TWO WAYS OF THINKING ABOUT FITNESS
AND NATURAL SELECTION*

The concept of fitness is, Philip Kitcher[1] says, "important both to informal presentations of evolutionary theory and to the mathematical formulations of [population genetics]" (*ibid.*, p. 50). He is absolutely right. The difficulty is to harmonize these very different ways of understanding its role. We examine here how fitness and natural selection relate 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 framework of population genetics. We argue that this model demands a far-reaching revision of some widely accepted notions of causal relations in evolution.

I. VERNACULAR FITNESS AND PREDICTIVE FITNESS

The two formulations of fitness are well illustrated by expositions given by Kitcher 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 (*ibid.*, p. 38).

[1] *Vaulting Ambition: Sociobiology and the Quest for Human Nature* (Cambridge: MIT, 1984). 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, Political, and Political Perspectives* (New York: Cambridge, 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.

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. The resulting differential growth of gene frequencies is natural selection. This much is central 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 shall call this measure of an organism's selective advantage 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 *comparative*, not *quantitive*, 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 equations of population genetics which 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 a statistical measure of evolutionary change, the *expected* rate of increase (normalized relative to others) of a gene, a trait, or an organism's representation in future generations, or, on another interpretation, its *propensity* to be represented in future generations, suitably quantified and normalized.[2]

Unlike vernacular fitness, predictive fitness is not a cause of selection, or of evolution for that matter. The expected rate of return on an investment is not a *cause* of its growth—it just *is* its growth, numerically estimated and projected into the future. Likewise, predictive fitness is a measure of evolutionary change not a cause. Nor do these conclusions have to be modified much if we treat these quantities as *propensities*. An investment's propensity to grow is a consequence of all the factors that economists identify as causes of growth—profit, growth prospects, dividends, and so on. From an economist's point of view, explaining an investment's propensity to grow is no different from explaining its expected growth. In exactly

---

[2] We do not mean to insist on the expected-value interpretation; there are many ways of defining predictive fitness within population genetics. For a review, see John Endler, *Natural Selection in the Wild* (Princeton: University Press, 1986).

the same way, even if predictive fitness is thought of as a propensity, and so a cause of evolutionary change, it is so closely identified with that change as to make no difference from a scientific point of view—explaining one is the same as explaining the other.

Now, what is the relationship between vernacular fitness and predictive fitness? As we have seen, conventional wisdom holds that vernacular fitness is a cause of evolution. If predictive fitness measures the rate of evolutionary change, then, vernacular fitness might be thought to be a 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 behavior when tossed. In other words, one is tempted to say that part of the reason a particular gene tends to increase in frequency is that it has a competitive advantage. On the other hand, vernacular fitness is not the *only* thing responsible for evolution. Such other factors as the underlying mechanism of inheritance and drift must be taken into account, as well as various constraints on selection. These supposedly "nonselective" factors would need to be combined with vernacular fitness in any adequate account of the "why" of predictive fitness. How are they to be combined? Any attempt to relate vernacular fitness to predictive fitness, and thus to harmonize Kitcher's "informal" and "mathematical" discourses, needs to come to grips with this question.

One standard approach to this question is to say that the contribution 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 through 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 and discussed in sections v through viii. We argue further, in section ix, that natural selection is not a process driven by various evolutionary factors taken as forces; rather, it is a statistical "trend" with these factors (vernacular fitness excluded) as predictors.[3] These theses demand a radical revision of received conceptions of causal relations in evolution.

---

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

II. THE THEORY OF EVOLUTION AS A THEORY OF FORCES

Elliott Sober[4] is the author of one influential suggestion about how the theory of evolution accommodates natural selection alongside drift, developmental constraints, architectural constraints, and other such determinants of evolutionary change. He proposes that we should think of the theory of evolution as a theory of forces along the lines suggested by Newton's mechanics (*ibid.*, p. 31), and of natural selection (or vernacular fitness) as *one* (but only one) of the forces that may act on a population.

Sober urges that the following structural elements of Newton's 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 (*ibid.*, p. 32) that the corresponding steady state in evolution is specified by the Hardy-Weinberg law of population genetics. This is not exactly right—the Hardy-Weinberg law tells us about genotype frequencies only 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, 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 force—gravitation, electrostatic force, magnetic force, impact, drag, and so on. In classical physics, "source laws," as Sober calls them, describe how these forces emerge, one by one, from various physical situations. The law of gravitation, Coulomb's law, the law of elastic bodies, the laws of aerodynamic resistance, and the like 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 (*ibid.*, p. 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 which predict the selective value of various traits and the laws that govern how genes located on the same chromosome assort during recombination. Accordingly, causal responsibility for total evolutionary change can be 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. Evolutionary forces combine in this way, according to Sober. He offers the case of *heterosis* to illustrate the point, contending that here selection is added to the genetic recombination found in sexual reproduction (*ibid.*, p. 47). (A description of this phenomenon, and our own analysis, will be found in section v.)

