causal conditions that make this possible. Properties of the substrate are relevant to the speed and sustainability of selection, and the manner in which it affects evolutionary change.²¹

is thus obliged to introduce considerations not mentioned in the Darwin-Wallace principle.

19

See David Hull, Rodney Langman, and Sigrid Glenn, "A General Account of Selection: Biology, Immunology, and Behavior," *Behavioral and Brain Sciences* 24 (2001): forthcoming.

20

Donald E. Campbell, "Evolutionary Epistemology," in P.A. Schilpp (ed.) *The Philosophy of Karl Popper* vol 1 (LaSalle III: Open Court, 1972): 412-63.

21

In "Teleology and the Product Analogy" (*Australasian Journal of Philosophy* 75 [1997]: 21-37), Matthen argued that the diversity of processes in these substrates undermines the claim that teleological talk of functions reduces to selection, and proposed that the latter is merely analogical. The argument is strengthened here by the contention that selection is not a single

The phenomenon of *heterosis* provides us with a vivid illustration of substrate dependance. Sexually reproducing organisms have two genes at each chromosomal locus. Homozygotes have the same gene occur twice, heterozygotes have different genes opposite one another. Now, it sometimes happens, with respect to a particular locus, that the heterozygote is fitter (in the vernacular sense) than either homozygote, a phenomenon known in the plant world as "hybrid vigour." Let's imagine an extreme case of this: the vast majority of each homozygote type are aborted immediately upon conception, the majority of heterozygotes survive to reproduce in the normal way. If so extreme a difference of viability were to occur (a) among asexually reproducing organisms, the viable phenotype would go to fixation in short order, and the others would disappear. If it occurred (b) with blending inheritance, the whole population would similarly settle into some non-lethal type intermediate between the two homozygotes. In the actual case, however, homozygotes continue to be produced despite the extreme negative selection against them. This is a consequence of the fact that (c) when two heterozygotes mate, they contribute matching (particulate) genes to some proportion of their offspring, thereby producing a homozygote. Of the three substrates just considered, heterosis is manifested only in connection with sexual reproduction with particulate inheritance. But this does not mean that the substrate is working against natural selection to preserve non-viable phenotypes. After all, selection has no way to cause a change in gene frequencies in such populations except through the underlying mechanism of inheritance, in this case (c); it cannot act "on its own."²² The retention of the heterozygote is a consequence of the way in which ecological differentiation is

type of cause.

22

As Walsh, Lewens, and Ariew argue in "The Trials of Life," the equilibrium state here is different from that of a feather that floating in the air, or of an oil-drop motionless under the combined effect of gravity and electrostatic force. In these cases, the opposed forces can be modified, resulting in a change in the overall effect—you can switch off the electrostatic force and watch the oil-drop descend. This is closely related to our point in section IV concerning the non-resolvability of resultant fitness.

imprinted on the gene pool-but there is no additive relation between selection and its substrate.

The multi-realizability of natural selection parallels that of statistical thermodynamics. For as Lawrence Sklar points out, thermodynamics has "universal applicability" across heterogeneous domains and, in marked contrast with other theories of physics, describes neither the constitution of entities nor their dynamics.²³ Like the theory of natural selection, statistical mechanics has a formal component: this is concerned with probability distributions on large ensembles of entities interacting with each other in certain ways. This formal character accounts for its applicability to heterogeneous domains. Thus, as Sklar says, a quantity like temperature is "realized in radically different physical ways" (194). It is not so much a physical variable as "a way of characterizing ensembles," "a parameter appearing in some appropriate probability distribution over the microscopic states of individual systems" (195). Taken in conjunction with the characteristics of different systems, the general principles of thermodynamics yield different distributions, or statistics. If two regions in contact are of equal temperature, then they will be in equilibrium. If they are of different temperatures, there will be heat transfer from higher to lower. The character of the equilibrium and of the heat transfer will depend on the nature of the underlying system. But it would be infelicitous, to say the least, to treat of these differences as arising from factors interfering with, or amplifying, the thermodynamical influence of temperature.²⁴ It is the same in the case of natural selection.

