Selection and Causation*

Mohan Matthen and André Ariew†

We have argued elsewhere that natural selection is not a cause of evolution, and that a resolution-of-forces (or vector addition) model does not provide us with a proper understanding of how natural selection combines with other evolutionary influences. These propositions have come in for criticism recently, and here we clarify and defend them. We do so within the broad framework of our own 'hierarchical realization model' of how evolutionary influences combine.

1. Introduction. In Matthen and Ariew 2002, we argued for the following:

(A) Natural selection is not a cause of evolution. (We are not opposed to Darwin and the modern synthesis—quite the contrary. Rather, we argue that the relationship envisaged by these theories is not a causal one.)
(B) A resolution-of-forces (or vector addition) model does not provide us with a proper understanding of how natural selection combines with other evolutionary influences such as drift, mutation, migration, etc.
(C) Our own 'hierarchical realization model' offers a better understanding of the interplay of these evolutionary influences.

Propositions (A) and (B) have come in for criticism recently, and our aim in this article is to respond to these criticisms. We want to do this in a systematic way, however, not piecemeal. Thus, we structure our replies with an eye toward our positive view (C), which we adumbrate and extend as we proceed.

2. Trait Variance and Causation. The first thing to do is mark the difference between (A)—the claim that natural selection is not a cause of

*Received June 2008; revised May 2009.

†To contact the authors, please write to: Mohan Matthen, Department of Philosophy, University of Toronto, 170 Saint George St., Toronto, ON M5R 2M8, Canada; e-mail: mohan.matthen@utoronto.ca. André Ariew, Department of Philosophy, 434 Strickland Hall, University of Missouri–Columbia, Columbia, MO 65211; e-mail: ariewa@missouri.edu.

Philosophy of Science, 76 (April 2009) pp. 201–224. 0031-8248/2009/7602-0005$10.00

evolution—and a different, but superficially similar, proposition. Consider a common example. A population of moths contains two variants, a dark-colored kind *D* and a light-colored kind *L*. *D*-type moths are well-camouflaged in a sooty environment in which the tree trunks on which they sit are darkened by soot. *L*-type moths are not well camouflaged in this environment. Consequently, *D*-type moths are less subject to predation in a sooty environment and their frequency increases in the population compared to *L*-type moths. In a cleaner environment, *L*-type moths would have had the advantage with regard to camouflage and would have increased in frequency.

In this case, it is clear that the variation in camouflage causes evolutionary change in the moth population. This can be seen by appealing to various tests that philosophers standardly use to identify causal relations.

The probability raiser test (T1) says that

> *C* causes *E* only if the C raises the probability of *E*. (An 'only if'—necessary—condition is not a test, but we shall ignore this here, since our main critical concern is with situations where the condition is *not* met; hence something is found *not* to be a cause.)

Applying this to the case: *D*-type moths are better camouflaged than *L*-type moths in a sooty environment. Since better-camouflaged moths are more apt to be picked off by predators, this variation in camouflage raises the probability of *D* increasing in frequency relative to *L*. Thus, this test does not rule against the claim that variation in camouflage causes the increase in the frequency of the *D*-type moth—which is an evolutionary change.

More tellingly, consider the manipulability test (Woodward 2003). (We use a version of this test that is appropriate to probabilistic causation, and where cause and effect are variables that take numerical values.)

Manipulate *C* 'exogenously', that is, do not just examine cases where *C* takes on a different value in normal contexts where the whole nexus of surrounding variables also changes. Rather, create such a case by actively manipulating the value of *C* while holding everything else the same.[1] Now observe what happens to *E*, a putative effect of *C*.

The manipulability test (T2) says that

1. Suppose, for example, that you are testing whether falling atmospheric pressure or falling barometer readings cause wind. According to the proponents of the manipulability test, you don't want to examine cases where both the barometer readings and the atmospheric pressure are (as normal) falling together. Rather you want to put the barometer in a pressure-controlled box so that its readings can be held at a low value independently of atmospheric pressure. If, when the barometer reading has been manipulated in this way, a wind rises up, one may conclude that low barometer readings cause wind.

> *C* causes *E* if and only if we can manipulate the probability-distribution of *E* by means of some exogenous manipulation of (or 'intervention' on) *C*.

Suppose that we exogenously manipulate the camouflage advantage that *D* enjoys by painting all the trees white. Presumably, this intervention would reduce the probability that *D*s increase in proportion. By the manipulability test, this would imply that variation in camouflage causes the *D*-type to increase.

These tests give us excellent reasons for endorsing

> (Â) *Variation in camouflage* causes evolutionary change (of the moth population).

Nobody today doubts this result, least of all us, though (as we will show in Sections 5 and 6) there are some puzzles surrounding the exact nature of this causal relationship, and its proper analysis turns out to be philosophically interesting.

The reason why we have dwelt on this obvious fact is that the proposition that we have just endorsed is not the same as

> (Ā*) *Natural selection* causes evolutionary change.[2]

And it is, of course, (Ā*) that is required to show that (A) is false.

The difference between (Â) and (Ā*) is not marked by some of our critics. Roberta Millstein (2002, 2006), who claims that natural selection is a population-level cause, contents herself with arguing in support of propositions like (Â)—which, we repeat, we do not contest. In her article, Millstein employs tests of causation just like the ones mentioned above. Her aim is to show that natural selection and drift are causes of evolutionary change. To establish her point, she considers various tests of the form '*C* causes *E* if and only if *R*(*C*, *E*)'. But she does not show us how the term 'natural selection' is supposed to take the place of '*C*' in these formulae. Millstein says that she wants to establish a 'causal basis' for evolution. We have no quarrel with this—(Â) is an example of just such a causal basis. Our claim is that natural selection is the wrong sort of thing to play a causal role.

We hope it is clear that we are concerned with an ontological question, not a scientific one. We do not want to debate whether variation in camouflage actually caused evolutionary change, or whether differences in fitness-relevant traits actually played a role in evolution. We are rather concerned with the meaning and ontological commitments of the claim

2. We mark propositions we contest with an asterisk.

that natural selection caused evolution. What is natural selection and how does it influence gene or trait frequencies?[3]

3. Reifying Natural Selection. In this section, we consider some preliminary reasons for being puzzled by the claim (Ā*), that natural selection causes evolution. We do this by once again applying the tests mentioned above, but this time to natural selection. First, the probability raiser test. Does the operation of natural selection raise the probability that the *D*-type moths will increase in frequency? If not, it is not a cause. Second, is there a way of manipulating the probability that the *D*-type moths will increase in frequency by intervening on natural selection? If not, it is, once again, not a cause.