(4) *A law describing change under resultant force.* Newton's second law of motion ($F = ma$) 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.

Now, it is of central importance in Newton's theory that the mathematical device of "resolving" forces by vector addition does not compromise the *separate* operation of "component" forces like gravitation and drag. Newton's second law of motion predicts change given force. It makes no difference, however, whether we first combine forces by vector addition and then use Newton's second law to derive acceleration, or, reversing the order of these operations, first feed the component forces into the second law one by one and then use vector addition to combine the separate acceleration vectors that result. In other words, acceleration under the sum of component forces is just the same as the sum of accelerations due to the component forces taken separately. Resultant force is, in this respect, dispensable. (It is needed, however, to preserve the universality of Newton's second law by positing a force that corresponds to total acceleration.) Physically, the component forces act independently of one another; there is no mechanism that creates a new force out of them.

This independence condition for component causation, which we regard as indispensable to the Newtonian apparatus of resolved force, fails in the evolutionary case. The consequence laws are, as we have seen, the index of total evolutionary change in population genetics. So, if Sober's conception of evolutionary forces is right, one should expect that vernacular fitness and the other causes of evolution add up to a resultant force. If the analogy with Newtonian mechanics is to be maintained, these components should, despite such summation, retain their separate causal influences. But, as we shall argue in the next two sections, *it is incoherent to think of the component factors contrib-*

uting to evolutionary change by separate action. As a consequence, the analogy with Newtonian forces collapses: the first three structural elements mentioned above find no counterpart in evolutionary theory, and the fourth, the notion of a consequence law, makes sense only in terms of probabilistic quantities like predictive fitness, with no role for vernacular fitness.[5]

III. DRIFT AND THE ZERO-FORCE LAW

The trouble starts with the zero-force law. In Newton's theory, inertia—the property of each massive body to resist changes to its "state of rest or uniform motion along a straight line"[6]—ensures 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.[7]

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

---

[5] Michael Strevens forced clarity here, and in the section I discussion of predictive fitness.
[6] *Principia,* Definition III.
[7] *The Nature of Selection,* p. 34.

the outcome *two heads*? Answer: the physical setup of the coin-tossing trials. What would explain the outcome *four heads*? Answer: the same thing. Although the second result is less probable, the same setup explains both outcomes. The set of possible worlds in which fair coins are tossed four times form what Wesley Salmon[8] 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.[9] Perhaps one might say that the four-head outcome was less predictable, therefore less well explained within this setup. However that may be, there is nothing available with which to strengthen the explanation of 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" that occasionally favor the less likely outcome.

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 is not the sort of thing that can play the role of a force—it does not 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.

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

    [9] 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.

Histories of natural selection consist, after all, of collections of con-
crete individual events: births, deaths, matings, mutations, and so on.
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 indi-
vidual events and determining when discriminate sampling has been
at work, and when indiscriminate. In this way, one apportions the
outcome to fitness and drift, respectively.

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

(C1) O2's bad eyesight 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 gene frequency 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 evolu-
tion, 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-
biased) cases like (C1), but that it excludes anomalous (or fitness-
indiscriminate) cases like (C2). Then it might be thought plausible to
say that something else—drift? indiscriminate, or neutral, selec-
tion?—is operating in (C2).

But this violates sound probabilistic thinking. Probability enters
into the picture because the theory of evolution abstracts away from
concrete individual events like (C1) and (C2) in order to isolate the
causal factors that make a probabilistic *difference* to evolution.[10] Since,

---

[10] 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

ex hypothesi, your chances of being killed by a lightning strike are the same whether you have good eyesight or bad (other factors being held equal), this factor cannot be conjoined with either good or bad eyesight to yield finer predictions. This does not mean that evolutionary theory assigns cases like (C2) to some special category of "uncaused events." Lightning *has* been taken into account when we say (a) that bad eyesight and the like causes earlier death with probability p (less than one), and further (b) that there is *no other factor* that will enable us to refine this estimate.[11] That is, it has been taken into account when we judge that there is no variance among individuals in terms of their vulnerability to lightning strikes.

