What Darwinism Explains

Peter Godfrey-Smith

Harvard University

For the Darwin Festival, University of Cambridge, July 2009

Did Darwin really do what Kant said was impossible, and serve as a Newton for the biological world? In assessing this question we need to look at both the structure of evolutionary theory and the structure of our explanation-seeking minds. The short answer to the question is yes. Both underestimates and overestimates of the significance of Darwinian explanations derive from psychological habits which may stem from our own evolutionary history.

1. Introduction

The 18th century German philosopher Immanuel Kant said that there would never be a Newton for the biological world. He said:

> [It is] absurd for human beings… to hope that there may yet arise a Newton who could make conceivable even so much as the production of a blade of grass according to natural laws which no intention has ordered. (*Critique of Judgment,* 1790*,* §75)

A lot of people would say that Darwin was a Newton for biology. If that is true, then Kant was proved wrong a mere 69 years after his prediction. But was Darwin in fact a Newton for biology?

First we need to ask: what did Newton do for physics? Or rather, what did he seem to do at the time? He introduced a theory that unified a huge range of phenomena, and made those phenomena explicable in terms of causes acting in the absence of foresight or purpose. Did Darwin achieve a breakthrough of that kind? My short answer is yes. I won't approach this by wrangling over the Newton-Darwin analogy, though. Instead I will offer some specific views on what Darwinism explains.

What is "Darwinism"? And what is "explanation"? We can distinguish two sets of ideas with special explanatory power in Darwinian biology. One is summarized in the "tree of life," the postulation of a network of common ancestry linking all life on earth. The second is the theory of evolution by natural selection within a population. In this paper I will just discuss the second, evolution by natural selection, the process in which a population changes due to the interaction of three factors: variation, heredity, and differences in how well organisms survive and reproduce.

Much of what I'll say will be neutral between various philosophical views about explanation. I will assume that describing causes is often important in explanation, and finding patterns that unify phenomena is also often important. At the end of the paper I'll say something more substantial on this part of the topic, though. Explanation has both objective and subjective sides. Objectively, the aim is to represent causal facts, and to find real patterns. Subjectively, there is a psychological *feel* to explanation, often an "aha" feeling. In the latter part of the paper I will argue against some views in this area by looking at explanation from a psychological perspective – in particular, attending to the *evolved* psychology of explanation, a set of psychological habits which make evolutionary sense.

2. Distribution Explanations and Origin Explanations

I will make use of a distinction between two kinds of explanations: *origin explanations* and *distribution explanations.*[1]

When we give a distribution explanation we assume the existence of a set of variants in a population, and explain why they have the distribution they do, or why their distribution has changed. Some variants may be common, while others are rare. Some may have been lost from the population, having been present before. A distribution explanation explains facts of that kind.

An origin explanation, in contrast, is directed on the fact that a population has come to contain individuals of a particular kind *at all*. It does not matter how many there are. The point of the explanation is just to tell us how there came to be some rather than none. So now we are explaining the original appearance of the variants that are taken for granted when giving a distribution explanation.

[1] For a more detailed discussion see Godfrey-Smith (2009). This terminology is modified from one due to Karen Neander (1995). Neander distinguished "creation" and "persistence" explanations in a debate with Elliott Sober over what natural selection can explain. I broaden Neander's two concepts, as well as rename them. Explaining distribution facts of all kinds is broader than explaining "persistence," and Neander tied her treatment of "creation" explanations to the idea that genes contain "programs" for traits, which is not needed.

When thinking about what Darwinism explains, origin explanations are especially important. They are also relevant to Kant's impossibility claim – Kant's passage is sometimes translated by saying that the goal is to explain the "genesis" of a blade of grass. This might mean explaining a particular blade's development from a seed, or explaining how we came to have grass at all. The second one is a kind of origin explanation. And on a Darwinian view, the characteristics that mark differences *between* species originate in differences *within* species.

Simplifying both Darwin and modern biology, I will think of evolution by natural selection as a process in which variation appears in a population in an undirected or "random" manner, some of the variant individuals are able to survive and reproduce better than others, and their quirks, both good and bad, tend to be passed from parents to offspring in reproduction. The focus will be on the role of natural selection itself – the part of the process where traits are preserved or lost because some individuals do better or worse than others in survival and reproduction.