Before we can tackle these questions we need to address a prior question. What exactly is natural selection? What entity or process are supporters of (Ā*) talking about here as a cause of evolution? In this section, we assume Sober's fairly orthodox treatment (1984, 21–22): natural selection is evolution due to heritable variation in fitness.[4] To this, we add an important clarification adapted from Millstein (2002) and Christopher Stephens (2004).

Millstein and (following her) Stephens make a distinction between drift-as-process and drift-as-product (or outcome).[5] Let us say for present purposes that evolution is a longterm change in gene frequencies in a population.[6] For Millstein and Stephens, drift-as-product is the culmination of a drift-driven evolutionary process, the new gene frequency that results. Drift-as-product is not a cause of evolution; in this conception, it is the result of a certain kind of evolutionary process. Drift-as-process, on the other hand, is supposed to be a causal process that drives a population to a new gene-frequency level. "It is the process notion of drift that is needed to understand its role as a cause," Stephens says (2004, 557).[7] He

3. We see our argument as opposing Elliott Sober's in Chapters 1 and 3 of *The Nature of Selection* (1984)—Sober's argument is ontological in the sense intended here.

4. Millstein has a related but subtly different account, to which we shall return later.

5. Millstein uses the term 'outcome', but we like the resonance of the traditional process-product distinction.

6. Gene frequencies fluctuate from generation to generation. Our concern is with the overall trend line: evolution occurs when there is a change in gene frequencies after short-term fluctuations have been smoothed out.

7. Godfrey-Smith (2007) seems to overlook this distinction. He observes that there are ways of estimating drift and uses this to admonish us for our skepticism about drift-as-process. He fails to notice that he is appealing to methods of estimating drift-as-product.

goes on to suggest that drift-as-process is a variable-strength force that drives evolutionary change. According to Millstein in particular, drift-as-process is something over and above the sequence of births, deaths, and matings that add up to that new frequency; it is a motor that acts on a population, driving it to new gene-frequency rates. (Millstein holds that drift-as-process and selection-as-process are two different motors working on a population; later we'll comment on this briefly.)

Sober holds that natural selection is evolution due to heritable variation in fitness. By analogy with Millstein's notion of drift-as-process, it is selection-as-process that he is talking about. Selection-as-process would be not merely the sequence of events leading from variation in heritable fitness and culminating in a new gene frequency, but a variable-strength force that appears or is generated when organisms have heritable differences that lead to differential reproductive success—a motor that drives fitter types to greater frequencies within a population. The greater the variation in these traits, the more powerful the motor.[8]

Sober appears to take the causal efficacy of this motor quite literally: "Selection for is the causal concept par excellence. . . . An organism's overall fitness does not cause it to live or die, but the fact that [for example] there is selection against vulnerability to predators may do so" (1984, 100).

Selection for a trait causes animals to die, Sober says.[9] (Stephens [2004, 561] endorses the thought.) Since causes and effects are distinct, it follows that

(1) Selection-for-trait-T $\neq$ the death of animals.

Again, fitness (unlike selection-for-trait-T) does not cause anything to die, according to Sober, and presumably he thinks that the same is true of variations in fitness. So,

(2) Selection $\neq$ variation in fitness.

8. Robert Brandon (1978) advances a variation on this idea. He thinks that fitness is a differential strength propensity in individuals, which drives them to reproduce in different numbers. Population-level changes in frequencies of types result from the differing fitnesses of individuals, according to Brandon. See also Bouchard and Rosenberg 2004 and Rosenberg and Bouchard 2005. The argument that we provide against Sober's view could be modified to take in Brandon's view as well.

9. Well, he says that it may do so, but this should be taken in the context of the emphatic first sentence of the quotation. Actually, we suspect that Sober slipped, and did not mean to say that selection for caused an animal to die, or even that the-fact-that-there-is-selection-against-vulnerability-to-predators did this. But since Stephens takes off from this thought, and since, further, we are not sure what Sober really meant to say, we will simply let the quotation stand.

Obviously, also,

(3) The death of animals $\neq$ variation in fitness.

These propositions imply that there are three things here. Evolution does not occur merely because types of organism reproduce and die in differing amounts. Rather it is driven by natural selection. Natural selection is a *tertium quid* on Sober's account,[10] an intervening variable that drives the process. The causal diagram in evolution-by-selection would go like this, according to his account:

> heritable variation in trait $T \rightarrow$ selection of magnitude proportionate to variance in heritable fitness due to heritable variation in $T \rightarrow$ birth and death of animals $\rightarrow$ evolution.

If the quote from Sober above is to be taken literally, selection is a Grim Reaper, killing a weak organism here, preserving a strong one there.[11] And Stephens would have drift be its Grimmer Brother, killing at random. (Of course, selection and drift have a more erotic side also—they cause animals to mate and reproduce.) Millstein, for her part, believes that selection is a process by which heritable variation leads to evolution—heritable variation in fitness is the input to this process—but she believes that it acts on populations, not individuals.

Does it make sense to think of selection in this way—that is, as a *tertium quid* that drives evolution? The manipulability test of causation employed above directs us to envisage situations in which we intervene on natural selection in the above causal diagram, leaving heritable variation in T and the environment unchanged. The question we are supposed to ask is, would the probability of evolution in the moth population due to variation in camouflage change if we changed or eliminated the intervening variable, selection? The answer is surely no. For if the *D*-type and *L*-type moths have differential reproductive success, the probability of evolutionary change will be high even in the absence of any such *tertium quid*. This outcome does not depend on the existence or the magnitude of a *tertium quid*. Take a coin that is biased towards heads. The bias implies that it is probable, in a series of tosses of this coin, that heads will come up more often than tails. No process of 'toss selection' is needed for this result. Similarly, in the presence of heritable variation, it is math-

10. Actually, it is selection for and selection against that are so treated, not selection as such—we will take this as read.

11. Brandon (see note 5) is, for his part, committed to the causal efficacy of individual fitnesses. He and Ramsey claim in a current project that, in effect, it is a cause of births over and above causes such as sexual intercourse, fertilization, etc., and similarly a cause of death over and above disease, predation, etc.

ematically necessary that evolutionary change is probable.[12] No intervening variable is required.

For exactly the same reason, the probability raiser test also fails: given a certain level of heritable variation in fitness, increasing or decreasing the strength of intervening variables does not change the probability of evolution taking place. (Shapiro and Sober object to this argument; we'll turn to their objection presently.)