23

24

There are considerable complications here. Sklar says: "Various probabilistic assumptions . . . fail to have their complete grounding in either the theory of the constitution of matter or in the underlying dynamics" ("Reduction (?)," 190). Consequently, as he says elsewhere, "Probabilistic assumptions can be built into [the] theory *at its own level*, instead of bringing [them] in at the level where atomicity and micro-mechanics are introduced." (*Physics and Chance: Philosophical Issues in the Foundations of Statistical Mechanics* [New York: Cambridge University Press, 1993]: 23, our emphasis.) It has never been suggested that *sui*

Sklar, "The Reduction (?) of Thermodynamics to Statistical Mechanics," *Philosophical Studies* 95 (1999): 187-202, at 189.

VI. ABSTRACT FORMULAE OF SELECTION

If selection is realized in many substrates, and has different characteristics in these substrates, what is it in general? What do the multiple realizations share? We begin now to give consideration to the formal properties of natural selection by introducing an extremely general characterization.

Li's Theorem In a subdivided population the rate of change in the overall growth-rate is proportional to the variance in growth-rates.

Li's theorem can be made vivid by means of an economic example. Imagine that you put one hundred dollars into mutual fund A, and one hundred into B. A grows steadily at six percent per annum, B (to your chagrin) at two. Now, the average growth-rate of your total investment starts out at four percent. But as your sub-investment in A gets larger and larger relative to that in B, the representation of A in your investment account grows. This, if you like, is selection: A "reproduces" faster than B, and there are proportionately more A-dollars in the account. The consequence is that the average growth-rate of your whole investment account becomes more and more influenced by A relative to B. For example, at the end of the first year, your investment stands at \$106 in A and \$102 in B; consequently, the overall growth rate is 4.04%. At the end of 40 years, there is \$970 in A and \$216 in B, and the overall growth rate is 5.25%. Over time, the growth rate of the whole approaches the A-rate, six percent. Now, the *variance* of growth rates as between A and B was maximal at the very start of the process: as the overall mean gets closer to A and A comes to contain a greater proportion of the whole, the variance drops because the greater proportion of the account is closer to the mean. At the same time, the increase in the overall growth rate slows down. Thus the increase in the overall growth rate tracks variance.

Anthony Edwards reports that this "growth-rate theorem" was presented by C.C. Li as a

generis probabilities should be built into the theory of natural selection

"simplified version of Fisher's fundamental theorem of natural selection."²⁵ It is certainly simplified in that it does not take on board such complications as sexual reproduction, dominance, linkage, ecological change, etc., while Fisher's theorem and its successors contain parameters that sum up the effect of such factors.²⁶ Consequently, Li's theorem cannot be applied to phenotypically defined sub-populations except under special circumstances: the growth rate of a trait masks the growth rate of the underlying genes, and variance in the former will not predict changes in gene frequencies. But this is just a reappearance of the point made in the last section, namely that ecological interactions are imprinted on a population of genotypes through a substrate. Li's theorem does apply to populations of genotypes, though what is needed for the application is values for the growth of each genotype, or gene.²⁷ The theorem captures

25

A.W.F. Edwards, "The Fundamental Theorem of Natural Selection," *Biological Reviews of the Cambridge Philosophical Society* 69 (1994): 443-74 (here 444), citing C.C. Li *Population Genetics* (Chicago: University of Chicago Press, 1955). Readers are referred to Edwards for a brief, but reasonably simple, general proof of the theorem; it is, however, quite instructive to chart the example given above in a spreadsheet program.

26

See the beautiful exposition in G.R. Price, "Fisher's 'Fundamental Theorem' Made Clear," *Annals of Human Genetics* 36 (1972): 129-40. Price himself has a general mathematical characterization of natural selection in "Selection and Covariance," *Nature* 227 (1970): 520-21. Steven A. Frank has a useful comparative discussion of general approaches to selection in "The Price Equation, Fisher's Fundamental Theorem, Kin Selection, and Causal Analysis," *Evolution* 51 (1997): 1712-29.

27

Despite this predictive limitation, Li's theorem is sufficient if one is content to operate at the genic level, assuming that all relevant differences at other levels will show up there. For as Kim Sterelny and Philip Kitcher ("The Return of the Gene," *Journal of Philosophy* 85 (1988): 339-61) show, George Williams's "book-keeping argument" implies just this. The fancier

the essential nature of selection; it *defines* selection as change in a population divided by growth rates. What it says, in effect, is that the faster growing sub-types increase their representation in the whole. Moreover, the speed with which this takes place is proportionate to how much they vary from the whole. The definition, and its mathematical consequence, is substrate neutral.