This highlights a confusion of levels which can occur when we speak of probabilistic causes. On the one hand, we say that lightning strikes are not a factor in natural selection. By this, we mean something like the following: lightning strikes do not, over large populations, over the long haul, result in the differential retention of heritable traits. This is what we shall call an *ensemble*-level fact, or *e*-fact: it is concerned with frequencies, probabilities, or propensities which derive from aggregates. On the other hand, we need to acknowledge that *any* death, from whatever cause, results in evolutionary change since it results in a change of gene frequencies. Consequently, we can say that lightning strikes cause evolutionary change in cases like (C2). This is a statement which derives from a causal relation at the level of concrete *individuals*, not ensembles: it is an *i*-fact. The pair of statements seems contradictory only when the difference of levels goes unnoticed.

Now, it is often not sufficiently well understood in discussions of causation that there is no straightforward translation between levels with regard to causal statements. 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. It states an e-fact. Now consider a particular individual X, who has good eyesight and has lived a long life. Here we are considering an i-fact. We might wonder what exactly the e-fact just alluded to tells us about the causation of the i-fact concerning X. How many years did good eyesight contribute to *her* long life, and how much of her life should we attribute to good luck? Generally, this question has no determi-

---

discussing. As with the fair coin, one would not be justified in claiming that the individual events above contained elements of indeterminacy.

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

nate answer. 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. "Unassigned
variation" has no i-counterpart, however. Although 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 in general differentiate between individual events
on the basis of how much they are attributable to chance.

For exactly the same reason, one cannot *in general* determine, even
retrospectively, whether discriminate or indiscriminate sampling was
responsible for an actual sequence of historical events. In an example
discussed above, *T* and *T'* were traits that 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 also know things like this: the chances of the less viable trait
going to fixation are greater if the population is smaller—this appli-
cation 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 unas-
signed variation to event types. There is, however, no such thing as
unassigned variation with respect to *concrete* sequences of events
(barring true indeterminacy). 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.[12] Even if we are able to identify each
and every i-event in that history, and determine which involved the
(deterministic) operation of factors, like good eyesight, that make an

---

[12] Alexander Rosenberg—*Instrumental Biology or The Disunity of Science* (Chicago:
University 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 evolutionary scenarios. But he is wrong if he is
talking about evolutionary *theory*, which abstracts away from individual causes.

e-level difference, it will still not make sense in general to quantify the residual role of chance.[13]

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 may not be well defined in this instance. Suppose that over a period of time a population stays exactly the same, or changes in some determinate way. The proposition that drift was involved to degree p in this history generally has no determinate truth value.

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[14] 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." 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 behavior" (*ibid.*, p. 256).

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 shall 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.

Gould and Lewontin think this view is mistaken because in fact natural selection is strongly opposed, indeed often thwarted, by these

---

[13] Strevens points out (personal communication to Matthen) that in some cases, it may be possible to determine whether chance was involved. Strevens himself is interested in a range of cases where the i-level causes can be definitively identified. In some such cases, it might be possible to assess the exact involvement of various probabilistic causes.

[14] "The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme," *Proceedings of the Royal Society of London*, CCV (1978): 581-98; reprinted in Sober, ed.,*Conceptual Issues in Evolutionary Theory* (Cambridge: MIT, 1984), pp. 256-70. Page references are to the latter version.

other forces—in particular by genetic, architectural, and develop-
mental 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
nonselective constraints, and how much to natural selection, we need
first to appreciate the action of the latter acting alone. 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 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 selec-
tion acts in certain underlying causal media, and the so-called con-
straints 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 shall return to this point in the next section.) If this
is right, we have no way in this context of making sense of the
Newtonian idea that, if an effect is to be analyzed 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 which 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 sex determination 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, 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.[15]

Since we cannot add up the values produced by the source laws, the quantification of fitness in the laws of population genetics is conceptually independent of their occurrence in source laws in a way not paralleled in the Newtonian treatment of force. The overall fitness values demanded by consequence laws must be estimated statistically, that is, 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.[16] But 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—

---

[15] Donald Davidson makes structurally the same point in "How Is the Weakness of the Will Possible?" in Joel Feinberg, ed., *Moral Concepts* (New York: Oxford, 1971), pp. 93-113. He argues (pp. 108-11) 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" judgment on whether one should brush one's teeth.

[16] "Fitness," this JOURNAL, LXXX, 8 (August 1983): 457-73, at p. 459.

helps explain why *general* source laws are hard to come by."[17] We agree
that "fitness is not a single physical property" (see section VII). But the
multiplicity of realizations, whether with regard to source or conse-
quence laws, does not mark a disanalogy between the physics and
evolutionary biology: Newtonian force has multiple physical bases, too—
charge, mass, collisions, and so on. 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 geneticists are forced to estimate fitness by
measuring population change.