It should by now be clear at least that the following inferential slide by Millstein is problematic: "If . . . we introduced a new beetle genotype into the population that was able to withstand a greater range of temperature than any of the existing genotypes, we would expect that the reproductive successes of the other genotypes would decrease. . . . Thus, selection satisfies a manipulability account of causation" (Millstein 2006, 633; emphasis added). Millstein argues that manipulating variation in temperature resistance would result in a change of genotype frequencies. She infers that manipulating selection would result in the change of genotype frequencies. This is precisely the move that we oppose.[13]

4. The Force-Body Paradigm. Millstein moves from (4) to (5*):

(4) Variation with respect to trait *T* caused evolutionary change in population *P*.

(5*) Natural selection caused evolutionary change in population *P*.

How does she think that this inference is to be justified?

In an earlier article, Millstein says that random drift is "an indiscriminate sampling process in which physical differences between organisms are causally *irrelevant* to differences in their reproductive success" (2002, 171). And natural selection, by contrast, is supposed to be a discriminate process in which physical traits are relevant to differential reproduction rates. In other words, she holds

12. According to J. H. Bennett, R. A. Fisher wrote to his publisher of his own "impudence in treating the subject as a branch of mathematics" ([1930] 1999, ix). Fisher was well aware, in other words, of how he had placed a mathematical theorem at the heart of his treatment of evolution. Some (e.g., Brandon) think that this trivializes the subject. We agree with Sober that on the contrary, treatments of Fisher's ilk, though mathematical, illuminate an aspect of the causal structure within which evolution is embedded, and should hence be regarded as explaining evolution. "Explanation, to be illuminating, need not single out the events that actually did the causing," Sober says (1984, 96).

13. In Matthen and Ariew 2002, 62, we say, eyesight leads to differences between [phenotype fitnesses]—but not drift or 'neutral' selection. We meant to be clear throughout that it is selection that is being denied causal status, not traits such as good eyesight.

(6*) Natural selection is the process by which variation in fitness-relevant traits causes evolutionary change in a population.

(6*), together with the uncontested proposition (4), implies the contested proposition (5*). Note that there is a certain similarity between the causal diagram implied by (6*) and the one that we encountered earlier in Sober and in Stephens. There may be a difference between the two positions, however, because Millstein believes that selection operates on populations.

One reason for reifying natural selection in this manner lies in a (faulty, as we shall argue) analogy between equations of population genetics—such as Fisher's Theorem—and certain equations of physics. The equations of population genetics describe the behavior of ensembles of organisms in terms of ensemble-level parameters.[14] For example, Fisher's theorem relates the variance of fitness in an ensemble to the rate and magnitude of increase of its average fitness. Similarly, the equations of physics describe the behavior of particles in terms of particle-level parameters.

Think, for example, of Newton's Law of Gravitation: $F = G(m_1^* m_2/r^2)$. This law governs a certain interaction between pairs of massive particles. But most physicists think that this interaction is mediated by a *tertium quid*—the force of gravitational attraction. In other words, they hold that

(7) A massive particle P_1 caused another massive particle P_2 to move.

is true in virtue of

(8) The gravitational attraction exerted by P_1 caused P_2 to move.

14. We prefer the term 'ensemble' to 'population' here. Millstein adopts Futuyma's definition of population as "a group of conspecific organisms that occupy a more or less well-defined geographic region and exhibit reproductive continuity from generation to generation" (Millstein 2002), and she seems to think that in mathematical 'population' genetics, variables such as size, variance, and growth rate are defined over populations in this sense. This is not exactly correct. A lot of the discussion within population genetics proceeds in quite abstract terms, and the restriction to populations is unnecessary. Theorems such as Fisher's Fundamental Theorem or Price's equation apply indiscriminately to any set of organisms, conspecific or not, connected by region and ancestry or not. Consider a set S consisting of ten organisms randomly drawn from each mammalian species. S is not a population in Futuyma's sense; it fails all of the tests in his definition. Nevertheless, the average fitness of S will increase, in accordance with Fisher's Theorem, in proportion to its additive genetic variance in fitness. Of course, there are some results in population genetics that apply only to populations in Futuyma's sense: for instance, the disappearance of an allele from S does not imply fixation of the alternative allele. Similarly, species and speciation depend on such populations. All the same, one should not lose sight of the purely mathematical and abstract side of population genetics.

Thus, physicists diagnose gravitation in terms of the following causal diagram:

> mass of two particles → gravitational attraction of magnitude proportionate to product of masses → motion.

Let us call this the 'force-body paradigm'. In physics, this paradigm looks for forces such as gravitation to mediate the interactions of particles.

This is where the analogy between equations of physics and of population genetics comes into the picture. Millstein, in common with many biologists, seems to apply force-body paradigm to populations. She moves from variance with respect to some trait T influencing evolutionary change to a force that mediates the influence that variance with respect to T has on evolutionary change. This intermediary, she contends, is natural selection. Likely, this kind of reasoning is the ultimate source of the causal diagram that we attributed to Sober and Stephens in Section 3.

In our opinion, the force-body paradigm is misapplied to natural selection and drift. It is a background assumption of most philosophical theories of causation that logical or mathematical implications do not fall under the notion of causation. Manipulate a man's marital status while leaving all else the same, and you will change whether he is a bachelor. (For that matter, manipulate his marital status, and you will change his marital status.) But surely one does not want to say that being unmarried causes him to be a bachelor (or that his marital status causes his marital status). Being unmarried is for him no different from being a bachelor. Similarly, intervene on the number of items in a collection by adding one member, leaving all else the same, and you will change whether that collection is odd- or even-numbered. But surely it would be a mistake to think that number is a cause of being odd or even. Causes have to be distinct from their effects.

It is precisely in order to avoid errors of this sort that Hausman and Woodward say of the manipulability test: "Mechanisms are distinct if and only if it is in principle possible to interfere to disrupt one while leaving the other alone. . . . Causation is connected to manipulability and that connection entails that separate mechanisms are in principle independently disruptible" (1999, 539–40). In the same vein, Hall says: "I will also take it for granted that we can adequately discern when two events fail to be wholly distinct—that is, when they stand in some sort of logical or mereological relationship that renders them unsuited to stand in causal relationships" (2004, 226). The force-body paradigm works in physics. Gravitation satisfies the tests for causation, and is, moreover, distinct, in the sense of the two quotes given above, from the interaction of the two particles. However, natural selection is not in the requisite sense distinct from heritable variation of fitness.

Let's look at the gravity case first. In order to test, for example, whether the gravitational attraction exerted by the moon is the cause of the motion of the tides, one must hold the masses of these bodies constant, while intervening on gravitational attraction between them. At first sight, the test works. Suppose that we cranked up the gravitational attraction of the moon while holding its mass and the mass of the oceans constant. Then the tides would be higher.

