How can an explanation say why one event rather than another took place, if it would equally well explain whichever alternative came true?

The idea that likelihood is a measure of explanatory power helps provide an answer. An account explains why one event rather than another occurred only if it says that the former was more probable.[8] This leaves it open that even events said to be improbable can have their contrasting alternatives. The fact that the population began with 2N-1 copies of the A allele and only a single copy of a and then experienced random drift explains the novel mutant's subsequent universality. It does *not* explain why a went to fixation rather than A. But it does explain why it was a rather than some third (hypothetical) allele that spread. Thank heavens that improbable events have still more improbable events with which to contrast!

Our concept of explanation does not presuppose that determinism is true. Events may be explained by fitting them into causal structures that are probabilistic in character; even improbable events may be explained. But consistent with all this is the fact that our notion of explanation is oriented toward a deterministic ideal. Although propositions may explain an event even when they say that it was highly improbable, there is nevertheless an element of truth in one-half of the symmetry thesis. An explanation need not be a prediction, but one property that makes an explanation good is that it facilitates prediction. That is, likelihood is one dimension along which explanatory power increases.

A common thread unites the present point about explanatory power and the idea developed at the beginning of this section concerning why fitness differences can be difficult to establish. Philosophers of science have often inferred that a science's inability to explain or predict a particular event must reflect some logical defect in the theories deployed. For example, if evolutionary biologists find it difficult to assign fitness values to characteristics, this must imply some imperfection in the theories they use. This same idea, turned on its head, asserts that if a theory that is beyond reproach issues in explanations of a certain sort, then such explanations cannot fall short of any explanatory ideal we might reasonably maintain. So we find it asserted that there is nothing wrong with explaining an event by showing that it was very improbable, since quantum mechanics produces explanations of just this sort. There is only one thing wrong with this point of view: it ignores the way *the world* can guarantee that explanations fail to live up to reasonable standards. Besides blaming the victim (the scientists stuck with the fact

8. Van Fraassen (1980, p. 128) notes that this idea was suggested in an unpublished manuscript by Bengt Hannson.

that the world has forced them to accept certain theories) and readjusting the standards of evaluation to mask the fact that there is a shortcoming, another response is possible. We may acknowledge that the world plays a crucial role in determining the extent to which the theories we accept will permit us to construct accounts that measure up to our standards of good explanation.

5.2 Variational and Developmental Explanation[9]

The idea, developed in Section 5.1, that an event is explained by showing why it rather than some contrasting alternative occurred, throws an interesting light on the concept of natural selection. Darwin not only offered a new explanation of old phenomena—namely, the observed adaptedness of organisms to their environments. Additionally, he redescribed the world so that the very propositions that called for explanation were different from the ones that earlier theories had focused upon.

It is standard for a scientific innovation or revolution to change our picture of what phenomena require explanation. Typically, this transformation occurs because new empirical concepts are introduced. For example, evolutionary theory now seeks explanations for the amount of heterozygosity in a population, whereas Darwin did not. The reason is simply that we now possess a descriptive vocabulary that Darwin lacked. The explanatory innovation I want to explore in this section is different. In a sense I will try to make precise, it involves a change in *logic*, not a change in *concepts*.

As noted in Section 1.1, Darwin's was certainly not the first evolutionary theory. The revolution was not the idea of change but the mechanism that produces it.[10] Seventeenth- and eighteenth-century evolutionary theories have been described by Lovejoy (1936) as "temporalizing" the Great Chain of Being. The hierarchy from "lower" organisms to "higher" ones, with human beings at the pinnacle, was thought to correspond to a chronological pattern. The evolutionary theory associated with the name of Lamarck (1809) was in this mold (Mayr 1976a). For Lamarck, there is a predetermined path that evolution tends to pursue. Each lineage begins with very simple life forms and

9. The contrast between variational and developmental theories elaborated in this section is due to Lewontin (1983).
10. This is not to deny that part of Darwin's achievement was showing how the hypothesis of evolution could explain certain phenomena better than the doctrine of special creation. It is important to separate Darwin's defense of the hypothesis of evolution from his defense of the idea that natural selection has been the principal cause of evolution. This is done with admirable lucidity in Kitcher (1984a).

gradually gains in complexity. The evolutionary progression continually begins anew, with spontaneous generation constantly replenishing the store of primitive organisms. Within a lineage, lower rungs on the evolutionary ladder are vacated and higher rungs are eventually occupied.

