DUBLIN CITY
UNIVERSITY
Ollscoil Chathair Bhaile Átha Cliath

Essays on Darwinism
2: Organismic Darwinism

Barry McMullin

April 1992

School of Electronic Engineering

TECHNICAL REPORT: `bmcm9202`

Barry McMullin,
School Of Electronic Engineering,
Dublin City University,
Dublin 9,
IRELAND

Telephone: +353-1-704 5432
Fax: +353-1-704 5508
E-mail: `McMullinB@DCU.IE`

Abstract

Artificial Life is, by its nature, an interdisciplinary research programme; it will involve biologists, of course, but also philosophers, mathematicians, chemists, computer scientists—and perhaps even (as in my case) engineers. Success in our endeavours will require some of us, at least, to venture into foreign territory. This essay is a log of my personal expedition into evolutionary biology. I attempt to provide a comprehensive review of Darwinism in the biological world, and to do so as an *outsider*—a non-professional in the field. My purpose is twofold. Firstly, I hope that real biologists may take this opportunity to correct at least the worst of my errors. Secondly, I offer this to other non-specialists as a sort of map—a record of my particular exploration. I hope it might at least provide some insight into the kinds of questions which need to be asked, even if the particular answers suggested here are less than satisfactory. Above all, I want to convince any who may be in doubt that *Darwinism* encompasses a complex and subtle system of interrelated theories, whose substantive transplantation to any artificial medium will be very far from easy.

This essay draws on abstract concepts introduced in a previous essay (McMullin 1992); the two essays are therefore best read in conjunction.

Contents

1 Introduction

> Another curious aspect of the theory of evolution is that everybody thinks he understands it!
>
> Jacques Monod
> (Quoted by Dawkins 1989b, p. 18)

This essay provides a review of (neo-)Darwinian theory, in its native, biological, setting. I suggest that this is an essential pre-requisite for any attempt to apply Darwinian principles outside their original context. The essential idea is to view the biological world as an example of what I have previously called a *D-system* (McMullin 1992), with biological organisms as *D-actors*—hence *Organismic* Darwinism.[1]

I make a distinction between two questions: firstly, whether there exists an evolutionary or phylogenetic *tree*; and secondly, given the existence of such a tree, the elucidation of particular aspects of it—preeminently, the emergence of "adaptive complexity" (or knowledge). It is the latter question which is of critical interest and importance, and which Darwin sought to answer with his theory of natural selection.

I discuss the detailed structure of the theory of Organismic Darwinism, establishing that some components may be much more easily tested and corroborated than others.

I review the various theories which seem to be plausible alternatives or competitors to Organismic Darwinism, but conclude that none of these currently provides a significant challenge.

I consider the *scientific status* of Organismic Darwinism—particularly claims that it may be tautological and/or metaphysical. I argue that these suggestions are mistaken, but *not* trivially so. Consideration of them will clarify some further important aspects of Organismic Darwinism. The resulting claim is that Organismic Darwinism is *incomplete* (as a theory or explanation of the growth of adaptive complexity) but that, in an essential sense, it is *incompletable*, even in principle. It is asserted that a complete or universal theory of the growth of adaptive organismic complexity would amount to a logic of induction, and will therefore remain forever beyond our grasp. This may be of relatively little significance within biology; it is, however, of crucial importance to any research programme aimed at realising the spontaneous growth, by Darwinian means, of adaptive complexity in *artificial* systems.

2 On the Origin of Species...

2.1 Two Theories

> In considering the Origin of Species, it is quite conceivable that a naturalist, reflecting on the mutual affinities of organic beings, on their embryological relations, their geographical distribution, geological succession, and other such facts, might come to the conclusion that each species had not been independently created, but had descended, like varieties, from other species...
>
> Darwin (1859, Introduction)

This is Darwin's original statement both of the theory of evolution, and the problem(s) which it is intended to solve. Remarkably, this is as valid and incisive a formulation today as when Darwin first presented it; indeed, it is still widely quoted (e.g. Lewontin 1978). Nonetheless, Darwin's sheer economy of expression disguises its subtlety. Therefore I should like to dissect it carefully, restating and/or elaborating the various elements.

