

organism ("if such-and-such happens, its chances of dying are thus-and-so"), I may also know what the probability is that predators will approach, that an epidemic will occur, and so on. I might summarize these facts in the judgment that the individual has a very low overall chance of surviving the next year. Now suppose the individual dies shortly after this overall assessment is made. Presumably, these facts together offer *some* explanation of why the individual failed to live through the year. Granted, they do not specifically single out the *actual* cause of death, and therefore may seem to be less than ideally complete.<sup>5</sup> Still, life expectancy describes the overall causal framework of factors that may impinge, and thus may afford *some* understanding of why the individual dies.<sup>6</sup>

So it isn't that explanation and causation are absolutely irrelevant to each other. For now, I will describe their connection a bit vaguely: The explanation of an event describes the "causal structure" in which it is embedded.<sup>7</sup> Explanation, to be illuminating, need not single out the events that actually did the causing. Even when a description does this—as when the cause of *Y* is described simply as "the cause of *Y*"—it may fail to be explanatory. The concepts of causation and explanation have yet to be fully clarified, but I hope that the contrast I have in mind is beginning to come into focus.

In this section, I have argued that an organism's overall fitness and the overall fitness of a trait (which, recall, is the average fitness of the organisms possessing it) do not cause survival and reproduction. Not that this constitutes any victory for those who disparage evolutionary theory or the theory of natural selection. As we will see in the next section, natural selection is causally efficacious, even if overall fitness is not.

5. Although in Section 6.1 we will examine a style of explanation—equilibrium explanation—in which the specification of the actual cause seems to contribute little or nothing to our level of understanding.

6. Also relevant here is the point made in Section 1.5 that fitter organisms need not outsurvive less fit ones; so when they do, the fact that they are fitter offers some illumination of the kind of causal process at work.

7. This claim concerns the explanation of single events, not laws of nature or mathematical truths. It would be desirable to have a theory of explanation that could subsume all these cases in a univocal format. However, I will not propose a view at this level of generality here.

### 3.2 Selection Of and Selection For

The overall fitness of an organism (with respect to those factors influencing viability) is one minus the sum of its chances of dying from numerous possible causes. Two organisms *A* and *B* may therefore have the same overall chances of dying, even though they have different vulnerabilities to specific possible causal factors. For example, this may be true because *A* is more vulnerable to disease than *B*, while *B* is more vulnerable to predators than *A*.

Let's imagine a large population of organisms each of which is a clone either of *A* or of *B*. The organisms in this population would be identical in overall fitness. Does it follow that no selection process can occur here, since the requisite variation in fitness is absent? In a sense that I will try to clarify, the answer to this question is *no*. When we look at the population, it wouldn't be surprising to learn that the percentage of *A*'s that die of disease exceeds the percentage of *B*'s that die from this cause. Nor would it be a shock to learn that the percentage of *B*'s that fall to predators exceeds the percentage of *A*'s that get eaten. Disease selects against type *A* and favors type *B*, while predation selects against *B* and favors *A*. It's true that there is no *overall* selective difference between the two types, because there is no overall difference in fitness. Nevertheless, there's a good deal of selecting going on.

We can think of the two sources of mortality at work in this population as component forces. One favors *A* at the expense of *B*; the other favors *B* at the expense of *A*. The net force is zero; so there is no net change in the frequencies of the two traits. The situation here is like the billiard ball case in which you push the ball north and I push it south with the same vigor. There are component forces at work, but no net force, and so no change.

The concept of overall fitness offers a rather coarse-grained description of the various selection processes at work in a population. It summarizes the various selection process in the way that the vector addition used in billiard ball physics combines the component forces into a single net force. The summary constructed involves a loss of information, in this sense: from the component forces, you can calculate what the net force is,<sup>8</sup> but you can't recover the components from the net. Given that the billiard ball has a zero net force acting on it, you can't tell whether there were lots of component forces that canceled out each other (and what these various influences actually were) or whether no forces impinged on the object at all. Given that *A* and *B* are equal in

8. Assuming, of course, that the theory of forces provides a rule for calculating the effects of forces when they act together. See the discussion in Section 1.2 concerning the nontrivial character of this compositional problem.

overall fitness, you can't tell whether there were various component selective forces at work in the population that canceled out each other, or if the population was simply not touched by selective forces of any kind.

These different possibilities will make no difference if one is interested only in calculating the *consequences* that the present state of a system has for its future. Whether the billiard ball is acted on by no forces at all or by a suite of forces that are mutually nullifying, the same thing happens—nothing. The same holds for the population that has no variance in overall fitness. It was pointed out in Section 1.4 that population genetical models are usually interested only in the consequences for gene frequencies of various configurations of forces. It is no wonder that such models make use of the concept of overall fitness and do not need a more fine-grained description. However, there is more to selection than its effects. One also wants to consider its sources. One frequently wants to know *why* the net selective force has the value it does, and this requires decomposing the net into its components. In the sickle-cell example (Section 1.4), we were interested not only in the overall fitness of each of the three genotypes but also in *why* the genotypic fitness values have the ordering they do. The mathematical model was written in the idiom of overall fitnesses, but the explanatory story that accompanied it provided further information.