This postulated progressive tendency predicts that if the whole history of life consisted in the development of a single lineage, we would not observe any diversity at all; there would be just one species. But since different lineages begin at different times, diversity is possible; different lineages, having begun at different times, have ascended to different heights. *Our* lineage must be the oldest, since human beings are the "highest" living forms.[11]

Lamarck's theory postulates a second, and secondary, evolutionary force. Besides being guided by a progressive tendency, a lineage may be influenced by "forces of circumstance." Local conditions may produce differences between populations that are at the same evolutionary level of development. This means that Lamarck did not have to arrange the entire diversity of life in a single linear ordering, as if each pair of species had to be related as more or less advanced. But since the central upward tendency was given the greater emphasis in his thinking, Lamarck was committed to thinking of the main features of morphology, physiology, and behavior as unfolding in a one-dimensional sequence.

It is significant that in his notebooks Darwin reminded himself "to never say higher and lower" (Gruber 1974, p. 74). I will not presume to say what this piece of advice meant to Darwin at the time, but its relevance to his mature theory is worth pondering. The point is not that there can be no differences in adaptedness between organisms or between species. As Ospovat (1981) has argued, Darwin had a genuine interest in and commitment to finding a measure of perfection. Rather, the idea that this remark makes salient is that the notion of a hierarchy of stages through which evolution proceeds plays no role whatever in the theory of natural selection.

Lamarck's theory was *developmental*. It explained the evolution of species by laying down a sequence of stages through which life forms are constrained to pass. Species evolve because the organisms in them are gradually modified. In contrast, Darwin's theory of the evolution

---

11. Pietro Corsi pointed out to me that Lamarck drew back from this explanation and eventually reached a position that was purely descriptive, in that no mechanism was suggested for generating the patterns he posited. As is usual, what comes to be known as "X's position" is, *at best*, X's position at a certain time. In the same vein, Mayr (1976a, 1982) argues that the rather standard interpretation of Lamarck as thinking that evolution is propelled by a kind of psychological striving on the part of organisms is the result of a mistranslation of "besoin."

of species is not developmental. Darwin explained change in a species by a mechanism that permits (and, in a sense to be made clear, even *requires*) *stasis* in organisms. In addition, the Darwinian paradigm views evolution as opportunistic, not preprogrammed. Selectional theories and developmental theories have fundamentally different explanatory structures.

A simple example helps bring out the contrast. You observe that all the children in a room read at the third grade level. What could be the explanation? Two strategies of analysis are possible. A developmental account will take the children one at a time and describe the earlier experiences and psychological conditions that caused each to attain that particular level of reading proficiency. These individual stories may then be *aggregated*. You may explain why all the children in the room read at the third grade level by showing why Sam, Aaron, Marisa, and Alexander each do.

A selectional explanation would proceed very differently. Suppose it were true that individuals would not be admitted to the room unless they could read at the third grade level. This would explain why all the individuals in the room read at that level. But, unlike the developmental story, the selectional account would not explain the population-level fact by aggregating individual explanations. The selectional theory explains why all the people in the room read at the third grade level, but not by showing why Sam, Aaron, Marisa, and Alexander do.

Notice that in the selectional explanation the individual children *do not change*. We may imagine that candidates for admission to the room are evaluated in an antechamber. Some read at the third grade level while others do not. A selection is then made for the ability to read at the third grade level. This was the criterion for admission, *and it is assumed that the individuals do not lose their ability to read at the third grade level as they enter*. In this account, all individuals are static, yet a selection among static individuals can produce change in the composition of a population. The change in the population is not due to the fact that the individuals in it *develop*; rather, what is crucial is that they *vary*. This is the essence of variational explanation.

Both the developmental and selectional accounts explain why all the children in the room read at the third grade level. But they do so by placing that proposition into different contrasting contexts. The developmental story says why each individual has one reading level rather than another. The selectional story, on the other hand, shows why the room is filled with individuals reading at the third grade level rather than with different people with different reading abilities. It would be misleading to describe the two explanations as both discriminating between the following alternatives: ⟨All people in the room

read at the third grade level, Not all the people in the room read at the third grade level). This representation glosses over the fact that the developmental explanation construes "all the people in the room" as encompassing the same set of individuals in both contrasting propositions. The selection account, however, interprets that expression as picking out different individuals in the two contrasting alternatives.