(Li's theorem is an abstract expression of the effects of differential growth. Taking growth rate as a surrogate for fitness, it is possible to appreciate the meaning of Fisher's "Fundamental Theorem": a population increases in fitness proportionate to the genetic variance in fitness among sub-populations. Note that this does not tell us much about *vernacular* fitness, that is, about the causes of growth in a sub-population, or about progress.²⁸ In fact, the theorem does not predict that any given type will grow faster in time. It just tells us that the type that grows fastest will contribute most to the mean growth rate of the population.²⁹)

In calculating the effect of this "selection" pressure on the investment account, we were dealing with a very simple case. There is no differentiation between genotype and phenotype:

formulations of Fisher and Price add nothing to our understanding of the essential character of selection.

28

Fisher himself was a eugenicist for whom the vernacular associations of the concept of fitness resonated too loudly. He worried (as late as the revised post-war edition) that the fittest people—those from the upper social classes, according to him—were making themselves less fit by limiting the size of their families.

29

It would be a mistake to that Li's theorem (or Fisher's) depends on the growth rate definition of *fitness*. They are neutral as to whether fitness is to be defined as the "per capita rate of increase" (as Fisher understood it) or as "expected contribution of the type to the next generation" (as Sewall Wright did). (Sahotra Sarkar, personal communication to AA.) In fact, one of us (AA) argues that the Fisherian definition is misleading as a definition of statistical fitness. (Ariew and Lewontin, in preparation)

selection occurs directly on the kind of item being counted. Indeed, there is no "inheritance" here as such; the accumulation and continuity of "dollars" serves to preserve the effects of differential growth rates. When we deal with real biological cases, there will be effects that Li's theorem does not predict, just as there are substrate-related effects not predicted by the Darwin-Wallace formula. The "force" conception of evolution attempts to accommodate these additional effects by positing additional influences over and above bare selection. This is a misconception. Li's theorem tells us nothing about causes of growth: it is a general truth about growth regardless of how it is caused. Bare selection has nothing to do with "the constitution of actual entities and their dynamics." It is not a cause or a force.

VII. A HIERARCHICAL REALIZATION SCHEME FOR SELECTION

We propose to accommodate substrate influences by conjoining various additional conditions to a basic defining formula of selection—we'll put Li's theorem in that role, though there are other candidates such as Fisher's theorem, Price's equation, and Richard Michod's "Darwinian Dynamic."³⁰ The difference between the proposed approach and that of a theory of force is that a conjunction of propositional conditions does not imply, as the force analogy does, the existence of separate additive causal influences corresponding to each propositional conjunct.

Michod *Darwinian Dynamics: Evolutionary Transitions in Fitness and Individuality* (Princeton: Princeton University Press, 1999).

Accordingly, we now introduce the notion of a natural selection *formula*.

A natural selection *formula* is one of the form (L & C), where L is the antecedent of Li's growth rate theorem—L posits a population subdivided by growth rates—and C is a *substrate specification* which states the causes and/or effects of differential growth rates of particular populations and their parts under specified conditions of inheritance, development, and environmental interaction.

There is a great variety of natural selection formulae, as the following examples demonstrate. (Li's theorem is taken an implicit conjunct in each.) The Darwin-Wallace principle belongs at a very abstract level—"abstract" because the substrate specification is minimal—but less so than Li's because it is restricted to the biological domain (though Darwin was himself influenced by economic parallels). The substrate specification here is that the source of differential growth rates is variance of adaptedness, that is, of heritable traits which give organisms better ways to exploit environmental resources than those possessed by organisms with different traits. August Weismann proposed that changes to an organism's body could not be transmitted to its "germplasm"; consequently, the germ-plasm is passed on to subsequent generations unchanged by environmental interaction. This is an important addition to the Darwin-Wallace substrate specification, and results in a new natural selection formula, which is distinct from Lamarck's, for instance. Again, as noted in the last section, most principles of population genetics make some assumptions about the mechanism of inheritance; the Hardy-Weinberg formula assumes particulate inheritance, sexual reproduction, and constant gene frequencies-each of these assumptions can be summarized in propositional form to yield a substrate specification C in accordance with the definition given above. Fisher attempted to focus on fitness by introducing regression coefficients for fitness, assigning any definite value to these coefficients yields a natural selection formula applicable to a particular trait in a particular organism.³¹ Again, optim-

See Price, "Fisher's 'Fundamental Theorem'," 130. Fisher counted dominance relations and other chromosomal effects as "environmental." Sterelny and Kitcher make the same move in

ality analyses, which implicitly make environmental assumptions, are substrate specifications.