Almost everyone will agree that natural selection can figure in distribution explanations; it can explain why a characteristic has become common or rare, for example. The more controversial matter is whether and how selection can figure in origin explanations. It might seem that selection itself has no role in origin explanations. Selection has to do with sorting things that already exist. That process cannot possibly bring new things into being. This would not mean that evolutionary biology could not give origin explanations. Rather, it would be show that origin explanations are given by a different part of the package of factors that evolutionary biology recognizes. Origin explanations would be given in terms of what we now call "mutation" (along with the sexual recombination of characteristics). They would be given in terms of the processes that directly give rise to new variations in a population. So maybe we should say that modern Darwinism contains a package, and mutation is the "creative" part of the package. Selection is a distribution-explainer, while mutation is an origin-explainer.

Hugo De Vries, writing in 1909, for an earlier commemoration of Darwin's birth, expressed a view like this.

> It should here be pointed out that *in order to be selected, a change must first have been produced*. This proposition, which now seems self-evident, has, however, been a source of much difference of opinion among Darwin's followers. The opinion that natural selection produces changes in useful directions has prevailed for a long time. In other words, it was assumed that natural selection, by the simple means of singling out, could induce small and useful changes to increase and to reach any desired degree of deviation from the original type. In my opinion this view was never actually held by Darwin. It is in contradiction with the acknowledged aim of all his work,— the explanation of the origin of species by means of natural forces and phenomena only. *Natural selection acts as a sieve*; it does not single

> out the best variations, but it simply destroys the larger number of those which are, from some cause or another, unfit for their present environment. ("Variation," 1909, italics added)

Again, in a slightly earlier passage:

> Darwin discovered the great principle which rules the evolution of organisms. It is the principle of natural selection. It is the sifting out of all organisms of minor worth through the struggle for life. It is only a sieve, and not a force of nature, not a direct cause of improvement....
>
> [W]ith the single steps of evolution it has nothing to do. Only after the step has been taken, the sieve acts, eliminating the unfit. The problem, as to the manner in which the individual steps are brought about, is quite another side of the question. (*Species and Varieties: Their Origin by Mutation,* 1906, pp. 6-7).

De Vries was not ungracious about Darwin's achievement, but he saw it as part of a larger picture which was only just starting to become clear. All Darwin could do was study observable tendencies in the production of new variations, and tentatively postulate a process by which they are produced by shocks to the reproductive system. The term "mutation" was De Vries' own term, and as he saw it, it was only well after Darwin that the origin of what he calls "the individual steps" in evolution was becoming clear, in large part through his own work.

De Vries himself had a "saltationist" view of evolution, one in which substantial steps can be taken in a single generation. Darwin emphasized gradual change. But even with very small steps, De Vries might say, "in order to be selected, a change must first have been produced." That point is unaffected by the size of steps.

Some of the things De Vries says here make good sense. It is true that in order to be selected, a change must first have been produced. And if this was the complete picture, it would deflate Darwin's achievement quite a bit. De Vries' picture is one in which the creative factor in evolution is the thing Darwin knew least about – the causes of new variations. So Darwin himself would not have done what Kant said was impossible, though evolutionary biology a whole might do it.

But the view outlined above is not the whole story. Selection has a role in origin explanations as well. How is this possible?

Selection is not an immediate, or proximate, cause of a new variant. The most important immediate sources of new variations are mutation and recombination. (From here on I will just discuss mutation.) But selection can reshape a population in a way that makes a given variant *more likely to be produced* through the immediate sources of variation than it otherwise would be.

Selection does this by changing the background against which the immediate sources of new variation operate. This affects what those sources of variation can produce. More specifically, selection can make a new characteristic more likely to be produced by making intermediate stages on the road to that new characteristic common rather than rare. This greatly increases the number of ways in which a rare mutation can suffice to produce the characteristic in question.

I will give a very simple example.[2] Suppose we are explaining the evolution of the human eye. Building the genetic basis of the human eye involved bringing together many genes. Consider a collection of genetic material, Y, that has everything needed, as far as genes go, to make a human eye, except for one final mutation. So this background Y is such that *if* new mutation M arises against Y, it will finalize the evolution of the human eye. Initially, Y was rare in the population — it was the product of a single mutational event that produced Y from yet another precursor, X. Selection can make the appearance of the eye more *likely* by making Y more *common*. This increases the number of independent "slots" in which a single key mutational event will give us the eye. If the intermediate Y remains rare in the population, then additional mutations are much less likely to produce the human eye, because the right mutation has to occur in exactly the right place – in an lineage where Y is present.