V. SELECTION AND ITS SUBSTRATES

We have now argued in support of three propositions. First, there
is, in general, 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," in the way that gravitation acts alone
on an uncharged particle in a vacuum. So the decomposition-of-
forces condition fails. Third, there is no common currency in
which to compare the contributions of different evolutionary
"forces." Thus, the resolution of forces does not have any coun-
terpart in the 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 natural selection acts only in the
context of certain underlying causal determinants—chromosomal
structures, mode of reproduction, pre-existing body plans, devel-
opmental sequences, and so on. Instead of thinking of selection as
a force acting independently of these factors, either competing
with them or being reinforced by them, we shall revert to another
historically familiar way of thinking about selection, namely, as a
statistical trend emerging from events that occur in these "sub-
strates." To put this somewhat fancifully, we prefer the metaphor
of the hidden hand—that is, of e-level occurrences emerging from
i-level events without any inherent tendency in the latter to pro-
duce them—to any conception that reifies selection as an *influence*
with inherent direction. In the hierarchical-realization model that
we shall now begin to construct, selection is a *formally* character-
ized phenomenon, a statistical property of physical substrates that

---

[17] *The Nature of Selection*, p. 51.

possess certain metrical properties. The causally active physical properties that lie beneath this metric are different from, or even incommensurable with, the properties that form the subject matter of the theory of selection.

To illustrate this change of perspective, we now present two examples of the substrate dependence of selection. Recall first that Charles Darwin and Alfred Wallace hit on the principle of natural selection in an extremely general form: a heritable trait that enables its possessor to leave behind more descendants than those which lack the trait will increase its proportion in the population. Now, it turns out that, as R. A. Fisher[18] first argued in detail, the principle of natural selection is insufficient by itself to explain the facts of evolution. Fisher showed that under the system of inheritance envisaged by Darwin—"blending inheritance"—the 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 (*ibid.*, chapter 1). Mendelian inheritance, in which discrete genes are handed down unchanged from parent to offspring, is required to maintain variation in the face of ecologically imposed homogenization.[19] It is not natural selection, then, that is responsible for the "origin of species," but natural selection in a Mendelian substrate. Our argument in section IV shows that this statement should not be interpreted in terms of additive forces. We shall be suggesting that a more appropriate diagnosis is in terms of certain formal properties of the Mendelian substrate.

The phenomenon of *heterosis* provides us with another 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

---

[18] *The Genetical Theory of Natural Selection* (New York: Dover, 1958, second edition) (the first edition was published by Oxford University Press, 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" (p. 10).

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

sense) than either homozygote, a phenomenon known, when it occurs in the plant world, as "hybrid vigor." Let us imagine an extreme case of this: the vast majority of each homozygote type aborts immediately upon conception, but most heterozygotes survive to reproduce in the normal way. If so extreme a difference of viability were to occur among *a*sexually reproducing organisms, the viable phenotype would go to fixation in short order, and the others would disappear. If it occurred with blending inheritance, the whole population would similarly settle into some nonlethal 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 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.

These analyses demonstrate the way in which selection emerges from substrate interactions summed up over ensembles. In the force model, blending inheritance would be seen as opposing selection, while particulate inheritance would be seen as reinforcing it. In addition to the infelicities noted earlier, this has the additional disadvantage that it leads us to expect only that selection would be *slower*; it gives us no reason to think that the magnitude of evolutionary change would, in the long run, be diminished. Fisher's analysis addresses the issue at a different level. It looks at *variation* within the population, and tracks the effects of this variation on the gene pool transaction by transaction, as it were. Of course, Fisher did not *literally* track individual transactions; he simply used statistical methods to sum over them. But this is equivalent to simulating ensembles of individual transactions one by one on a computer. The important point is that Fisher did not ask about the goal toward which selection is supposedly directed. Rather, he asked: 'What is the effect of random mating in a population in which some organisms have a greater probability of reproduction'? To answer this question, he followed the sequence of events through at the genic level. In his way of looking at things, selection emerges as the mathematical *aggregate* of multiple interactions at the i-level, where the latter interactions are governed by laws appropriate to the causal substrate in which these interactions occur. A similar form of analysis is used in the heterosis case. The equilibrium there is not seen as an opposition

of matched forces either of which can be switched off or diminished.[20] It too arises as an aggregate of i-level occurrences.