Now, it might seem misguided to speak of 'cranking up' gravitation in this way. Since the gravitational attraction between two massive particles supervenes on their masses, one might think that one cannot manipulate gravitational attraction in the manner proposed. The only way to manipulate gravitational attraction is to change the masses that are attracting one another. Thus, one might want to understand the 'hold everything else the same' or 'independent disruption' stipulation of the manipulability test liberally, in order to permit changes to a supervenience 'base' when one wants to intervene on the supervenient property. This, at any rate, is how Shapiro and Sober (2007) would have us proceed. So to test whether the moon's gravitational force influences the tides, Shapiro and Sober would presumably have us change gravitational attraction by changing the masses. They go on to suggest that when we use the manipulability test to check whether natural selection causes evolutionary change, we should be permitted to change items in the supervenience base of natural selection—individual births and deaths, variation in fitness-related traits, and so on.

We do not agree that loosening the 'hold everything else the same' condition is the right way to address the supervenience problem. Let's start with the case of gravitation. How would you test whether gravitational attraction is an intervening cause in between mass and motion? If you intervene on the masses of the moon and the oceans in order to intervene on gravitational attraction, all that you test is whether these masses influence the movement of the tides, not whether gravitational attraction is required as an intermediary. So you have to hold these masses the same and intervene on gravitation. This does not pose as much of a problem as Shapiro and Sober suppose—at least not for the philosopher. For as Woodward argues in effect, it is possible to intervene on gravitation—if you understood 'intervention' and 'manipulation' properly. He writes:

> Consider, for example, the (presumably true) causal claim (G):
>
> (G) The gravitational attraction of the moon causes the motion of the tides.
>
> Human beings cannot at present alter the attractive force exerted by

> the moon on the tides (e.g., by altering its orbit). More interestingly, it may well be that there is no physically possible process that will meet the conditions for an intervention on the moon's position with respect to the tides. . . . It is nonetheless arguable we have a principled basis in Newtonian mechanics and gravitational theory themselves for answering questions about what would happen if such a surgical intervention were to occur and that this is enough to vindicate the causal claim (G). (Woodward 2008)

Woodward proposes, rightly in our view, that "a properly formulated version of a manipulability theory will thus allow us to talk about causal relationships in some contexts in which interventions are not physically possible" (2008).

We believe that it is possible to envisage a 'surgical intervention' on gravitation itself. Accordingly, we would like to say (paraphrasing Woodward) that we have a principled basis in Newtonian mechanics for answering questions about what would happen if we (or God) were to surgically intervene on the relationship between mass and gravitational attraction. If we made it, for instance, an inverse-cube relation instead of an inverse-square relation, or changed the gravitational constant, then we would change the gravitational force exerted by the moon without changing the mass of the moon or of the ocean. Newtonian mechanics would lead us to expect that in such a case, the motion of the tides would change. This is why we say that according to Newtonian mechanics, gravitational attraction is an intermediary between mass and motion—a *tertium quid* as the force-body paradigm urges. Physically impossible intervention is the proper analytic tool for supervenience cases, not loosening the 'hold everything else the same' clause.

As implied by the earlier discussion, the theory of natural selection is different from that of gravitation, for natural selection is mathematically necessary. When there are heritable differences in traits leading to differential reproduction rates, the probability of the fitter types increasing in frequency is greater than that of the less-fit types increasing. This is simply a mathematical truth. Applying the manipulability test, we have to tolerate hypothetical interventions that are physically impossible, but refuse to tolerate hypothetical interventions that are mathematically impossible. This is why the manipulability test cannot properly be applied to

> (5*) Natural selection caused evolutionary change in population *P*.

There is no room for a manipulable intermediary between advantageous heritable traits and the increase in frequency that results from their being advantageous. Such an intermediary would not be distinct from the increase in frequency. The analogy with gravitational attraction is flawed.

This is the point that Stephens (2004), Kenneth Reisman and Patrick Forber (2005), and Shapiro and Sober (2007) all miss. Stephens observes that drift is dependent on population size—that is, that the smaller the population, the less predictable the result of natural selection. He imagines the following situation:

> Suppose there are two sets of populations, each with two kinds of otherwise identical individuals. One kind of individual has a trait (T_1) is fitter that the alternative (T_2). In one population set, there are 20 (isolated) groups each with 6 individuals, whereas in the other set there are 20 (isolated) groups each with 1,000 individuals. Suppose that every group in each population set begins with exactly 50% T_1 and 50% T_2. Imagine that the population evolves, and, as expected, the number of groups in which T_1 goes to fixation in the second population set is much higher than in . . . [the first]. . . . *Why*? Answer: drift. (Stephens 2004, 564)

Along much the same lines, Reisman and Forber argue that drift is a cause of evolutionary change because when one manipulates drift by changing population size one thereby changes the probability value and/or magnitude of evolutionary change. Finally, Shapiro and Sober say:

> Selection and drift are distinct processes whose magnitude gets measured by different population parameters (fitness on the one hand, effective population size on the other). And changes in each of these parameters will be associated with changes in the probabilities of different outcomes. If you intervene on fitness values while holding fixed population size, this will be associated with a change in the probability of different trait frequencies in the next generation. And the same is true if you intervene on population size and hold fixed the fitnesses. (2007, 262)

What these critics miss is that the connection between population size/variation-in-advantageous-traits and drift/selection is purely mathematical. This connection is the same as that which holds between sample size and proportional variance from the mean in random sampling. Nobody would say that the numerical ratio between 4 and 100 is the cause, in any literal sense of 'cause', of there being a higher proportional variance from the mean when sets of 4 are randomly sampled than when sets of 100 are so sampled. Sample size and variance under random sampling are connected by a mathematical law, and thus they are not sufficiently distinct from one another to count as terms in a cause-effect relationship. Certainly, it is right to say that selection and drift are part of the explanatory apparatus used to explain evolutionary change. But as Sober says, "Explanation, to be illuminating, need not single out the events that actually

did the causing" (1984, 96; see our note 9). In other words, selection and drift can be explanatory (and thus answer Stephens' '*Why?*') without being causes of evolutionary change. This is what Stephens and the others fail to understand.

5. Causation and Locality. We said earlier that we do not contest on general principle propositions of the following form:

(4) Variation with respect to trait T causes evolutionary change in population P.

That some propositions of the form (4) are true and account for evolution should not be a point of contention between our critics and ourselves. There is, however, an interesting complication as to the nature of the causal link asserted by (4), and we shall go into this now.