Here we have looked just at the last step, and have also told the story looking backwards from a characteristic of interest. The process itself runs forwards, without foresight, and involves many of these steps. What is needed is for many intermediates – Y, X before Y, and so on – to be favored by selection and hence made common. This makes the eye *accessible* to mutation, in a way it would not be without the favoring of intermediates.

So despite the initial oddity of the idea, selection can be a "creative" force in evolution, even though it is true in every case that "in order to be selected, a change must first have been produced," as De Vries put it. What De Vries did not see is that the "production" of a change can be made more likely when selection has altered the background against which the immediate causes of variation operate.

Let's return to De Vries' expression of something that he thought was not true and not believed by Darwin:

> [I]t was assumed that natural selection, by the simple means of singling out, could induce small and useful changes to increase and to reach any desired degree of deviation from the original type.

Suppose we do some more "sieving" on the passage:

---

[2] A clear discussion of these issues is found in Dawkins' book *The Blind Watchmaker* (1986).

> ... natural selection, by the simple means of singling out, could induce small... changes... to reach any desired degree of deviation from the original type.

Then we have, roughly speaking, a truth.

3. Further Features of Origin Explanations

Next I will look more closely at how the distinction between distribution and origin explanations relates to natural selection.

(i) Survival and reproduction

Natural selection is often described in terms of differences in "fitness." And fitness is usually described in terms of two things, survival and reproduction. I talked that way above. But what is the relation between these two? People have looked at this in slightly different ways.

On one view, survival is only important because you need to be alive to reproduce. An organism's fitness is the number of offspring it has, or something like this (maybe a "propensity," and maybe a number that gives you some credit for your relatives' offspring as well). Roughly speaking, you have no fitness until you reproduce, replicate, or "multiply," and staying alive after reproduction is irrelevant to fitness.

On an alternative view, survival is "as good as" reproduction, in principle. The point is clearest for asexually reproducing entities, such as bacteria and individual genes. You can make a new copy of yourself, or just stay around. Either way you are represented in the population (Bouchard 2008).

Which view is right? Does it matter? I suggest that the relation between survival and reproduction is different in the context of distribution and origin explanations. In the case of distribution explanations, persistence in any form is enough. You can change the distribution facts in a population by living longer, without reproducing. But in context of origin explanations, reproduction has a special role. For mutation to give rise to whole new kinds of organisms, there must be a "fresh start" to the developmental process, a new turn of the life cycle (Bonner 1974, Dawkins 1982). So we should think about "fitness" a bit differently for the purposes of distribution and origin explanations.

(ii) Absolute and relative

In evolutionary biology, selection is normally seen as a comparative matter. If we have two types, A and B, and A is "favored by selection" over some time period, this does not have to do with how well A is doing in absolute terms, but with how well it is doing in relation to B (Millstein 2006). Much of the accounting in evolutionary biology is done in terms of *frequencies* of a gene or trait in a population.

This is a sort of "compression of information." We might have many populations, of different sizes. But if in all of them, the relative number of A's is increasing in the same way, we have the same phenomenon of selection. How well A is doing in absolute terms is not relevant. For some special purposes, the absolute numbers are taken to matter. The theory of random change by "drift" attends to absolute numbers, and questions about extinction certainly do. But for many purposes, all that is taken to matter when thinking about selection is relative frequencies.

That is fine in the context of distribution explanations. But things are different in origin explanations. Above I told a story about the evolution of the eye. I said the probability of the eye arising was raised when selection favors a precursor, Y, that is one mutational step away. But what mattered there was changing the *absolute* number of backgrounds that are of the right type. The probability of the eye arising in a population in one step depends on the absolute number of appropriate "slots" there are in which a final mutation will do the job, not on how numerous these slots are in relation to inappropriate ones. The same applies to the precursor of the precursor, and so on.

Selection is only important in origin explanations when it affects absolute numbers in the right way. Usually it does, but it does not have to. Suppose we have a population with some eye-precursors and some non-precursors. Selection can make the eye more likely by making the precursors more common and non-precursors less common, but the same effect can be achieved by increasing the numbers of *all* types. A global increase, in which frequencies do not change, is just as good as selection for the precursor. Similarly, there are ways in which selection favoring the precursor could *decrease* the chances of the eye arising. We might have a population with some precursors and some non-precursors, and selection is introduced favoring the precursors – but it reduces the absolute numbers of all types. The precursor becomes more common relative to its rivals, but less common absolutely. So introducing selection favoring the precursor can, in principle, reduce the probability of evolving the eye.