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. Variations in rate of reproduction may occur at the phenotypic level, in which case effects on gene frequencies are indirect, and also at the genic level, in which case the effect is direct. Selection can be more or less determinate: the probabilities involved 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, and so on. Selection also occurs in nonbiological realms: in the economic domain, for example, as well as in "clonal selection" in the mammalian immune system, in classical conditioning,[21] and, according to some, in the propagation of theories and other cultural artifacts.[22] While the Darwin-Wallace principle continues to hold true under all these conditions, it is silent on *how* events in one generation leave their mark on subsequent generations. The properties relevant to the speed and sustainability of selection are shared by causally heterogeneous substrates, which is why these properties must be formally characterized.[23]

The multirealizability of natural selection parallels that of statistical thermodynamics. For as Lawrence Sklar[24] 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 (*ibid.*, p. 189). Like the theory of natural selection, statistical mechanics has a formal component: this is concerned with probability distributions on large ensem-

---

[20] Walsh, Lewens, and Ariew argue (*op. cit.*) in a closely related argument that equilibrium states in selection, such as the one arrived at in the heterosis example, are different from those of a feather floating in air, or of an oil drop motionless under the combined effect of gravity and electrostatic force.

[21] See David Hull, Rodney Langman, and Sigrid Glenn, "A General Account of Selection: Biology, Immunology, and Behavior," *Behavioral and Brain Sciences*, XXIV (2001): 511-28.

[22] Donald E. Campbell, "Evolutionary Epistemology," in P.A. Schilpp, ed., *The Philosophy of Karl Popper*, Volume 1 (LaSalle IL: Open Court, 1972), pp. 412-63.

[23] In "Teleology and the Product Analogy," *Australasian Journal of Philosophy*, LXXV (1997): 21-37, Matthen argues that the diversity of processes in these substrates undermines the claim that there is a distinctive form of causation involved in the genesis of teleological functions, and proposes that they are based on an analogy.

[24] "The Reduction (?) of Thermodynamics to Statistical Mechanics," *Philosophical Studies*, XCV (1999): 187-202.

bles 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" (*ibid.*, p. 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" (*ibid.*, p. 195). Like those found in evolution, thermodynamical changes are the accumulation of a vast number of underlying events.[25] As Fisher kept emphasizing, it is statistical thermodynamics—not Newtonian dynamics—that provides the closest parallel in physics to the theory of natural selection.

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 abstract 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 $100.00 into mutual fund *A*, and $100.00 into *B*. *A* grows steadily at 6% per annum, *B* (to your chagrin) at 2%. Now, the average growth rate of your total investment starts out at 4%. But as your subinvestment 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 $106.00 in *A* and $102.00 in *B*; consequently, the overall growth rate is 4.04%. At the end of forty years, there is $970.00 in *A* and $216.00 in *B*, and the overall growth rate is 5.25%. Over time, the growth rate of the whole approaches the

---

[25] There are 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" (*op. cit.*, p. 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, 1993), p. 23, our emphasis. We shall see at the end of section IX that somewhat analogous points can be made with regard to natural selection.

A-rate, 6%. 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. As a consequence, the increase in the overall growth rate slows down. Thus the increase in the overall growth rate tracks variance.

Anthony Edwards[26] reports that this "growth-rate theorem" was presented by C. C. Li as a "simplified version of Fisher's fundamental theorem of natural selection" (*ibid.*, p. 444). It is certainly simplified in that it does not take on board such complications as sexual reproduction, dominance, linkage, ecological change, and so on, while Fisher's theorem and its successors contain parameters that sum up the effect of such factors.[27] Consequently, Li's theorem cannot be applied to phenotypically defined subpopulations 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. It does apply to populations of genotypes, however, though what is needed for the application is values for the growth of each genotype, or gene.[28] In effect, the theorem *defines* selection as what happens to the parts of a population when these parts have different growth rates. It implies that the faster growing subtypes increase their representation in the whole, at a speed proportionate to their variance from the mean. (Note that

---

[26] "The Fundamental Theorem of Natural Selection," *Biological Reviews of the Cambridge Philosophical Society*, LXIX (1994): 443-74, citing C. C. Li, *Population Genetics* (Chicago: University Press, 1955). Readers are referred to Edwards for a simple proof of the theorem; it is, however, quite instructive to chart the example given above in a spreadsheet program. It should be noted that the proof assumes that change in growth rates is always continuous. Patrick Maher suggests (personal communication to Matthen) that it cannot handle discontinuous changes of growth rate or discrete population growth events (for example, an annual birthing, where there are rate changes from one year to the next).