Part of Darwin's revolution was to embed the problem of "explaining organic diversity" in a new contrastive context. Lamarck and Darwin both could have interested themselves in explaining why giraffes have long necks. Lamarck's developmental theory would have interpreted that problem as requiring that one show why existing giraffes have long necks rather than short ones. A progressive tendency would be invoked to show how individual ancestral giraffes were modified, thereby producing a change in the giraffe population. Darwin's selectional account proceeded differently. Rather than aggregating individual developmental stories into an explanation of the population-level fact, Darwin took his question to have an irreducibly population-level character. Population change isn't a consequence of individual change but of individual stasis plus individual selection. The explanatory question became one of saying why the giraffe population is composed of long-necked individuals rather than of other individuals who are not long-necked. The theory of natural selection created a new object of explanation by placing the population fact in a new contrastive context.

In the selectional explanation of why the children read at the third grade level, the individuals do not change their reading abilities as they cross the threshold into the room. Selectional explanations in a more explicitly biological context do not require this sort of absolute stasis, but they tolerate it quite comfortably. To see why, we need to examine how mortality and fertility selection require certain "stability assumptions," if selection is to produce change.

What does it take for mortality selection to produce a change in trait frequencies in a population? Suppose the transition from zygote to adult in a population has individuals with characteristic F dying twice as frequently as individuals with the alternative characteristic not-F. Does this mean that the frequency of F individuals will decline? For this to follow, we must make the usual *ceteris paribus* assumption: Migration exerts no decisive influence, nor does mutation, and so on. However, in addition to these, there is another assumption we need to make. Roughly, we need to assume that an F individual at one time is an F individual later on (and similarly for those who are not-F). Individuals need not be absolutely stable in their characteristics for selection to produce change, of course. But stability there must be;

otherwise, selection may produce no net change, even in the absence of mutation, migration, drift, and so on.

The same sort of stability assumption is more obviously required in the case of fertility selection. If F individuals have twice as many offspring as non-F individuals (and no other evolutionary forces impinge), this will produce an increase in the frequency of F's in the next generation only if the trait in question is *heritable*. This doesn't mean that the trait has to be "genetically encoded," in the sense described in Section 4.1 having to do with the norm of reaction. Heritability simply ensures the right sort of *resemblance* between parents and offspring.[12] Will a greater birthrate among cowboys in one generation increase the frequency of cowboys in the next? Only if there is a resemblance between the occupations of parent and offspring. This "matching" may be due to the cultural fact that parents teach their children. Heritability does not require a "gene for cowboyhood."

I have already noted (Section 1.1) that selection doesn't imply evolution. The present point goes beyond this: Selection does not imply evolution even when selection is the only evolutionary force at work. Selection implies evolution only when no evolutionary force counteracts it *and the trait being selected is heritable*.

Issues of selection and heritability are often confused. To check on the selective advantage of a trait in an evolving population is one thing; to see whether it is heritable is another. For example, if you wanted to know whether heterozygotes have a higher fertility than homozygotes, you might count up the number of offspring that heterozygotes produce per capita and compare this with the number of offspring that each of the two homozygote forms produces. In doing this calculation, it would be irrelevant what the genotypes of the offspring were. A parent who is heterozygotic may produce offspring of all three genotypes, but *all* its offspring, regardless of their genotypes, are counted as part of the parent's reproductive output.

Another way of estimating genotypic fitness values—a fallacious one—is to compare the frequency of heteroyzgotes in the parental generation with the frequency of heterozygotes among their offspring and to reason that if the trait declines it must be at a selective disad-

---

12. A quantitative trait like height is heritable when parents who are above or below the mean tend to produce offspring who deviate in the same direction. Imagine a "scatter diagram" of midparent heights (i.e., the average of the parents' heights) on the x axis and offspring heights on the y axis. If parents and offspring are related in the way just described, a regression line drawn through these data points will have a positive slope; indeed, the heritability (h^2) just is the slope of this regression line. For an explanation of what heritability is and how it influences the response to selection for a quantitative character, see Roughgarden (1978).

vantage. A simple example of what is wrong with this procedure is furnished by the balanced lethal system described in Section 1.4. Suppose that all homozygotes die in the passage from egg to adult stage and that all heterozygotes survive and then reproduce. The frequency of heterozygotes at the adult stage is 1.0. The frequency of heterozygotes among the next generation of offspring at the egg stage is 0.5. Does this show selection against heterozygotes? Absolutely not.