First, Darwin summarises the problem, which is:

P_b **(Biology):** To explain the various *regularities* which are manifested in the biological world, especially the relationships which exist between species,[2] both as they exist now, and as they have existed in the past.

Then Darwin introduces two theories which may compete to solve this problem:

T_t **(Phylogenetic Tree):** Species arise by descent from other different species.[3]

In this case there must exist a unique evolutionary (phylogenetic) tree, ultimately leading back to a single common ancestor of all species. The process by which this putative original species came into existence is left open.

T_c **(Creation):** All species arise *separately*, and are immutable (cannot give rise to other species by descent). The process by which species do, in fact, arise, is left open.

---

[1] My usage of "organismic" here is loosely related to the *Organismic Evolution* of Beurton (1981), and the *Organismic Selection* of Wright (1980).

[2] There are difficulties with the exact interpretation of *species* in biology; for my purposes it will be sufficient to take the fairly simplistic line that a species is a set of organisms whose members can be reliably identified in some (unstated) fashion.

[3] I shall speak, loosely, of species descending from other species, as a shorthand way of saying that organisms of one species are (ultimately) descended from organisms of another species. I postpone any consideration of how many generations may be required to complete such a speciation event.

> The distinguishing feature of this theory is that organisms of any one species are not related *by descent* to organisms of any other species.

T_t is, precisely, what I had in mind when I said that Darwin had here stated *the theory of evolution.* It is worth emphasizing that Darwin did not *originate* T_t—any more than he originated T_c—and nor did he claim to. As Burrow put it, in his editor's introduction to a modern reprinting of *The Origin*:

> The theory of evolution was already an old, even a discredited one. Darwin, in later editions of *The Origin*, listed over thirty predecessors and was still accused of lack of generosity.
>
> Burrow (1968, p. 27)

In any case, I shall avoid any further reference to "the theory of evolution", as such, and will use the more austere "T_t" instead. I do this for several distinct reasons:

1. Darwin himself did not originally use the word "evolution" at all, because, at that time, both conventional and technical usage of this word was quite different from Darwin's notion of descent (with modification). This has been discussed in detail by Gould (1978, Essay 3).

2. T_t prescribes only that new species may arise from earlier species by descent; it does not specify that this process is necessarily incremental or gradual—i.e. that it is literally *evolutionary* as the term is now commonly understood. In particular, the fossil record (Darwin's "geological succession") does not offer direct evidence capable of deciding this issue with precision (even assuming that there is a single "right" answer).

3. Although I have characterised the two theories as involving *creation* versus *descent*, *both* theories have to posit *some* "creation" event(s)—in the technical sense of species coming into existence *other* than by descent from earlier species. The significant feature of T_t relative to T_c is not that it eschews creation altogether, but that creation is invoked only in a relatively minor, secondary, rôle. As we will see, this is decisive for the comparison between the two theories.

4. Finally, phrases like "the theory of evolution" (or perhaps "The Theory of Evolution"), "evolutionary theory", and even just "evolution", have become vague and confusing through popular use and abuse. Thus, for example, Gould (1983, Essay 19) argues forcefully for the view that the object I have identified as T_t should *not* be called a "theory" at all but rather a "fact"—on the basis that it has been "confirmed to such a degree that it would be perverse to withhold provisional assent". Gould reserves the term "theory" in this context for hypotheses relating to the *mechanisms* regulating the development of the phylogenetic tree. I do not disagree with the substance of Gould's position, but, for the purposes of discussing whether or not something is a "fact" in his sense, we must first formulate it, and it is, at that stage, only a "theory"—and that is the stage I am at in this section, so to start off by using Gould's terminology would, itself, be perverse.

At this point I should like to compare T_t and T_c—to decide, if possible, which we should *prefer.* To do this, I shall rely (albeit loosely) on Popper's notion of *degree of corroboration* (Popper 1959; 1983).

First let us consider the relative *corroborability* of the two theories—that is, the *a priori* degree to which they *might* be corroborated. This is equivalent to their *testability*, which, in turn, is more or less determined by their *information content*—the sets of non-tautologous consequences which flow from them.

Now, as it happens, T_c has zero corroborability—there are *no* non-tautologous consequences of it. That is, if we suppose that all species arise separately, and not by descent, then, in the (self-imposed) absence of some theory of the creation mechanism itself, we cannot predict *anything* about what relationships should exist between species, present and/or past; which is to say that T_c does not (yet) address P_b at all. In Darwin's formulation, this aspect is implied by his reference to *independent* creation.