[27] See the beautiful exposition in G. R. Price, "Fisher's 'Fundamental Theorem' Made Clear," *Annals of Human Genetics*, XXXVI (1972): 129-40. Price himself has a general mathematical characterization of natural selection in "Selection and Co-variance," *Nature*, CCXXVII (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*, LI (1997): 1712-29.

[28] 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 Kitcher show—"The Return of the Gene," this JOURNAL, LXXXV, 7 (July 1988): 339-61—George Williams's "book-keeping argument" implies just this. The fancier formulations of Fisher and Price add nothing to our understanding of the essential character of selection.

blending inheritance would result in a faster diminution of variance, which would damp change out faster.) This definition of selection is mathematical in nature, and independent of the particular causal laws that produce growth. (Of course, substrates with different causal laws may for that reason have different formal properties.)

(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 subpopulations. Note that this does not tell us much about *vernacular* fitness, that is, about the causes of growth in a subpopulation, or about progress.[29] 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.[30])

In calculating the effect of "selection pressure" on the investment account, we were dealing with a very simple case. There is no differentiation between genotype and phenotype: 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. Li's theorem tells us nothing about causes of growth: it is a general truth about growth regardless of how it is caused. Moreover, selection is not a cause of growth (or of the change in population characteristics) in this conception; it is the mathematical aggregate of growth taking place at different rates.

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 shall put Li's theorem in this role, though there are other

---

[29] 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 postwar 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.

[30] It would be a mistake to think that Li's theorem (or Fisher's) depends on a growth-rate definition of *fitness*. These propositions 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 Ariew). In fact, one of us (Ariew) argues that the Fisherian definition is misleading as a definition of statistical fitness (Ariew and Lewontin, in preparation).

candidates, such as Fisher's "fundamental theorem," Price's equa-
tion, and Richard Michod's[31] "Darwinian dynamic." The difference
between the proposed approach and that of a theory of force is that
a conjunction of propositional conditions does not entail, as the force
analogy does, the existence of separate additive causal influences
corresponding to each propositional conjunct. One can talk of con-
sequences emerging from domains in which p and q are satisfied,
without implying that these propositions correspond to separate
forces.

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—that is, L posits a population
> subdivided by growth rates—and C is a *substrate specification* which states
> properties of population (including properties of its members or of their
> parts), and/or the causes of differential growth rates in these popula-
> tions and their parts, and/or conditions of inheritance, development,
> and environmental interaction.

There is a great variety of natural-selection formulae, as the following
examples demonstrate.[32] (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, and further that these
traits are inherited. Weismann proposed that changes to an organ-
ism's body could not be transmitted to its "germ plasm"; conse-
quently, the germ plasm is passed on to subsequent generations
unchanged by environmental interaction. This is an important addi-
tion to the Darwin-Wallace substrate specification, and results in a
new natural-selection formula, which is distinct from Darwin's own,

---

[31] *Darwinian Dynamics: Evolutionary Transitions in Fitness and Individuality* (Prince-
ton: University Press, 1999).

[32] Hull et alia (*op. cit.*) attempt to define a formula that captures Darwinian
natural selection exactly, omitting all that is extraneous and including everything
essential, a formula that is neither "too broad or too narrow." We are skeptical
about the value or feasibility of coming up with such a formula: our hierarchical
conception undermines the idea of a unique way of specifying "gene-based evolu-
tion."

or Lamarck's. 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 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.[33] Again, optimality analyses, which implicitly make environmental assumptions, are substrate specifications.

Corresponding to each of these natural-selection formulae is the 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 nonsexual particulate inheritance was at work, and possible (but nonactual) 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, and so on, these being items that satisfy the defining formula of the 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 will satisfy more than one such specification. (Note that an individual material object, and also a natural-selection history, could in principle belong to more than one type.) Functional types have subtypes—*refectory tables, combination locks, digital cameras.* The subtypes 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. These types have subtypes, too. This is why we call our model a *hierarchical-realization scheme.* Note that the novelty of our

---

[33] See Price, "Fisher's 'Fundamental Theorem'," p. 130. Fisher counted dominance relations and other chromosomal effects as "environmental." Sterelny and Kitcher make the same move (*op. cit.*).

proposal does *not* lie in our making selection multirealizable; earlier, we quoted Rosenberg and Sober as already committed to this. The new feature in the scheme outlined above is its hierarchical arrangement, with various factors taken to be explanatory of evolution figuring in the substrate specifications.

We arrive at an adequate explanation of the evolution of a biological phenomenon by subsuming it under the *most specific* formula that applies to it, that is, 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. We understand why the phenomenon came about by comparing this probability with those yielded by natural-selection formulae which 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.