Corresponding to each of these natural selection formulae is a set of possible histories that satisfy the formula. Let us call such a set of possible histories a *natural selection type*. Natural selection types are nested and overlapping. The type SR that corresponds to sexual reproduction will be contained in the type that satisfies the Darwin-Wallace formula, the type that corresponds to *particulate* sexual reproduction will be contained in SR. On the other hand, the natural selection type that merely specifies particulate inheritance will overlap with SR, but not be contained in it, since there are (actual and possible) histories in which non-sexual particulate inheritance was at work, and possible (but non-actual) histories in which sexual reproduction operates through blending inheritance. We conceive of the hierarchical relationship between more and less inclusive natural selection formulae as similar to that between functional kinds and their realizations. Functional types such as *table*, *lock*, *camera* have concrete realizations—individual tables, locks, cameras, etc, these being items that satisfy the definitions of functional types. (Note that an individual material object could in principle belong to more than one functional type.) Similarly, each history in a natural selection type is a *concrete realization* of Li-selection, subject to the substrate specification C; a particular history could satisfy more than one such specification. Functional types have sub-types-refectory tables, combination locks, digital *cameras*. The sub-types of natural selection are sets of histories that satisfy a particular substrate specification. These are kinds of Li-selection—Li-selection with Mendelian inheritance, with sexual reproduction, and so on. This is why we call our model a hierarchical realization scheme.

It is vital to appreciate that the way we conceive it, natural selection is not a factor *within* histories. Rather, it is the differential growth of organism-types in a concrete history. The way we see it, natural selection is not a cause but an effect. The differential growth of organismic types in an evolutionary history is causally explained by factors such as adaptive value, development, and inheritance. It follows that predictive fitness is not a cause either; since it is a growth rate, as we noted at the outset, it too is explained by the same factors.

[&]quot;The Return of the Gene."

A particular pattern of differential growth is explained by subsuming it under a natural selection formula. Suppose, for instance, that we wish to explain a phenomenon like sickle-cell anaemia, which is a disadvantageous trait that arises from heterosis. The natural selection formula invoked will specify several factors-the ecological interactions of homozygotes and heterozygote types, and sexual reproduction. Note that this explanatory formula does not invoke just the optimality considerations involved in the evolution of this hereditary disease, but also the mechanisms of inheritance through which optimality influences the composition of the next generation. In general, we arrive at an adequate explanation of a evolutionary phenomenon by subsuming it under the *least general* formula that applies to it, the formula that posits all the substrate factors relevant to it. The probability of the target phenomenon is estimated relative to the histories that constitute the corresponding *least inclusive* natural selection type (and drift is the unassigned variation in this type). We *understand why* the phenomenon came about by comparing this probability with those yielded by natural selection formulae impose relevantly different substrate specifications. For instance, we understand why a deleterious hereditary condition like sickle-cell anaemia was not eliminated, by comparing the probabilities in its natural selection formula with those with relevantly different ones-ones in which malaria was not a factor, ones in which reproduction is not sexual, and so on.³²

Consider now how we approach questions of interaction among evolutionary factors. How, for instance, does sexual reproduction interact with selection? Where Sober's force model posits an additional force for sexual reproduction, we capture its effects by moving from a natural selection formula in which sexual reproduction is not specified to one in which it is. The advantage is, as we said before, that the added conditions need not be either separable or additive. Separability is not implied because a formula that fails to specify a mechanism of inheritance does not, for that reason, specify selection *without* inheritance. Rather, its type includes *all* forms of inheritance (possible and actual). Thus there is no implication that

We are relying here on Salmon's account of inductive-statistical explanation in *Scientific Explanation*, chapter 2.

selection can operate in the absence of any form of inheritance.

We do not see the theory of natural selection as competing with other explanations of biological traits. Brian Goodwin³³ suggests that certain patterns in plants result from chemical periodicities in morphogenesis. He suggests that our understanding of natural selection has to be modified to accommodate such facts, which, according to him, require us to diminish the role of historical accident in evolution. In our scheme, there will be a natural selection formula that constrains Li-selection by the relevant chemical laws. This formula may well have the consequence that, as Goodwin demands, certain organismic features assigned to chance in more standard formulae turn out to be determined by chemical periodicities instead. What does not follow is Goodwin's additional claim that the theory of natural selection is somehow threatened by the discovery of these laws of morphogenesis. This is based on the confused idea that natural selection is a cause competing for mastery with chemical laws of morphogenesis, and the rank misconception that the theory of natural selection is simply the differential growth of types constrained by morphogenetic rules.