T_t, on the other hand, makes some fairly definite, non-tautologous, predictions, and is therefore testable and corroborable. The basic prediction is that some single, unique, phylogenetic tree must be compatible with *all* manifestations of that tree. These manifestations are, precisely, the regularities cited in P_b. T_t is, at least, confronting our problem, if not actually solving it.

So, T_t has a higher corroborability than T_c. Indeed, there exists a sort of continuum of theories, between T_c and T_t, parameterized by the number of "original" species which are posited; the corroborability is (more or less) inversely related to this. I have deliberately placed T_t at the limit of this range; that is, in the strongest, boldest, form of stating that *all* species, present and past, are descended from just *one* original species. This may seem stronger than necessary—but there is no particular merit in caution *at this stage* (i.e. unless and until T_t fails one or more critical tests). This is essentially the position originally adopted by Darwin himself:

> I believe that animals have descended from at most only four or five progenitors, and plants from an equal or lesser number.
>
> Analogy would lead me one step further, namely to the belief that all animals and plants have descended from some one prototype. But analogy may be a deceitful guide. Nevertheless all living things have much in common, in their chemical composition, their germinal vesicles, their cellular structure, and their laws of growth and reproduction. We see this even in so trifling a circumstance as that the same poison often similarly affects plants and animals; or that the poison secreted by the gallfly produces produces monstrous growths on the wild rose or oak-tree. Therefore *I should infer from analogy that probably all the organic beings which have ever lived on this earth have descended from some one primordial form,* into which life was first breathed.
>
> Darwin
> (1859, Chapter XIV, emphasis added)

Now it is important to note that, even though T_t has a higher corroborability than T_c this is, as yet, of limited significance; we have not yet spoken of the degree to which T_t has *withstood* its tests to date—i.e. its actual (as opposed to potential) degree of corroboration. In particular, even though T_c is prescientific (metaphysical), we should still prefer it if T_t fared very badly in its tests (if, in effect, T_t was *refuted*).

Technically, while T_c has zero corroborability, and therefore zero corroboration, T_t could have a *negative* corroboration (ultimately, if it were decisively refuted, the corroboration would be exactly -1); we could thus find ourselves actually *preferring* T_c.

Of course, our response to such a situation would depend on the exact manner of T_t's refutation; we might be able to rescue it simply by moving a little way along the continuum toward T_c. However, if the refutation were profound or extensive, such *ad hoc* immunization, though always possible, would be counterproductive. In that situation we would want to confront T_c, by augmenting it with some specific theories of the "creation" process—theories which *would* make some definite predictions relating to P_b, and would thus have non-zero corroborability.

In any case, this is something of a digression, because T_t has, in fact, withstood its tests very well, and I should say that it is currently "well" corroborated.

This conclusion includes, of course, the fossil evidence. As Dawkins has recently put it: "If a single, well-verified mammal skull were to turn up in 500 million year-old rocks, our whole modern theory of evolution would be utterly destroyed" (Dawkins 1986, Chapter 9, p. 225). On the other hand, the fossil record is by no means complete, and this limits the severity of the testing by this means—as was known and acknowledged by Darwin himself (Darwin 1859, Chapter IX). In fact, the primary tests of T_t are not based on fossil evidence, but on evidence from the observed evolution of species, from biogeography, and from the hierarchical structure of organismic taxonomy. These are essentially the same kinds of tests as Darwin himself formulated, although the contemporary versions have become a good deal more severe, as discussed in detail by Ridley (1981).

Additionally, it has recently become possible to test T_t in a manner which would not even have been intelligible in Darwin's time—by the comparison of protein sequences in putatively related species. T_t has, in fact, continued to withstand these even more severe tests (see Penny *et al.* 1982).

So we can conclude that we should rationally prefer T_t to T_c, because of its decisively greater degree of corroboration; but let me reiterate, one last time, that this is nothing to do either with some *inherent* "implausibility" of T_c, nor with some refutation of it. T_c, as it stands, *cannot* be refuted, and it would be pointless to pretend otherwise. Granted, T_c is not (yet) scientific; but it could be elaborated into a variety of scientific theories. Our reason for not pursuing that avenue is simply that T_t is *already* a stronger (scientific) theory, which *is* successfully solving the relevant problems (i.e. has not been refuted). We will *need* to re-assess this position only if, or when, some additional elements are added into the problem situation.