To appreciate the nature of the complication, let us look first at the following simple example.

CM Example.

i. Let S be a system consisting of two particles of equal mass. At time t these particles are moving in exactly opposite directions at exactly the same speed. At t, therefore, the center of mass (CM) of S is stationary.
ii. Now, suppose that at time $t + d$, one of the two particles collides with a barrier and starts to move in the opposite direction. Since the particles are now moving in the same direction, the center of mass (CM) of S begins to move. This change occurs at $t + d$; that is, it is simultaneous with the reversal in the particle's direction of motion, though the two events are separated in space.

It seems clear that in the CM example the collision at $t + d$ has two effects. First, it causes the colliding particle to change its direction of motion. Second, it also causes the CM to change its direction of motion. Now, this second causal link seems at first glance to violate a certain condition on causation, to wit:

Locality Principle. Causal influence is transmitted from one physical entity to another in a spatiotemporally continuous manner—that is, by every point in between the two entities being affected.

It seems, that is, that a causal influence was transmitted from the collision (or from the barrier) to the CM instantaneously, and without any changes in the medium in between the barrier and the CM. This, moreover, seems

to violate a fundamental principle of the Special Theory of Relativity—namely, that causal influence cannot be transmitted from one physical entity to another at a speed faster than that of light.

It may seem that way on the surface, but in fact the Locality Principle is not violated. The key to seeing why is to observe that the CM is not a physical entity—it has no mass, for one thing. It is a derivative entity that depends, as a matter of mathematical necessity, on other physical entities—to wit, the two particles that make up system *S*. To influence the CM it is (mathematically) necessary and sufficient to influence these two particles. It is not necessary to transmit causal influence from the collision site to the CM in order to get the latter to move. (Indeed, it makes little sense to talk about a physical cause acting on a nonphysical, or 'fictional', entity such as the CM.) One should, therefore, distinguish between causal transmission, which must respect the Locality Principle, and a more inclusive sense of causation, which need not do so. In the more inclusive sense of causation, if x causes y, and y mathematically implies z, then x inclusively causes z, even though the connection between x and z is nonlocal.

Now, let us take an example from the theory of natural selection. An (unfit) animal is caught in a predator's jaws and killed. The squeezing of the predator's jaws causes the death of this organism, and in so doing immediately affects the mean fitness of a population. There are two causal connections here: the first between the squeezing of the jaws and the death of the prey, the second between this same cause and the change of fitness. The first of these connections respects the Locality Principle. The second, between the squeezing and the change of mean fitness, however, does not. This is not an anomaly: it can be accounted for in exactly the same way as the CM example. Mean fitness is not a physical entity; rather, it is ontologically dependent on physical entities. Thus, it is not the kind of change that has to be mediated by causal influence that is transmitted from one place to another. In fact, it does not even make sense to ask where the mean fitness of the population is, much less to ask how fast the influence was transmitted to it from the squeezing of the predator's jaws. Along the same lines, suppose that a new mutation occurs that protects smokers from cancer. This too will affect the mean fitness of the population instantaneously: the link between these events is not a process that has to travel from one place to another with finite speed. In each case, the relation of nonlocal causation is the logical product of the relation of local causation and the relation of mathematical implication. (See Figure 1.)

Stephens contests the account just given: "It is far from clear that [the Locality Principle and other such conditions] are requirements on fundamental physical processes. My contention is that these are ad hoc re-

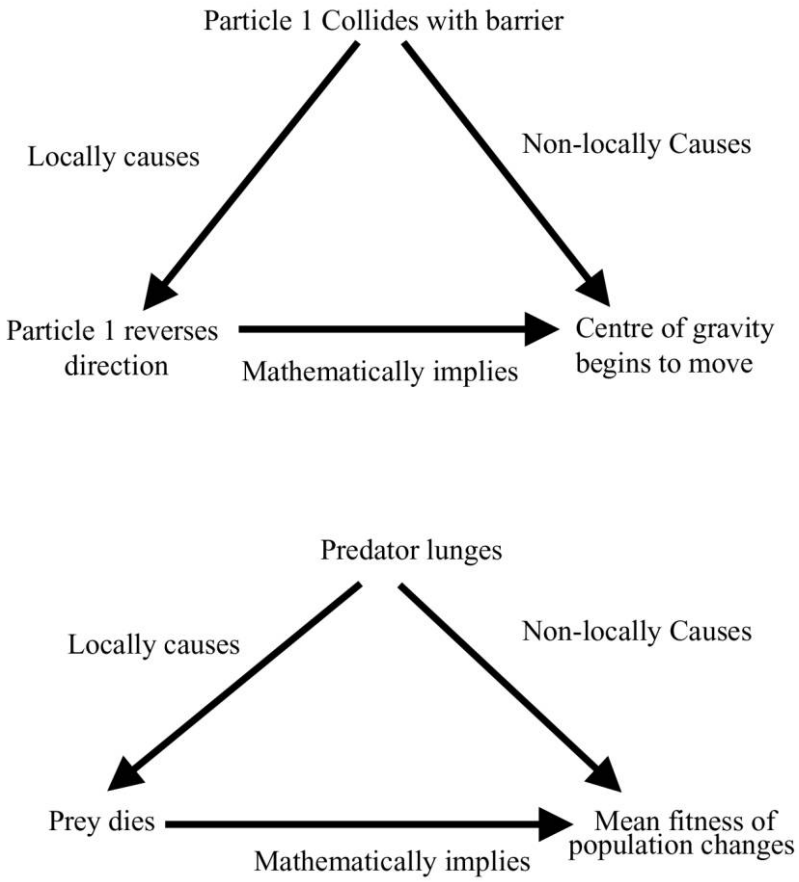

Figure 1.

quirements. What if one of our fundamental physical theories turns out not to meet these criteria? . . . Quantum mechanics throws doubt on [such criteria], since there are important senses in which quantum phenomena is [*sic*] discontinuous" (Stephens 2004, 567–568; Millstein endorses and relies on this argument as well).

This strikes us as completely unmotivated. Perhaps quantum processes do, as Stephens claims, violate the Locality Principle and other classical principles of causation. But is he really suggesting that the CM example and the changes of mean fitness discussed above involve fundamental physical processes that violate the Locality Principle and other such principles? Here, we recall Hall's witty remark: "Is this all it takes to achieve

nonlocality? (And to think that philosophers have been fussing over Bell's Inequalities!)" (2004, 236).[15]

The account that we have just offered—that CM and mean fitness are nonphysical derivative entities, affected by nonfundamental causal connections—is not only cribbed from treatments in classical statistical theory of gases (Matthen and Ariew 2002, 79–80) but completely conforms with biological orthodoxy. Natural selection is ontologically derivative on individual-level events such as births, deaths, and mutations—even Shapiro and Sober are committed to this by their supervenience claim. It occurs in ensembles as a mathematical consequence of events that involve members of those ensembles. The latter events are brute physical occurrences—they occur by the transmission of physical influence from one place to another.