To see whether a characteristic is being selected against, we discover whether organisms with the trait on average produce fewer offspring. But to see whether the trait is heritable, we must see whether there is resemblance between the characteristics of parents and the *characteristics* of offspring. Here the traits of the offspring must obviously be taken into account.

The examples of selectional explanation (whether they involve survival and reproduction or admission into a room) have a selectional component and a stability component. What exactly does the selectional component explain? In the schoolroom example, selection does not explain why any child can read at the third grade level. The fact that an individual now in the room can do this is accounted for by the stability assumption: That individual could read at the third grade level at some earlier time, and the trait persisted. Selection is not what does the explanatory work here.

The same holds for biological examples. Natural selection does not explain why I have an opposable thumb (rather than lack one). This fact falls under the purview of the mechanism of inheritance (Cummins 1975). There are only two sorts of individual-level facts that natural selection may explain. It may account for why particular organisms survive and why they enjoy a particular degree of reproductive success. But phenotypic and genotypic properties of individuals—properties of morphology, physiology, and behavior—fall outside of natural selection's proprietary domain.

Yet, at the population level, these limitations disappear. When conjoined with assumptions about heritability (stability assumptions), natural selection may account for why 50 percent, or all, or none of the individuals in a population have opposable thumbs. The frequency of traits in a population can be explained by natural selection, even though the possession of those traits by individuals in the population cannot. This reflects the fact noted earlier that selectional explanations, unlike developmental ones, do not explain population-level facts by aggregating individual-level ones. Selection may explain why all the individuals in the room read at the third grade level, but not by showing why each individual can do so.

The alternative to developmental explanation I have labeled "vari-

ational," following Lewontin (1983). Natural selection is the obvious prototype here, but other evolutionary forces are conceptualized in the same way. Drift may be thought of as a process of "random selection."[13] Sampling error may transform a population without any of the organisms in it changing at all. If a population begins with the gene A at 99.5 percent and a at 0.5 percent, where these are equal in fitness and no other evolutionary force impinges, it is very probable that A will go to fixation. When drift modifies the composition of the population in this way, it is not because the individual organisms change but because they vary. The grip of the variational paradigm on evolutionary thinking goes deeper than the Darwinian commitment to the historical hypothesis that natural selection is the preeminent force of evolution.

Developmental theories are familiar in the human sciences, where the entities investigated are individual organisms. Piaget's (1959) theory of cognitive development and Chomsky's (1975) theory of language acquisition seek to explain changes in human beings by postulating a mechanism that transforms them. In both cases, the organism is viewed as "internally constrained." As a result, not just any sequence of stages is to be expected.[14]

Developmental stage theories are predicated on the assumption that the sequence of states an organism occupies is not the fortuitous result of the experiences that happen to impinge. Regardless of wide possible variation in the character and order of experience, the organism will change in a certain way. The idea of a developmental pathway is precisely the idea of regularly occurring changes that are insulated from environmental influence.

So the mere fact that an organism goes through a sequence of changes is no powerful argument for the possibility of a stage theory. Radical environmentalism is also consistent with the fact of change. What is important to a stage theory is the idea that the organism's internal state approximates a sufficient parameter in the theory of transformation. The state of the environment need not be entirely irrelevant, of course. But the more irrelevant it is, the more attractive this kind of developmental theory will be. Powerful endogenous constraints make change look more like an unfolding than like a buffeting.