Finally, note that the hierarchical realization scheme makes room for non-additive influences on selection. Suppose that we conduct a regression analysis on the fitness of a certain organism or gene. We find that $f1 \dots fn$ are variables that measure the additive influences on the growth rate of a trait. (The f's might be genes, environmental factors, etc.) Thus we have:

$$F = clfl + c2f2 + \ldots + cnfn$$

With the help of such an equation, we are able to predict what happens when one of the *f*s is modified or absented—what happens when the selective environment is modified, or when a particular gene is removed or modified, etc. Now, introducing a new condition like sexual

Brian Goodwin, *How the Leopard Changed Its Spots: The Evolution of Complexity* (New York: Simon and Schuster, Touchstone Books, 1994), chapters 3-5.

reproduction does not necessarily take the form of adding a new factor to such a list with a coefficient of its own. For sexual reproduction might add interactive effects to the influences already noted, thus necessitating a modification in their coefficients, or the underlying exponent of one of the *f*s, or require the addition of other variables. This is why we should not necessarily think of sexual reproduction, chromosome structure, etc. as vectors added one to another. The hierarchical realization scheme that we are recommending countenances the possibility of a *different* equation of the above sort for each natural selection type.

This paves the way to a new interpretation of "adaptationism," namely that it consists not so much in over-estimating the "power" of natural selection to overcome the other determinants of evolution, but in over-estimating the explanatory powers of natural selection formulae that lean too heavily on optimality analyses concerning traits, omitting consideration of other factors. We noted before, for instance, that evolution would be minimal if there was blending inheritance rather than particulate. For this reason, the Darwin-Wallace formula fails to predict the consequences of differences in vernacular fitness—it lacks an adequate substrate specification. As a consequence, Darwin over-estimated the power of optimality analyses to predict evolutionary change, and was in that way an adaptationist.

VIII. PROCESS VS. TREND IN EVOLUTION

We return now to our question concerning the relationship of predictive and vernacular fitness. In what sense is overall competitive advantage a factor in natural selection? This requires some consideration of causality and process. Can the increase of fitness in a population by natural selection be considered a process that contributes to evolution? What about the spread of adaptive features through a population? Here it helps to consider parallels from thermodynamics. Is increase of entropy a process? What about the transfer of heat from high to low temperature?

Let's start by considering the case of a pot of water brought to a boil on a stove. At a gross phenomenal level, this seems like an orderly process. Heat flows from stove to liquid, the liquid circulates, carrying heat with it; gradually, it comes to a boil. However, closer examination shows that the transformation is not an orderly one at all. This is most evident perhaps in the

actual phase transition, which is highly jerky. Parts of the liquid heave up and down, bubbles form more or less randomly. When the liquid is actually boiling, its surface is chaotic. A microscopic examination would display similar disorder at earlier stages. The energy transfer from the bottom of the pan to the top is disorderly, with myriad local exceptions (for example, adjacent small regions with unequal temperature with no mutual heat transfer), reversals (energy flow from low to high temperature regions), and other fits and starts. Such *discontinuities* violate the spatio-temporal continuity required of physical processes. To halt, reverse, delay, or accelerate a genuine process takes energy and work. But in the pot of water such changes of direction occur spontaneously. This shows that heat flow is not genuinely a physical process. The same point can be made (and historically was made) by reference to *reversibility*.³⁴ A genuine physical process can occur in reverse; heat transfer cannot. You cannot get the heat to flow spontaneously back from the boiling water into the stove top.

Statistical mechanics handles conductive and convective heat transfer stochastically. The pot of water consists of a large number of molecules in random motion. Let us say that a *mechanical* state description of such a system is the set of position and momentum vectors for each particle it contains. Each thermodynamic state description of the system—the set of values for its energy, temperature, and other such thermodynamic variables—supervenes on a mechanical state description thereof. However, this supervenience relation is *one-many*: that is, the same thermodynamic mechanical state of a given system supervenes on many different mechanical states. Now, some thermodynamical states comprehend a larger number of distinct underlying mechanical states than others; those that comprehend more are "more probable," those that comprehend fewer are "less probable." The basic principle of statistical thermodynamics is that less probable thermodynamic states give way in time to more probable ones, simply by the underlying molecules following their own inertial trajectories and interactions. This explains why statistical thermodynamics has "universal applicability": it is

An accessible account of reversibility can be found in P.W. Atkins, *The Second Law* (New York: W.H. Freeman, *Scientific American Books*, 1984).

silent about the constitution of entities and their dynamics, and mathematically demonstrates a phenomenon that occurs regardless of these physical details. The heat-flow described above is stochastic in nature; it is the consequence of the mathematics of large numbers, not of the specificities of physical law.