This paves the way to a new interpretation of adaptationism. It consists not so much in overestimating the "power" of natural selection to overcome the other determinants of evolution, but in overestimating the explanatory powers of natural-selection formulae that lean too heavily on optimality analyses concerning traits, omitting specification of other factors. We noted before, for instance, that evolution would be minimal if there were 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. Darwin overestimated the explanatory power of natural-selection formulae that omit specific mention of inheritance and mention only optimality analyses. To this extent, he was an adaptationist.

VIII. THE DISAPPEARANCE OF VERNACULAR FITNESS

Consider now how to 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. Separability is not implied: a formula that fails to specify a mechanism of inheritance does not, for that reason, specify selection *without* inheritance. When we consider evolutionary effects that do not depend on sexual reproduction, or when we wish to consider an effect in abstraction from mode of reproduction, we

might ascend to a more inclusive type that includes, for instance, *all* forms of inheritance (possible and actual). In this more inclusive selection type, the growth rates are averaged out over all forms of inheritance. There is no selection in this or any other type in the absence of some sort of continuity that could be considered a fill-in for inheritance.

In this way of looking at things, the distinction between evolution (the total change of gene frequencies due to all causes), and natural selection (the portion of evolution due to differences in competitive advantage) is unmotivated. Natural selection is, as Li's theorem tells us, the aggregative result over time of differential growth rates in a population. These growth rates are explained by considering *all* of the factors posited by the most specific relevant natural-selection formula, competitive advantage acting in concert with all the others. In histories that conform to this formula, certain trends get established at the e-level as accumulations of multiple concrete events—births, deaths, mate choices, as well as events at the cellular and molecular level. These trends constitute natural selection. There is no difference between these trends and evolution. Further, since these trends constitute total change from generation to generation, predictive fitness is their measure, not the supposed partial value that some posit as vernacular fitness. Like the distinction between natural selection and evolution, the distinction between vernacular fitness (an organism-type's overall competitive advantage traceable to heritable traits), and predictive fitness (its expected growth rate, or the propensity underlying it), is unmotivated.

Some might resist this conclusion on the grounds that actuarial practice regularly apportions causal responsibility to various factors. (After all, this is what multivariate analysis is all about.) In responding to this concern, we want clearly to acknowledge that it *is* legitimate to ask, *in a statistical sense,* how much of the causation of B is due to competitive advantage, though as we noted in sections III and IV, one needs to be prepared for a degree of indeterminacy in the answer, since the composition and decomposition of probabilities is mathematically undefined in many cases. This question is similar to the questions asked when determining insurance premiums. For example: How much does being a male youth contribute to road accidents in which male youths are involved? There is some sort of answer to this question in statistical correlations. But this does not imply that being a male youth operates in an ontologically separate way in these accidents. This is clearest when we think about the individual events. An eighteen-year-old male is involved in a collision at night, in the rain, after drinking three pints of beer. Are these factors, specified in

this way, separable influences on the outcome? We think not: the same goes here as for the factors involved in cases like (C1) and (C2) in section III.

IX. PROCESS VERSUS TREND IN EVOLUTION

We turn now to some considerations of causality and process. We have been insisting that natural selection is a mathematical aggregate of individual events. This seems to imply that it is not a causally connected process. The increase of fitness in a population by means of natural selection is, as we see it, a temporal sequence of aggregates, a running tally of lower-level events; it is not a process within which the earlier events cause the later. The same holds for the spread of adaptive features through a population. This may seem counterintuitive. 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 us 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. Closer examination shows, however, 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 spatiotemporal continuity required of *fundamental physical processes*, which are strictly law governed. To halt, reverse, delay, or accelerate a fundamental process takes energy and work. But in the pot of water such changes of direction occur spontaneously. This shows that heat flow is not a fundamental process. The same point can be made (and historically was made) by reference to *reversibility*.[34] Played backward, a fundamental physical process conforms to the laws of physics; heat transfer does not. You cannot get heat to flow spontaneously back from the boiling water into the stove top by reversing thermodynamic

---

[34] An accessible account of reversibility can be found in P. W. A. Atkins, *The Second Law* (New York: Freeman, 1984).

determinants, as you can get a particle to retrace its path by reversing all the momenta in its ensemble.

In physics, discontinuous processes of the above kind are treated as aggregates of fundamental processes. 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. But 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 which comprehend more are "more probable," those which 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 entities participating in fundamental processes. This explains why statistical thermodynamics has "universal applicability," as Sklar says (see the end of section v, above): it is 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 a trend predicted by the mathematics of large numbers, not a fundamental process governed by the specificities of physical law.