There is another way in which the concept of overall fitness provides an impoverished representation of the causal structure of a selection process. The fitness of a trait, recall, is the average fitness of the organisms having that trait. This implies that two traits attaching to precisely the same organisms must have the same fitness value. But to treat such traits as equivalent in fitness is to ignore the fact that one of them may have great selective importance, whereas the other may simply be "neutral." In Section 1.2, I mentioned the example of the human jaw and chin. Imagine that individuals with a certain jaw structure always had chins (and conversely). Hence the overall fitness of the two traits must be the same. Yet there was selection for jaws, but not for chins. To express this idea, we clearly need a concept that allows us to say more about a trait than simply describe its overall fitness.

Thus the concept of overall fitness is limited in two ways. It describes the net force of selection, not its components. And it cannot identify differences in the selective significance of traits that apply to the same individuals. These inadequacies are consequences of the deeper fact that fitness is causally inert. But in its division of explanatory labor, evolutionary theory provides a causal concept to do the work that fitness cannot perform. This is the idea of *selecting for and against properties*. In the example considered above of the population composed

of two types of individuals *A* and *B*, the traits were identical in overall fitness, even though there was *selection for* the ability to evade predators and *selection for* immunity from disease. In the second example, although jaws and chins are indistinguishable in terms of their overall fitness, the present concept identifies a difference: there was *selection for* having a jaw of a certain kind, but none for possessing a chin.

We must explore the logic of this new concept—the idea of selection for and against characteristics. It is pleasant that the commonsense notion of "selecting" has the structure and marks the distinctions we need to identify. Figure 2 shows a toy that my niece once enjoyed playing with before it was confiscated to serve the higher purposes of philosophy. Each horizontal level contains holes of the same size. The holes on each level are larger than those on the level below. The balls also vary in size. If the balls are at the top, shaking the toy distributes them to their respective levels. This is a selection machine. Balls are

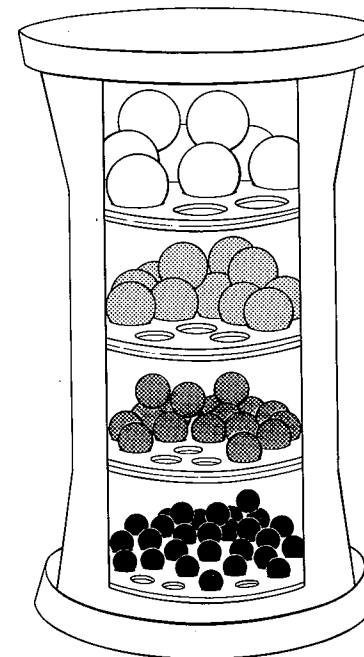


Figure 2. A selection toy in which the name of the game is getting to the bottom. The green balls are the smallest and therefore have the best chance of descending. After a thorough shaking, the small balls—i.e., the green ones—are selected. There is selection for being small, not for being green. This illustrates the difference between the concepts of *selection of objects* and *selection for properties*.

selected for their smallness. The smaller a ball is, the more successful it is at descending. Balls of the same size happen to have the same color. The smallest balls, and only they, are green. So the selection process selects the green balls, because they are the smallest.

There are *two* concepts of selection that we must pry apart. There is *selection of objects* and there is *selection for properties*. The smallest balls are the objects that are selected; it is equally true that the green balls are the objects that are selected. However, the concept of selecting for properties is less liberal. There is selection for smallness, but there is no selection for being green.

"Selection of" pertains to the *effects* of a selection process, whereas "selection for" describes its *causes*. To say that there is selection for a given property means that having that property *causes* success in survival and reproduction. But to say that a given sort of object was selected is merely to say that the result of the selection process was to increase the representation of that kind of object.

When the green balls reach the bottom more frequently than the blue ones, we think that there must have been a reason why. Green balls were selected; so they must have had some property that was selected for. But the property in question was not their color. There was *selection of* green objects, but no *selection for* greenness. I offer the following slogan to summarize this logical point: "*selection of*" does not imply "*selection for*."<sup>9</sup> This idea will be important when the question of the "selfish gene" is considered in Part II.

"Selection for" is the causal concept *par excellence*. Selection for properties causes differences in survival and reproductive success, even though (as argued in the previous section) overall fitness is causally inert. An organism's overall fitness does not cause it to live or die, but the fact that there is selection against vulnerability to predators may do so. Overall fitness gives a summary picture of an organism's vulnerability to *possible* selection forces. There being selection for a particular property, on the other hand, means that a certain causal process is *actually* in motion.

So far, I have discussed the fitness of objects, the fitness of properties, selection of objects, and selection for properties. The *of* concepts all aim at describing the effects of selection for properties but do so in slightly different ways. There can't be selection of objects without there being some sort of selection for properties.<sup>10</sup> But one can have selection for

9. Larry Wright (1973, pp. 163–164) draws the same distinction. I am indebted to Steve Kimbrough for discussion on this point.