When we care about origin explanations, we cannot choose to ignore absolute numbers, as we can in the case of many distribution explanations. This point leads to my third one.

(iii) The struggle for life.

In Darwin's own descriptions of how natural selection works, the "struggle for life" was central. Organisms live in a world of scarce resources, and compete incessantly for them. In more recent writing about natural selection, the idea of a struggle for life is often seen as inessential. It is true that resources are usually scarce, but the basic idea of change through differences in reproductive success does not require any sort of struggle. Here is Lewontin (1970, p. 1).

> [C]ompetition between organisms for a resource in short supply is not integral to the argument. Natural selection occurs even when two bacterial strains are growing logarithmically in an excess of nutrient broth if they have different division times.

The view is also expressed by Lewens, in his book *Darwin* (2007): "Modern evolution has no essential commitment to the Malthusian view that lies at the heart of Darwin's theory."

Again, I suggest the issue looks different across the contexts of distribution and origin explanations.

Suppose I have two offspring and you have one; I have more offspring than you. But it may or may not be the case that my having two rather than one prevented you from having two rather than one. There may or may not be a causal dependence between my reproductive output and yours, so that a "slot" I fill in the next generation is a slot that you do not fill. When this sort of causal relation is present, then if I am successful under selection, this implies not just higher numbers *than you*, but higher numbers *than I would have had* if selection had not been favoring me. Selective competition is only important to origin explanations insofar as the "winner" is able to produce more absolute numbers of offspring than it would otherwise. This we saw in my second point above. What is it that produces this sort of causal dependence between reproductive success of different individuals? Roughly speaking, it is scarce resources, what Darwin called a struggle for life.[3]

So what "natural selection" involves in the context of origin explanations is not the same as what it involves in the context of distribution explanations. People often describe selection in a way that is most appropriate for distribution explanations. Perhaps this helps explain why it has often not been clear to people that selection has the role in origin explanations that it really has.

---

[3] Not all the special cases in (ii) are handled with this idea. The case where selection for an intermediate harms the chances of the eye arising through global population decrease is not handled.

4. The Psychology of Explanation and Darwinian Paranoia

Darwin showed how to explain the origination of diverse and complex organisms from a simpler and less diverse stock. He did not explain the origin of life itself, and the explanations he gave were schematic, because he knew so little about the mechanisms of variation, development, and inheritance. Some of what Darwin left unaddressed has been filled in by modern biology. But it might also be argued that regardless of the details, what Kant saw is that mechanistic explanations are *in principle* insufficient here. Maybe a mechanistic biology needs to be supplemented with a different kind of understanding?

I accept that we might have *intuitions* that something is missing in a mechanistic biology. What those intuitions indicate is a different matter.

I said earlier that explanation has both objective and subjective sides. Here we need to look at the subjective side. One part of psychology studies the frameworks and habits that we naturally tend to use to make sense of different kinds of phenomena. This includes recent work on how we deal with both living things and intelligent systems (Medin and Atran 1999, Csibra 2003). Humans are primed to detect agents and agency in the phenomena around them, and will do so on quite slim evidence. These habits make evolutionary sense; they are plausible parts of an evolved psychological tool-kit. In our ancestors' lives, it would have been very important to be tuned in to the intentions, plans, and designs of other agents, both seen and unseen.

This may have a number of effects on how we think about biological systems. In one context, it shows up as a sense that Darwinian explanations and mechanistic biology miss something about living systems. Kant tried to add some detail here; he was interested in the sense we have that in a living system, the parts exist *for* a whole. But the idea that there is an extra kind of purpose or "intendedness" in biological systems can be rejected as an illusion with a psychological basis that itself makes evolutionary sense. The fact that we can explain the intuition's origins in this historical way does not itself imply that the intuition is wrong, of course. But if we have independent reasons to think it is wrong, this helps explain its persistence.

There is also a different manifestation of the same set of psychological habits. This is the tendency to re-describe evolutionary processes themselves in purposive and agential terms. This is seen within quite diverse pictures of how evolution works. The agents recognized might be organisms, species, or individual genes.