All of this extended argument, now coerced into the Popperian model of scientific discourse, was, then, already implicit in Darwin's straightforward assertion that it is "quite conceivable" that one would prefer T_t to T_c, as a solution to P_b. Nonetheless, I believe that, in the intervening years, there has been such a degree of confusion if not obfuscation of this original argument, that the detailed analysis offered here is justified.

2.2 Two Digressions

2.2.1 Creationism

At this point, I should like to briefly digress to relate the development so far to the doctrine of *Creationism.*

First consider what I shall call *Biblical Creationism* i.e. a more or less literal interpretation of the Book of Genesis. Note that this is *not* equivalent to T_c; it is stronger than T_c, and is, in fact, a perfectly good scientific theory (in the sense of having testable consequences).

In particular, Biblical Creationism posits that all biological species were created concurrently and more or less instantaneously (for my purposes it does not matter how, why, or when), and that the only change since then has been the possible extinction of some of the originally created species.

Certainly, this leaves a lot of questions open (particularly as regards the nature of the creation event), but, insofar as it goes, it is still clear and testable. Specifically, it entails clear predictions about what we should expect the fossil record to look like (the *only* allowed change over time would be a thinning out of the varieties of fossil observed). Since the fossil record has been found, repeatedly and consistently, to refute this prediction, this theory has been decisively refuted (relative to, among others, T_t).

Now, the theory could be saved by dropping the requirement for instantaneous, or simultaneous, creation (albeit, we would have to sacrifice our literal conformity to Genesis, to do this). But such a *Creeping Creationism* as it might be called, essentially *is* identical to T_c; and, as we have seen, T_c is not (currently) a serious competitor to T_t. Of course, Creeping Creationism is typically dressed up a lot more than T_c—by saying, for example, that the creation events have been decided or determined by some god, in her infinite wisdom; but, in the absence of a (testable) theory of the powers and motivations of this god, this is not logically distinguishable from T_c—it is merely more verbose.

Of course (and this really *is* a bold leap!) we could still hypothesize, in effect, that we are the hapless (but rational) victims of some diabolical practical joke—that Biblical Creationism really is true, but that some agent, with a rather warped sense of humour, has deliberately contrived the assorted empirical evidence which contributes to the refutation of Biblical Creationism and the corroboration of T_t. Presumably, such an elaborate hoax would have been designed with the specific *intention* of tricking us mere mortals into erroneously adopting T_t. Such an idea seems to me to be on a par with the theological speculations of Douglas Adam's fictional philosopher *Oolon Colluphid*.[4] On the other hand, there has apparently been at least one serious proposal to this general effect (reported by Little 1980).

In any case, I suggest that it serves no useful purpose to unilaterally *deny* or denigrate theories like T_c (e.g. by *insisting* on describing T_t as a fact, as discussed with reference to Gould in the previous section), or to pretend that they can be (or have already been) absolutely refuted. I consider that it is better simply to observe that such (metaphysical) theories *cannot* be refuted, that some such theory *may*, in fact, be true, but that, notwithstanding all that, T_t can and should be rationally *preferred* to them.

2.2.2 Saltationism

Within the scope of T_t we can identify various more specific theories distinguished by claims about the "rate" of evolution.

Firstly we may consider *gradualism* which asserts that species are formed gradually and cumulatively over many generations—i.e. the differences which may arise between parent and offspring in any *single* generation are much smaller than the differences which ultimately distinguish or separate species. In contrast to gradualism, we may consider *saltationism* which asserts that new species can arise in a small number of generations, perhaps even just one. Darwin certainly adopted the former, gradualist, view (particularly by his invocation of the slogan *Natura non facit saltum*, or *nature does not make leaps*; the slogan is usually attributed originally to Linnaeus—see, for example, Gould 1980, Essay 17). For the moment I simply note that the fossil record is too fragmentary to decide this issue, but that other sources of evidence (observed evolution, biogeography) seem to favour something closer to gradualism than saltationism.