6. Why Selection Is Not a Force. Once the force-body paradigm has been applied to natural selection, another idea follows rather naturally. It is a commonplace idea of evolutionary theory that evolutionary change is caused not only by differences of heritable fitness, but also by drift, mutation, genetic constraints, and migration. If you think that natural selection is a force that mediates differences of heritable fitness, then you might think that drift, etc., are forces that act in different directions, thereby opposing or assisting natural selection, or taking evolution in different directions.

Elliott Sober is one person who has developed this idea in some detail. The opening chapter of his classic book *The Nature of Selection* (1984) is entitled "Evolutionary Theory as a Theory of Forces." We have criticized this conception in some detail (Matthen and Ariew 2002). We assume that forces are vectors—entities that have magnitude and direction (though the direction does not have to be spatial in a literal sense). A theory of forces is, in essence, theory of how these vectors combine when several of them are acting together. We argued that evolutionary theory has no zero force law, no singleton force models, and most importantly, no law of resolution of forces. We will not repeat these arguments here,

15. Stephens is not the only one of our critics who posits entirely new and mystical forces of nature. Brandon and Ramsey (2007) say, "The propensity interpretation of fitness is committed to the fundamental indeterminacy of the lives, deaths, and ultimately reproductive successes of individual organisms." The term 'fundamental' suggests that the indeterminacy of these events is like the indeterminacy of quantum processes—no hidden variables will explain them. Along the same lines, it has sometimes been remarked to us that in General Relativity, there are causal influences that fail the Locality Principle. The question, once again, is: do we really need these exotic kinds of causation just to understand what is happening in the simple cases we are describing?

but we do want to reply briefly to some criticisms, and then to conclude by restating our alternative.

Our basic reason for denying the 'resolution of forces' model was this. We have no way of estimating from the theory of evolution how two factors combine. We said, "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? (Matthen and Ariew 2002, 67). Stephens has two objections to this line of thought. The first is the following: "Empirical observation is how one determines how these factors 'add up'. If we sample the members of the population, the estimated fitnesses may tell us that the traits interact—the relationship between these various possible traits may be additive or nonadditive" (2004, 553). Actually, Stephens is simply rephrasing our own point.[16] Suppose that a force of one dyne is acting eastward from a point, and a force of two dynes northward. The magnitude and direction of the resultant force is determinate. (By the 'parallelogram of forces', the resultant acts in a roughly north, north-easterly direction with a magnitude of $\sqrt{5}$ dynes.) The resultant force is a mathematical function of the components. But suppose that two organisms of species S that differ by x in strength will differ by y in fitness. Suppose further that if they differ by a in speed, they differ by b in fitness. Now suppose that two organisms differ by x in strength and a in speed. By how much will they differ in fitness? This is the kind of question that probability theory and population genetics cannot answer.[17] It isn't that there is no answer in particular cases—rather, it is that there is no general way of combining probabilities, unless the occurrences are independent of one another. The joint fitness is simply not a function of the separate fitnesses. Just as the probability of p & q depends not just on the probability of p and the probability of q, but also on the probability of p given q, so also the joint fitness of two traits depends not just on the individual probabilities, but on a variety of interaction factors at both the genic and the phenotypic levels. This is why one has to resort to 'empirical observation' as Stephens puts it: one has to estimate the interaction factors independently. The first disanalogy with force, then, is that forces combine in a predictable way across the board,

16. Cf. Matthen and Ariew 2002, 67: "The overall fitness values . . . 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."

17. What is the relevance of probability theory here? Population genetics is a calculus of frequencies, which interact in ways formally isomorphic to probabilities.

while distinct evolutionary influences combine case by case. We shall return to the consequences of this point in the following section.

Stephens attempts a second point against us. He says that mutation and selection can add up in the way that forces do: "We can talk about cases where the force of mutation *more or less strongly* opposes selection, and cases where mutation operates in the *same direction* as selection. It makes perfect sense to say that as the force of selection gets stronger relative to an opposing mutation rate . . . the equilibrium frequency . . . gets smaller" (2004, 554–555)

Robert Brandon echoes the point:

> Consider mutation and migration. Both are similar in that they are already net forces that could be broken down into their components. Migration consists of immigration into a population and emigration from it. Both can have effects on gene frequencies, which can then be added together to get the force of migration; similarly for mutation. In a simple system with one locus and two alleles, A1 and A2, there is an actual number of mutations from A1 to A2, and an actual number from A2 to A1. The change in A1 is simply the second number minus the first. There is no problem in conceiving of these as forces. (2006, 322)

Brandon and Ramsey (2007) extend the point. Suppose that we have two alleles A and a. There is a net mutation rate of A to a of m and a net migration rate of A-type organisms of n. Surely we can add mutation to migration in such a case. Stephens and Brandon and Ramsey seem to be giving us a general purpose method for compounding distinct evolutionary influences. Thus, Brandon (2006, 334) concludes: "What I have found surprising is how useful is the comparison" between evolutionary theory and a Newtonian theory of forces.[18]

Stephens and Brandon are wrong about this. Once again, they are overlooking interdependencies. Suppose that the rates of mutation, migration, and selection were dependent on one another. Then we couldn't calculate net change in the simple way that these authors suggest. That is, we could not simply add mutation to migration, or to selection, and assume that we have calculated net change. For as migration occurred, the mutation rate would also be changing. To calculate net change, therefore, we would need to know how mutation varied as a function of mi-

18. Brandon agrees with (but, unfortunately, has never acknowledged) our contention that "drift should not be regarded as a force that can be added to others acting on a population" (Matthen and Ariew 2002, 61). He thinks, as we do, that drift ensues even when there are no differences of heritable fitness, no mutation, migration, etc. Brandon calls this 'Biology's First Law'.

gration, and vice versa—and of course we would need the same quantities for selection as well. Here too, combined change under selection, migration, and mutation is not a function of change under selection alone, migration alone, and mutation alone. Additional independent variables have to be specified.[19]

There is another problem in the Stephens-Brandon approach. In order for the simple additive process they conceive of to work, selection, mutation, and migration have to be cited in terms of genes. If these rates are cited in genic terms, and if they are independent of one another, we can add selection rates to one another and solve the above-mentioned problem of combining strength and quickness.[20] Or we can just count net changes in gene numbers after the evolutionary influences have operated. The problem, however, is that the theory of evolution does not always specify selection and migration rates in terms of net results on allelic frequencies. For the theory of evolution is not just interested in alleles: it is also interested in adaptation and trait-selection (Lewontin 1974, introduction; Walsh 1998). Differences of trait fitness are responsible for changes in gene frequency: differences with respect to these traits are responsible for differential reproduction rates, and the latter account for changes of gene frequency. As well, phenotypes are at issue in discussions of selection for, frequency-dependent selection, assortative mating, reproductive strategies, body plans, and so on—and even migration. What would it mean, then, for only gene frequencies and not trait frequencies to be at issue? We would have to ignore trait interactions and look directly at genes. Goodbye, Darwin. So long, Modern Synthesis.