The agential perspective on evolution has always been an uneasy mix of the metaphorical and the literal. It takes different forms, but all talk of benefits and agendas comes with a peculiar psychological power, as it engages the concepts and habit we use to navigate the social world. Once switched on, these habits are hard to switch off. When we are introduced to the possibility of understanding something in terms of a *rationale*, we become reluctant to settle for less. One set of goals and agendas might be exchanged for another, but this becomes the *kind* of understanding we are after. Other information is relegated to a secondary role.

Agential description of evolution can accompany real theoretical progress. This language can be used to quickly express contrasts between theories and models that can then be described more precisely. It can be used to steer the listener away from one family of models and explanations, to another. I don't deny its communicative role, and perhaps a heuristic role in exploring options quickly. In the mid to late 20th century, a change in the application of agential terms to evolution accompanied shifts in evolutionary thinking that were important. Some mid 20th century biology had seen an uncritical treatment of high-level entities, such as groups, species, and ecosystems, as evolutionary units. The reaction against that tendency featured close attention to evolutionary change at the level of individual genes (Hamilton 1964, Williams 1966). Genes became new objects of agential description – tiny and invisible strategists.

Two kinds of agential narrative have a special psychological potency. The first is a *paternalist* schema. Here we posit a benevolent agent, often a large one, who intends that all is for the best. This category includes various Gods, the Hegelian "World Spirit" in philosophy, and stronger forms of the "Gaia" hypothesis, according to which the whole earth is a living organism. The second schema is a *paranoid* one. Now we posit hidden agents, often small, pursuing agendas that cross-cut or oppose our own interests. Examples include demonic possession narratives, the sub-personal creatures of Freud's psychology (superego, ego, id), and selfish genes and memes. And while it is true that sometimes there *are* large and kind agents or small and nefarious ones at work, the psychological appeal of these ideas means that we tend to take up such stories too readily and run with them too far.

The account of evolution in terms of "selfish genes" (Dawkins 1976) is a paranoid narrative of this kind. It relegates other entities in evolution, such as whole organisms, to the role of mere "vehicles."[4] This is a situation where a communicative device or heuristic has been allowed to take on too substantial a role; it becomes a kind of foundational description. Instead, the way to think of gene-level evolutionary processes is like this. Any collection of entities which vary, inherit characteristics in reproduction, and differ in how much they reproduce will evolve by natural selection. These include entities bigger than us, like social groups, entities smaller than us, like cells and genes, and organisms like us. As long as they satisfy the requirements of variation, heredity, and fitness differences, they will behave in a Darwinian way. The recognition

[4] We could try to look for even smaller plotters. Is the sugar-phosphate backbone in DNA there to serve the bases' purposes?

that genes have the necessary features – they vary, inherit features in replication, and differ in how much they are replicated – is the recognition of one of Darwinian population among others. It is not true that when we find small things doing this, inside us or underneath us, we're finding *what it's all about*, what it all means, the agents whose plots and programs are behind everything else we see.

*   *   *

References (Incomplete)

Bouchard, F. (2008). "Causal Processes, Fitness and the Differential Persistence of Lineages." *Phil. of Science.*

Csibra, G. (2003). "Teleological and Referential Understanding of Action in Infancy." *Phil. Trans. R. Soc.*

Dawkins, R. (1976). *The Selfish Gene.*

Dawkins, R. (1982). *The Extended Phenotype.*

Dawkins, R. (1986). *The Blind Watchmaker.*

De Vries, H. (1906). *Species and Varieties: Their Origin by Mutation.*

De Vries, H. (1909). "Variation." In *Darwinism and Modern Science* (ed. Seward)

Francis, R. (1996). *Why Men Won't Ask for Directions: The Seductions of Sociobiology*. Princeton.

Godfrey-Smith, P. (2009). *Darwinian Populations and Natural Selection.*

Hamilton, W. (1964). "The Genetical Evolution of Social Behaviour." *JTB.*

Kant, I. (1790). *Critique of Judgment.*

Lewens, T. (2007). *Darwin*.

Lewontin, R. (1970). "The Units of Selection."

Medin and Atran (eds.) (1999). *Folkbiology*.

Neander, K. (1995). "Pruning the Tree of Life." *Br. J. Phil. Sci.*

Williams, G. (1966). *Adaptation and Natural Selection.*