Secondly we may consider whether the rate of evolution is itself constant, or varies significantly. Here we encounter the theory of *punctuated equilibria* of Eldredge and Gould (Eldredge & Gould 1972; Gould & Eldredge 1977). This asserts that the rate of evolution varies greatly, and that species remain static for most of the time (equilibrium), but with occasional bursts (punctuations) of relatively rapid evolution and/or speciation. The key phrase here is "relatively rapid". Eldredge and Gould are not

---

[4] Well known author of the trilogy of philosophical blockbusters *Where God Went Wrong, Some More of God's Greatest Mistakes* and *Who is this God Person Anyway?* (Adams 1979).

saltationists—they are operating within the framework of gradual, cumulative change over a large number of generations. This question of punctuated equilibrium versus constant speed evolution has been reviewed in depth by Dawkins (1986, Chapter 9). The conclusion is that the weight of the evidence favours something closer to punctuated equilibrium rather than constant speed evolution (although with some exceptions). The point for my purposes here is simply to note that punctuated equilibrium is not in any sense an alternative or competitor to T_t; it is rather a further refinement within the gradualist version of T_t.

3 ... by Means of Natural Selection

3.1 Darwin's Problem (P_d)

> ... Nevertheless, such a conclusion [T_t], even if well founded, would be unsatisfactory, until it could be shown how the innumerable species inhabiting this world have been modified, so as to acquire that perfection of structure and coadaptation which most justly excites our admiration.
>
> Darwin (1859, Introduction)

With this continuation of the earlier quotation from Darwin (section 2.1) the problem situation is now extended, in a crucial way, beyond P_b. We arrive at what I shall term *Darwin's Problem*:

P_d **(Darwin):** Given that species descend from other species (i.e. given T_t), how is it that, in some cases at least, this descent has been accompanied by an increase in adaptive complexity?

I introduce the term *adaptive complexity* here, following Maynard Smith:

> The main task of any theory of evolution (*sic*) is to explain adaptive complexity, i.e. to explain the same set of facts which Paley used as evidence of a Creator. Thus if we look at an organism, we find that it is composed of organs which are at the same time of great complexity and of a kind which ensures the survival and/or reproduction of their possessor. Evolution theory must explain the origin of such adaptations.
>
> Maynard Smith (1969, p. 82)

Popper (1978) has similarly formulated a version of P_d by reference to Paley's famous "argument from design"—i.e. that the appearance of "Design" in the biological world "proves" the pre-existence of a designer or creator (which is to say, *God*). Popper emphasizes that Darwin himself acknowledged a strong influence from Paley's formulation (and attempted solution) of this problem. At one point Popper explicitly calls this "Paley's problem" (p. 345); but elsewhere in this same paper he refers to "Paley's and Darwin's problem" (p. 342), which is thus consistent with my designation of P_d as "Darwin's Problem". I shall consistently use the latter name; partly because I have used Darwin's formulation, rather than Paley's, to introduce it; but mainly because there is a subtle difference between the problem as *originally conceived* by Paley, and as *actually solved* by Darwin. In my construction (P_d), the (conjectured) truth of T_t (the existence of the phylogenetic tree) is part of the problem situation; whereas, for Paley, the truth of T_t was neither a necessary part of (nor even particularly relevant to) the problem situation. I shall return to this point below.

Dawkins has adopted the terminology of "adaptive complexity" from Maynard Smith (Dawkins 1983, p. 404), and I shall use this phrase freely in what follows, but I shall also synonymously (?), and sometimes preferentially, refer to "inate" or "inborn" *knowledge* in the sense introduced (somewhat separately) by Popper:

> I assert that every animal is born with expectations or anticipations, which could be framed as hypotheses; a kind of hypothetical knowledge. And I assert that we have, in this sense, some degree of inborn knowledge from which we may begin, even though it may be quite unreliable. This inborn knowledge, these inborn expectations, will, if disappointed, create *our first problems*; and the ensuing [somatic time] growth of our knowledge may therefore be described as consisting throughout of corrections and modifications of previous knowledge.
>
> Popper (1961, pp. 258–259)

A more comprehensive and formal analysis of the idea of an *anticipatory system* has been independently developed by Rosen (1985a).