If you base your theory of forces simply on the addition of gene frequencies, you ignore the trait-based or phenotypic factors that are at the heart of evolutionary theory. Evolutionary theory is a two-level theory—it deals with both genotypes and phenotypes. The moment phenotypic properties enter the picture, the resolution of forces becomes problematic. We concede to Stephens and Brandon the important qualification that gene additions and deletions can be handled in the manner of forces, when these additions are specified in a manner that makes them independent

19. Our critics persistently ignore the problem of interdependencies. Brandon and Ramsey, for instance, elaborate a number of different ways of estimating total evolutionary change under multiple causes—methods that detect selection in terms of its effects, methods that detect selection in terms of its causes, etc.—but fail to notice that they are all subject to this same difficulty.

20. Other philosophers seem uncritically to accept the Stephens-Brandon approach. We find instances in Godfrey-Smith 2007 and in Depew 2008.

of one another.[21] But evolutionary theory ubiquitously deals with phenotypes. Fisher's Fundamental Theorem is extremely difficult to state and to prove for precisely this reason—it deals with the fitness of traits, not genotypes and so he has to deal with dominance. The same holds true for Price's equation. Both theorems would have been much easier to state if gene-frequencies and their rates of change were all that were at issue.[22]

7. Hierarchical Realization. Take a nexus of factors at work in some evolutionary situation. Suppose that (i) tall people tend to mate with other tall people, and (ii) dark-skinned people with other dark-skinned people. Suppose that (iii) dark-skinned people tend at a higher rate than others to migrate to a certain location, *L*, where (iv) relatively speaking, there is selection in their favor, but (v) tall people are neutral with respect to *L*, where (vi) their fitness is neither enhanced nor reduced. Suppose finally that (vii) tall people tend overall to have more short offspring by reason of mutation and dominance than vice versa. Suppose that each of these tendencies is quantifiable in terms of the effect that each has on the rate at which tall/short/dark/light people occur in future generations, relative to the rate they occur in earlier generations. Suppose that these numbers are known. Can their joint effect be calculated? Is it determinate?

As we said earlier, it is a fundamental principle of probability/frequency theory that the rates of increase associated with each of these conditions are insufficient to determine the joint effect of the suppositions offered above. Consider tall people. The probability of their increasing in frequency is influenced by the rates of increase associated with (i), (v), (vi), and (vii). Correspondingly, the probability of dark-skinned people increasing in numbers is influenced by (ii), (iii), and (iv). Note, however, that the rate of increase in a certain subclass of dark-skinned people is influenced by (i), (v), (vi), and (vii)—namely, that of tall dark-skinned people; correspondingly, the rate of increase of a certain subclass of tall people is also influenced by (ii), (iii), and (iv). Moreover, these rates are influenced by dominance—how often tall/dark-skinned people have short/

21. Gene frequencies can be specified independently even when they are interdependent. If allelic additions and deletions are specified in brute numerical terms—and this, pretty much, is the trick that Stephens and Brandon rely on—rather than as functions of various other parameters, then they can be added. These brute numerical additions and deletions are not, of course, how population genetics or field evolutionists specify conditions. They don't go around counting alleles. They are interested in interconnections.

22. In Matthen and Ariew 2002, we discussed Li's Theorem, which is essentially a restriction of Fisher's Theorem to allele addition alone and on the assumption that growth rates are constant, hence independent. This theorem is very easy to state and understand, because dominance relations, etc., do not enter the picture.

light-skinned offspring without mutation—and the interaction of the various genes that make up each complex. Inter-trait connections of this sort are essential to the outcome, but their magnitude and influence are neither implied nor counterimplied by conditions such as (i)–(vii). In probability/frequency theory, the probability of conjunctions is not determined by the probability of the conjuncts alone: the interdependence relations among conjuncts, the relevant conditional probabilities or conjunctive frequencies, have to be specified in addition. The same holds here.

It is by contrast a fundamental principle of Newtonian force-theory that forces are vectors and combine by vector addition—the resultant force of two forces is wholly determined by the magnitude and direction of each force. Thus, forces have one kind of calculus of combination, and probabilities another. It is our contention (Matthen and Ariew 2002) that evolutionary influences such as migration, mutation, and selection combine in the manner of propositional conjuncts in probability/frequency theory, and not in the manner of forces combining by vector addition. These algebras are quite different from one another. Hence proposition (B) stated at the start of this paper: the resolution-of-forces model does not provide us with a proper understanding of how natural selection combines with other evolutionary influences, or for that matter how distinct selective influences combine with one another.

We offer Hierarchical Realization as a model of propositional conjunction in probability/frequency theory? Consider the set of all possible population histories, *N*. *N* assumes no facts about types of organisms, no fact of the type exampled by (i)–(vii) above: each such possibility concerning types is true in a subset of population histories, but since *N* includes all of the histories that conform to these conditions as well as all of the histories that violate them, *N* is neutral with respect to them. Call it, therefore, the 'neutral model'. Because the Laws of Natural Selection (Fisher's theorem, etc.) are mathematically true, we need no more than the neutral model to understand them.

Now, each condition among (i)–(vii) corresponds to a subset of *N*, the set of histories in which these conditions is satisfied. Each member of such a subset is a *realization* of the corresponding history. Further, the realization of the conjunction of conditions is the intersection of the subsets that realize each condition taken separately. The point that we have been emphasizing is that the frequencies that prevail in a realization of component conditions are not the same as those that prevail in the set out of which the realization is carved. Thus, the frequencies that prevail in *N* do not predict the frequencies that prevail in any condition-defined subset of *N*. And frequencies in the intersection of two conditions are not mathematical functions of the frequencies in the sets that intersect.