It now will perhaps be clearer why the Darwinian view of natural selection suggests that there can be no developmental theory of phylogeny. It isn't the mere fact that natural selection is accorded *some*

13. As long as the reader bears in mind that this is contradictory when "selection" is understood in its evolutionary rather than everyday sense!
14. Cohen (1978) interprets Marx's theory of history as a developmental hypothesis, in the sense involved here. Internal developmental constraints propel a society through a series of forms of economic organization.

role in the evolution of life. Rather, it is the idea that natural selection is the overwhelmingly most powerful force of evolution that makes the prospects dim for a developmental account of the origin of species. Natural selection stands in opposition to endogenous constraints. The Darwinian view denies that the evolution of species is "prepro-grammed." It admits no inherent tendency toward complexity or any-thing else. In contrast with Lamarckism, there is no preordained ladder of life that living forms are inherently disposed to ascend.

It is no accident that at the end of the nineteenth century, behaviorism in psychology saw Darwinian selection as a guiding metaphor (Boring 1950). Although Darwinism concerned itself mainly with "genetically encoded" characters and behaviorism focused on "acquired" ones, both doctrines viewed the engine of change as external to the object of change. Behaviorism in ontogeny and natural selection in phylogeny leave little room for developmental stage theories.

The idea of endogenous constraints on the changes a species may undergo is hardly unknown in evolutionary theory. My claim is that *natural selection* stands in opposition to this sort of mechanism. The Soviet biologist Vavilov, a victim of Lysenkoism, suggested a "law of homologous series," according to which similar patterns of evolution are to be expected in different lineages (Gould 1983). The orthogenetic theory of a number of anti-Darwinian paleontologists early in this century likewise tried to explain changes in a given species by thinking of the species as propelled through a series of stages from childhood to maturity to senescence. Lineages, it was thought, attain a kind of evolutionary momentum so that deleterious characteristics continue to develop until, like the paradigmatic horns of the Irish elk, they drag the species down to extinction (described in Gould 1977b).[15]

More recently, Gould and Lewontin (1979) have suggested that en-dogenous constraints may profoundly limit the power of natural se-lection. To result in an improved organism, natural selection must be able to tinker[16] with one part of the organism's phenotype without disrupting the rest of its organization. Lewontin (1978) has termed this requirement "quasi-independence." The idea of biological constraints asserts that the components of an organism's phenotype are so organized and intricately coordinated that it is very difficult to manipulate one piece at a time. On this view, species are pretty much locked into their phenotypes. Significant evolutionary change does not result from the

15. Kohn (1980) describes how Darwin worked with a theory of this sort before arriving at the mechanism of natural selection.
16. This happy phrase is from Jacob (1977).

slow accumulation of piecemeal selective modifications of organisms but has a quite different cause.[17]

Thus the idea of endogenous constraints is not alien to evolutionary theory, nor is it alien to the variational paradigm. Those who now argue that populations are incapable of evolving much under the pres-sure of individual selection are not urging a return to Lamarck's ladder of life. Darwinism asserts that populations can deploy a significant amount of variation and that selection among organisms can thereby produce a significant amount of evolution in a population. The idea of endogenous constraints is now put forward in macro-evolutionary theory not to displace the variational paradigm but to transpose it to another level of organization. The idea is that *species* are static entities, just as Darwinism believed individual organisms to be. Whereas Dar-winism thought of a single population evolving on the basis of variation among organisms, the alternative idea is that an *ensemble of species* evolves in virtue of variation among species. Hypotheses of group selection and species selection, our main concern in Part II, are perfectly at home in the variational framework.

So the variational paradigm should be understood as a very general and abstract form of explanation. One version of it is the idea of Dar-winian, individual selection. However, random genetic drift on the one hand and selection among populations on the other are also in the variational mold. The distinctive feature of variational explanation is a kind of antireductionism: Change in a set of objects is not accounted for in terms of changes in those objects. A population may change although it contains entities that are always static. Formulating this new kind of explanatory model involved rewriting the propositions that require explanation. An important element in this transformation involved placing the propositions to be explained in a new contrastive context.

5.3 Population Thinking and Essentialism[18]

In addition to creating a new object of explanation (Section 5.2), the theory of natural selection also transformed the status of the concept of variability. The fact of variability, both within and between species, has been recognized as something that requires explanation for as long as there has been biological theorizing. It is characteristic of that mode of thought that Mayr (1963, 1976b) has called "typological" or "es-

17. An alternative mechanism for generating large-scale diversity (which is hotly debated at present) will be discussed in Section 9.4.
18. Some of the ideas in this section are given a fuller treatment in Sober (1980).