The same is true of evolution. The stochastic aggregates involved take a much simpler form here than in the thermodynamical case, the ontology of the probabilities being relatively uncomplicated. (See notes 7 and 24.) Moreover, the discontinuities of natural selection are much more evident to the observer than in thermodynamics, gross enough, in fact, to be recordable by a careful observer.³⁵ These discontinuities show that natural selection, like thermo-dynamic change of state, is a time-asymmetric statistical trend on populations. Natural selection is a single phenomenon because it is amenable to the mathematics of ensembles, then, rather than on account of any causal similarity of the substrates in which it occurs. This is what the hierarchical realization scheme is intended to capture: natural selection is the consequence of different processes in different substrates sharing only a formal structure. Like thermodynamics, it is silent about "the constitution of entities and their dynamics."

The reduction (or whatever it is³⁶) of measurable and theoretically constructed thermodynamic variables—temperature, entropy, heat—to stochastic measures on mechanical states is an *elimination* of thermodynamic process. In much the same way, the mathematical-statistical treatment of natural selection eliminates evolutionary process.³⁷ Thus differences of

35

36

We are alluding, again, to Sklar, "The Reduction (?) of Thermodynamics."

For a famous and vivid study of discontinuities and reversals in selection, see Peter Grant *Ecology and Evolution of Darwin's Finches* (Princeton: Princeton University Press, 1986). Of course, there are climactic fluctuations in the Galapagos, and so one cannot say that the constantly fluctuating condition of the finch population is *spontaneous*. Nevertheless, the discontinuities demonstrate that long term selective trends there are pseudo-processes.

predictive fitness cannot be, in the physical sense at least, the *cause* of changes of gene frequencies that occurs in a population divided by growth rates. We do not wish to deny that they may be described as "causes" in some contexts—there is a class of probabilistic definitions of cause as raisers of probability, and predictive fitness is a "cause" in this sense. Nonetheless, fitness differences do not obey the laws of physical process. Let us define a *stochastic property* as one that belongs to ensembles as a *mathematical* (note: not *nomic*) consequence of the **i**-level properties of individuals in that ensemble. Further, define a *trend* as a change of an ensemble over a period of time with respect to one or more of its stochastic properties. The claim that we want to make is that while differences of predictive fitness are predictors of trends in populations, and are thus *stochastic causes*, they are not causes in the sense appropriate to processes, of these trends. Further, natural selection is a trend, but not a process.

What about vernacular fitness—the overall competitive advantage of an organism traceable to its heritable traits? Of course, it is not a process-cause either, since natural selection is not a process. And at the i-level, we have argued that it plays no determinate role: in a concrete history, one cannot apportion causal responsibility between advantage and chance. Moreover, as we noted before, we have no way of combining the different advantages that an organism possesses in virtue of its various traits into a single quantity. Now, the latter consideration is an obstacle to counting vernacular fitness even as a stochastic cause at the e-level. There are two legitimate ways of plugging fitness into the laws of population genetics. First, we can estimate how particular differences of traits will affect organisms that are the same in other ways, and by feeding these partial fitness values into the equations of population genetics arrive at estimates of selection pressure on one trait as opposed to an alternative. Second, we can use statistical estimates of *predictive* fitness in these equations, and arrive at estimates of future trends. Neither of these options involves appeal to the notion of overall competitive advantage. Thus, vernacular

In "Chasing Shadows: Natural Selection and Adaptation," *Studies in the History and Philosophy of Biology and the Biomedical Sciences* 31 (2000): 135-53, Denis Walsh argues, on grounds independent of those given here, that heat transfer and natural selection are "pseudo-processes."

fitness doesn't play much of a role in evolutionary theory; it is neither a physical cause nor a stochastic cause. It is useful in common sense expositions, but not much more. Scientifically, it is a notion ripe for discarding.