The same is true of evolution. The discontinuities of natural selection are much more evident to the observer than those of thermodynamics, gross enough, in fact, to be recordable by a careful observer.[35] These discontinuities show that natural selection, like thermodynamic change of state, is a time-asymmetric statistical trend instantiated by populations. In an important sense, the mathematical-statistical treatment of natural selection eliminates evolutionary process. That is, it shows that earlier and later states of population fitness or adaptedness in an evolutionary process should not be considered stages in the unfolding of a fundamental process.[36] This is what the

---

[35] For a famous and vivid study of discontinuities and reversals in selection, see Peter Grant, *Ecology and Evolution of Darwin's Finches* (Princeton: University Press, 1986).

[36] In "Chasing Shadows: Natural Selection and Adaptation," *Studies in the History and Philosophy of Biology and the Biomedical Sciences*, XXXI (2000): 135-53, Denis Walsh

hierarchical-realization scheme is intended to capture: natural selection is the consequence of heterogeneous processes in substrates sharing only a formal structure. 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 predictive fitness values are predictors of trends in populations, and may thus be considered *probabilistic causes*, they are not causes in the sense appropriate to fundamental processes.

We should distinguish this conclusion from two other propositions with which it might be confused. First, some will find that our confining fundamental 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; 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.[37] 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

---

argues, on grounds independent of those given here, that heat transfer and natural selection are "pseudo-processes."

[37] This is so even when the mathematical reason for trends is not explicitly stated. Consider the computer simulations used by Peter Danielson in *Artificial Morality: Virtuous Robots for Virtual Games* (New York: Routledge, 1992), and Brian Skyrms in *Evolution of the Social Contract* (New York: Cambridge, 1996). Ariew discusses the explanatory value of statistical patterns in "Are Probabilities Necessary for Evolutionary Explanations?" *Biology and Philosophy*, XIII (1998): 245-53.

interest of these patterns is not simply historical reconstruction, any more than it is in thermodynamics.[38]

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 distinguished two kinds of causal relations, fundamental and stochastic. 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 a cause 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 our opinion, this is misguided. It relies on the idea that fitness and natural selection are separable causes of evolution. We have argued for a hierarchical-realization model of selection, in which fitness is a probability defined on the class of selection histories that meet all the relevant explanatory conditions. Further, we have argued that the nature of causal relations is different at the e-level, where natural selection occurs, and the underlying i-level. Fitness and natural selection have no reality except as accumulations of more fundamental events. In our conception, it makes dubious sense to hold evolution is different from natural selection, much less that it is caused by it.

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. Sklar's discussion of thermodynamics is once again revealing here. Speaking of temperature, he 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.[40]

---

[38] Ariew makes similar points in "The Ultimate/Proximate Distinction" (in preparation).

[39] Here we concur with Sober, "Physicalism from a Probabilistic Point of View," *Philosophical Studies*, XCV (1999): 135-74, especially pp. 145-48.

[40] "The Reduction(?) of Thermodynamics to Statistical Mechanics," p. 194; see also footnote 25 above.

Something like the same problem arises in the theory of natural selection because of the variety of levels at which selection events occur—these range all the way from that of molecules to that of social groups. An example of this lies in the importance of both genetic selection and cultural selection in human evolution.[41] Here, as in Sklar's example, selection is occurring simultaneously in two quite different substrates. The question is: How do we understand the interaction of these two substrates? What is the nature of the equilibria that is reached in these cases? The imposition of such crossing relations on our hierarchical-realization scheme is no small challenge. It may well be that in the last analysis, fitness involves sui generis probabilities estimated not by causal, or even mathematical, analysis, but by census.[42]

MOHAN MATTHEN

University of British Columbia

ANDRÉ ARIEW

University of Rhode Island

---

[41] See P. E. Griffiths and R. D. Gray, "Developmental Systems and Evolutionary Explanation," this JOURNAL, XCI, 6 (June 1994): 277-304; and Susan Oyama, *The Ontogeny of Information: Developmental Systems and Evolution* (New York: Cambridge, 1985).

[42] The differences between physical and statistical ways of thinking about factors relevant to evolution have been investigated independently by Denis Walsh and Matthen. Ariew, who has been working with both, brought about considerable convergence both by cross-communicating and by his own critical contributions. The present article and Walsh, Lewens, and Ariew (*op. cit.*) are products of this process. But these articles are almost completely independent with respect to argumentation and even, to a surprising extent, motivation.