10. This doesn't mean, of course, that there can't be selection of green objects unless there is selection for greenness. It needn't be *that* particular property that drives the causal engine.

properties without the total selection process favoring one kind of object over another. The component forces may cancel out each other and generate no net selection effect at all. In the previous example, there was selection for predator avoidance and for disease resistance, but no overall selection of *A* individuals as opposed to *B*.<sup>11</sup>

The distinctions drawn here are not limited in their application to children's toys. The idea that there can be "free riders" on selection processes depends on the same bifurcation. Selection can result in a characteristic increasing in frequency without there being selection for it. As noted in Section 1.2, *pleiotropy* is one way that this can happen. If two phenotypic characteristics are both caused by a common underlying gene (or gene complex), then selection for one may increase the representation of both. Recall the example of the human chin; it emerged via a selection process but, apparently, without there being selection for it. A second way in which there can be free riders on a selection process is by way of *gene linkage*. A neutral gene and an advantageous gene may be close together on the same chromosome. Selection for one may increase the frequency of both. Whether the free rider is phenotypic (as in the pleiotropy case) or genetic (as in the case of gene linkage), we say that there was no selection *for* the character in question, meaning that its increase in frequency was not caused by its conferring an advantage.<sup>12</sup>

When a free rider characteristic increases in fitness owing to natural selection, its overall fitness (i.e., the average fitness of the organisms possessing it) must have been greater than the average fitness of the alternative characteristic(s). And, looking back on the way the process proceeded, we may correctly note that the free rider characteristic was selected (*cf.* the green balls in the toy). The concept of overall fitness and the idea that there was *selection of objects* fails to capture the causal structure that the concept of *selection for properties* is designed to characterize.

So natural selection—that is, selection for characteristics—is one of the causes of evolution. When selection for and against various properties of organisms produces evolution, it must be true that the organisms differ in overall fitness. However, there can be selection for and against

11. And even when there is a net selection effect, other evolutionary forces, like mutation and migration, may cancel *that* out, with the result that no evolution at all takes place. The *ceteris paribus* clauses are nested inside each other like so many Russian dolls.

12. Note that it may nevertheless be true that there is selection for linkage to advantageous characteristics. A gene's possessing this property may cause it to increase in frequency. This is quite consistent with saying that the gene is itself selectively neutral (i.e., fails to confer reproductive advantages on the organisms in which it occurs). Selection of the former sort will be discussed in Section 9.1.

properties without this being reflected in differences in overall fitness. Conversely, the mere fact that there is variation in overall fitness does not yet establish why it exists. Fitness differences among organisms or traits do not by themselves reveal which properties are selected for and which are selected against.

The contrast drawn in this chapter between fitness and natural selection does not depend on the claim, developed in Section 3.1, that fitness is causally inert. Even if an organism's fitness were taken to cause its survival and reproductive success, a conceptual difference between fitness and natural selection would still remain. To say that one organism is fitter than another is not yet to say *why* this is so. To say that there is selection for one characteristic and against another, however, is to give a more detailed description of the causal facts. We could abandon the claim that fitness is causally inert and still maintain the thesis that the fitness concept is causally *oblique*. Without wading into hard questions about whether (or in what circumstances) disjunctive properties are causally efficacious, we can still acknowledge that the proposition " $C_1$  caused  $E$ " gives us a more concrete picture of a causal process than does the proposition " $C_1$  or  $C_2$  or . . . or  $C_n$  caused  $E$ ." Not that this difference shows any *defect* in the fitness concept. As we saw in Section 1.5, that concept's lack of specificity allows it to carve out general patterns of explanation that would be invisible to more concrete concepts like "selection for predator avoidance" and "selection for disease tolerance." Evolutionary theory's conceptual economy deploys different concepts with different functions. Concepts complement each other; they do not put each other out of business.

## Chapter 4

### Chance

---

What is the role of the concept of chance in evolutionary theory? This question often elicits one of the two following responses:

1. Evolutionary theory says that the variation on which selection acts is due to mutations occurring in the genetic material. And it is a central proposition of the theory that such variation occurs "at random." Hence, mutation is a principal way in which the concept of chance enters evolutionary theory.
2. Natural selection is itself a process with an important chance component, and this constitutes a significant difference between the theory of natural selection and, say, the theory of gravitation. Whereas classical physics provided a deterministic theory of the phenomena it subsumes,<sup>1</sup> the theory of natural selection was an early example of the power of statistical thinking in the natural sciences.

Both 1 and 2 contain elements of truth, but both require clarification and supplementation.

In Section 4.1, I analyze the standard remark that mutation is "random" and discuss the difference between Lamarckian and Weismannian theories about the origin of variation. In Section 4.2, I explain how the terms "deterministic" and "stochastic" are used in evolutionary theory. In one very central sense, mutation and natural selection are conceived of as *deterministic* processes: They do not themselves incorporate an element of chance but stand in contradistinction to the stochastic element in evolutionary processes, namely, *random genetic drift*.

In Section 4.3, I address a question that has more traditional philosophical credentials. In evolutionary theory, probabilities are assigned to a variety of events. What is it that makes these probability assignments true? Until the twentieth century, one widely accepted answer was that

1. For a detailed examination of what is defensible and indefensible in this standard remark, see Earman (1984).