The frequencies that prevail in the realization set corresponding to any

condition can be ascribed to the effects of that condition. For example, if it turns out that the reproductive success of tall people depends on how many other such people are around, it is plausible to cite condition (i)—the assortative mating condition for tall people as a cause of the frequency dependency of the success of tall people. Now, (ii), the assortative mating condition for dark-skinned people, will combine with (i) in the intersection of the realization sets of (i) and (ii). But note that (i) and (ii) give us no reason to assume any particular pattern of assortative mating for tall dark-skinned people—the strength of their preference for one another might be quite different from that of dark-skinned people generally, or tall people generally. Further, these conditions may combine in a quite different way than other pairs of conditions—(iii) and (iv), for example. The upshot is that there is no shared underpinning for conjunctions. In Newtonian force-theory, the combination of forces is handed off to a shared mathematical operation, the parallelogram of forces. In frequency theory, the combination of conditions has to proceed case-by-case. Combination depends on local conditions, as it were—on the causal influences that operate with the particular traits or conditions that are thus combining.

The theory of natural selection—mathematical population genetics—is not as such concerned with added conditions such as (i)–(vii). Natural selection is an abstract phenomenon that obtains in all possible population histories. Natural selection is not even a biological phenomenon as such. It holds in any history in which the terms of the theory can be jointly interpreted in a way that accords with the abstract requirements of the theory. Suppose that you have two bank accounts, one yielding 5% interest and the other yielding 3%. One can treat units of money in each account as the members of a population, and the interest rate as an analogue of fitness. Provided that no money is transferred from one account to another, one can treat these 'fitness' values as heritable—that is, the fitness of any particular (nonoriginal) piece or unit of money can be ascribed to the 'reproductive' rate (i.e., interest) on preexisting units of money. Thus you would have, as between the monies resident in the two accounts, variation in heritable fitness. On this interpretation, Fisher's Fundamental Theorem of Natural Selection applies to your bank accounts: it predicts (correctly) that the average interest earned by the two bank accounts taken together will increase in proportion to the variance of interest rates earned by your money in the two accounts.

Natural selection is wholly abstract, then, but its realizations are shaped by concrete relations—these concrete relations are what determine the value of the abstract parameters of natural selection. This is why, as we have argued, variation with respect to a particular trait (such as camouflage) can be said to cause evolutionary change, but natural selection

cannot. This abstractness of natural selection is what our critics miss. Brandon, Millstein, and Stephens are interested in selection as a causal process. But they look for the causes in the wrong place—in the empirically vacuous conditions that determine the set of all possible population histories, rather than in the causally potent conditions that correspond to specific biological traits and conditions.

Mathematical population genetics is, in large measure, an application of probability/frequency theory. It is only with the concrete realization of abstract quantities such as growth rates that biology enters the picture. Once the hierarchical nature of realization is properly appreciated, it will also be seen that there is no one causal process that different realizations share. Bank accounts are not subject to the same causal processes as populations of moths. Yet both instantiate Fisher's Fundamental Theorem. This is why it is a mistake to think that Fisher's Theorem (or any other theorem of population genetics) describes a unitary causal process.

REFERENCES

Bouchard, Frédéric, and Alex Rosenberg (2004), "Fitness, Probability, and the Principles of Natural Selection", *British Journal of the Philosophy of Science* 55: 693–712.

Brandon, Robert (1978), "Adaptation and Evolutionary Theory", *Studies in History and Philosophy of Science* 9: 181–206.

——— (2006), "The Principle of Drift: Biology's First Law", *Journal of Philosophy* 102: 319–335.

Brandon, Robert, and Grant Ramsey (2007), "What's Wrong with the Emergentist Statistical Interpretation of Natural Selection and Random Drift?", in David Hull and Michael Ruse (eds.), *Cambridge Companion to the Philosophy of Biology.* Cambridge: Cambridge University Press, 66–84.

Depew, David (2008), review of *The Cambridge Companion to the Philosophy of Biology* edited by David Hull and Michael Ruse, *Notre Dame Philosophical Reviews,* http://ndpr.nd.edu/review.cfm?id=13388.

Fisher, R. A. ([1930] 1999), *The Genetical Theory of Natural Selection: A Complete Variorum Edition.* Reprint. Edited by Henry Bennett. Oxford: Oxford University Press.

Godfrey-Smith, Peter (2007), "Conditions for Evolution by Natural Selection", *Journal of Philosophy* 104: 489–516.

Hall, Ned (2004), "Two Conceptions of Causation", in J. Collins, N. Hall, and L. A. Paul (eds.), *Causation and Counterfactuals*. Cambridge, MA: MIT Press, 225–276.

Hausman, Daniel M., and James Woodward (1999), "Independence, Invariance, and the Causal Markov Condition", *British Journal for the Philosophy of Science* 50: 521–583.

Lewontin, Richard (1974), *The Genetic Basis of Evolutionary Change*. New York: Columbia University Press.

Matthen, Mohan, and André Ariew (2002), "Two Ways of Thinking about Fitness and Natural Selection", *Journal of Philosophy* 99: 55–83.

Millstein, Roberta L. (2002), "Are Random Drift and Natural Selection Conceptually Distinct?", *Biology and Philosophy* 17: 33–53.

——— (2006), "Natural Selection as a Population-Level Causal Process", *British Journal for the Philosophy of Science* 57: 627–653.

Reisman, K., and P. Forber (2005), "Manipulation and the Causes of Evolution", *Philosophy of Science* 72: 1113–1123.

Rosenberg, Alex, and Frédéric Bouchard (2005), "Matthen and Ariew's Obituary for Fitness:

Reports of Its Death Have Been Greatly Exaggerated", *Biology and Philosophy* 20: 343–353.

Shapiro, L., and E. Sober (2007), "Epiphenomenalism: The Dos and the Don'ts", in G. Wolters and P. Machamer (eds.), *Thinking about Causes: From Greek Philosophy to Modern Physics*. Pittsburgh: University of Pittsburgh Press, 235–264.

Sober, E. (1984), *The Nature of Selection: Evolutionary Theory in Philosophical Focus*. Cambridge, MA: MIT Press.

Stephens, C. (2004), "Selection, Drift, and the 'Forces' of Evolution", *Philosophy of Science* 71: 550–570.

Walsh, Denis (1998), "The Scope of Selection: Neander and Sober on What Natural Selection Explains", *Australasian Journal of Philosophy* 76: 250–264.

Woodward, James (2003), *Making Things Happen: A Theory of Causal Explanation*. New York: Oxford University Press.

——— (2008), "Causation and Manipulability", in Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, http://plato.stanford.edu/entries/causation-mani/.