We should distinguish this conclusion from two other propositions with which it might be confused. First, some will find that our confining causal process to the i-level, and insisting that e-level transitions are trends, reminds them of the idea that in the theory of natural selection, probabilities are *epistemic* only, used because we are ignorant of the individual deaths and births that constitute the *i*-level reality of natural selection. In fact, we are not sympathetic to the latter notion, and we have defined evolutionary probabilities on the sets of possible histories that constitute natural selection types. Moreover, we do not regard statistical theories merely as devices to deal with ignorance. The statistics employed in thermodynamics explains something about the nature of heat and work, something we would not understand by having a history of molecular interactions in a pot of water, even assuming that our minds could comprehend that kind of detail. Similarly, the statistics of natural selection tells us something deep about the patterns instantiated in diverse biological histories. By appreciating these patterns, we come to understand something that we are not able to see when we are given the full biographical details of organisms in diverse populations.³⁸ The theory of evolution is a historical science in the sense that it tries to retrodict significant events in the individual paths traced by actual species. Population genetics, however, is concerned with the formal characterization and mathematics of all such histories, taken as a group. The interest of these patterns is not simply historical reconstruction, any more than it is in thermodynamics.

Second, some may think that we are asserting that if a class of properties S supervenes on base properties B, then since all changes in properties S are wholly determined by properties B, there are no genuine causal relations at level S. In fact, we have not relied on the supervenience relation between **e**-level properties and **i**-level properties in making our point. We have distin-

See André Ariew, "Are Probabilities Necessary for Evolutionary Explanations?" *Biology* and Philosophy 13 (1998): 245-53

guished two kinds of causal relations, stochastic and process. We concede that stochastic causation occurs at the *S* level,³⁹ but deny that process causation occurs at this level. This denial is made *on grounds of discontinuity and irreversibility*, not of supervenience.

The view that natural selection is only a part of evolution is becoming increasingly fashionable. Some place the recently heightened interest in developmental constraints and morphogenetic process in this context; they say that these influences on organismic traits should be given more prominence relative to natural selection than they have been given hitherto. In the statistical way of thinking about natural selection, this is misguided. The way we see it, evolution involves nothing over and above natural selection: evolution is a cumulative change of gene frequencies, and natural selection is the pattern of differential growth of genes. Having said this, we do want to acknowledge that the construction of a genuinely all-inclusive conception of natural selection in the biological domain is extremely challenging. Lawrence Sklar's discussion of thermodynamics is once again revealing here. Speaking of temperature, Sklar says:

Systems of wildly divergent physical kinds can all have equilibrium states and, when allowed to transfer energy to one another, be in equilibrium with respect to each other. The moving molecules in a blob of matter, for example can form a system in equilibrium with electromagnetic radiation. So we are required to think of these systems as sharing common temperatures. ("The Reduction (?) of Thermodynamics, 194; see also note 24 above)

Here we concur with Elliott Sober, "Physicalism from a Probabilistic Point of View," *Philosophical Studies* 95 (1999): 135-74, at 145-48.

Something like the same problem arises in the theory of natural selection because of the occurrence of both genetic selection and cultural selection in human evolution.⁴⁰ Here, as in Sklar's example, selection is occurring simultaneously on two quite different substrates. The question is: how do we understand the interaction of these two substrates? The imposition of such crossing relations on our hierarchical realization scheme is no small challenge.⁴¹

40

See P. E. Griffiths and R.D. Gray, "Developmental Systems and Evolutionary Explanation," *Journal of Philosophy* 91 (1994): 277-304, and Susan Oyama *The Ontogeny of Information: Developmental Systems and Evolution* (New York: Cambridge University Press, 1985).

41

The distinction between physical and statistical ways of thinking about factors that are relevant to evolution has been investigated independently by Denis Walsh and Mohan Matthen, who arrived at the idea along different paths. André Ariew, who has been working with both, brought about considerable convergence both by cross-communicating and by his own critical contributions. The present paper and Walsh, Lewens, and Ariew, "Trials of Life" (see note 3) are products of this process. However, these papers are almost completely independent with respect to argumentation and even, to a surprising extent, motivation. "Two Ways of Thinking" was presented before the 2001 Congress of the Canadian Philosophical Association, and helpfully commented on there by Wayne Myrvold. We acknowledge helpful discussion with Paul Bartha, Bill Harper, Andrew Irvine, Dick Lewontin, Joel Pust, Patrick Rysiew, Elliott Sober, Larry Shapiro, and Catherine Wilson.