P_d refers to an increase or growth of adaptive complexity. I thus implicitly assume some notion of "degree" of adaptive complexity; but I require only that this can be defined *a posteriori*. That is, *given* two systems, I suppose that I will, in general, be able to rank them (at least roughly) in terms of relative adaptive complexity. I specifically eschew any attempt to define an *a priori* measure of adaptive complexity. To this extent I shall be following Rosen:

> ... complexity is not just complication, to be described by another number (e.g., the dimension of a state space or the length of a program), but a whole new theoretical world, with a whole new physics associated with it.
>
> Rosen (1985b, p. 202)

However, I should add that I shall not attempt to use "complexity" in precisely the strict technical sense intended here by Rosen.

von Neumann has described a similar idea as follows:

> There is a concept which will be quite useful here, of which we have a certain intuitive idea, but which is vague, unscientific, and imperfect ... I know no adequate name for it, but it is best described by calling it "complication." It is effectivity in complication, or the potentiality to do things. I am not thinking about how involved the object is, but how involved its *purposive operations* are. In this sense, an object is of the highest degree of complexity if it can do very difficult and involved things.
>
> von Neumann (1966, Fifth Lecture, p. 78, emphasis added)

Maynard Smith has similarly endorsed this somewhat loose approach:

> At the outset we are faced with a difficulty: we have no way of measuring the degree of complexity of a structure. Thus although most of us would readily agree that the organs of a man are more complex than those of an amoeba, and those of an amoeba more complex than those of a bacterium, we have no agreed criteria on which to base this decision, and no way of deciding by how much one organism is more complex than another.
>
> It may therefore seem odd to start formulating a theory of evolution by introducing a term which cannot be fully defined. However, I see no escape from doing so. *If organisms were not both complicated and adapted, living matter would not differ from dead matter, and evolution theory would have nothing to explain.*
>
> Maynard Smith
> (1969, pp. 82–83, emphasis added)

I consider that P_d is the central, perhaps even the defining, problem of evolutionary biology, and is the problem Darwin hoped to solve with his theory of Natural Selection. To this extent, the title of Darwin's book, *The Origin of Species* may be slightly misleading—it directs attention specifically at the phenomenon of *speciation*, rather than at the growth of adaptive complexity. While speciation is certainly an important issue we can say nonetheless that it is of secondary importance compared to the question of adaptive complexity. Diversity of species as such would be of relatively little interest if all species were equally rudimentary; conversely, even if there were only one species which displayed adaptive complexity, this one species would be of overwhelming interest.[5]

Before attempting to solve P_d we must be clear as to how P_d goes beyond P_b; that is, in what respect(s) T_t fails to address (much less solve) P_d.

P_b was solely concerned with *relationships* of various sorts between species—it took species (current and extinct) as *given* and asked why they should exhibit the various patterns of similarity and divergence which they do. T_t solved this problem by postulating the existence of a phylogenetic tree, involving more or less close relationships of common descent among all species; by a suitable choice of this tree (and T_t conjectures precisely that such a choice is possible) then all the other relationships can be correlated with (*reduced to,* if you will) the single relationship of common descent.

P_d, on the other hand, then goes on to ask why the phylogenetic tree has the particular structure that it has? Why, above all else, does it display at least some cases of increasing adaptive complexity?

Note that the claim here is not that *all* evolutionary lineages have involved increasing adaptive complexity, but merely that at least *some* have. Similarly, even for those evolutionary lineages where there has been a net increase of adaptive complexity, there is no claim that this increase has occurred at a steady rate, or that it has been monotonic, or (most especially) that it will continue into the future.

P_d thus cannot even be properly formulated *except* in the context of adopting (always *tentatively*)

---

[5] This is not, of course, to argue that speciation and the growth of adaptive complexity are *independent* phenomena, but merely that I shall concentrate on the latter.

T_t. P_d may be viewed, in this sense, as a *consequence* of the particular solution we have adopted to P_b.[6] This is essentially the point made earlier, that the problem situation perceived by Paley was not (necessarily) quite identical with that perceived by Darwin.

Contrariwise, for a Creationist (in the sense of one who rejects T_t), P_d simply does not arise; some *other* problem(s) certainly do arise, but not P_d. To this extent, the Creationist "debate", whatever its subject matter, is not about P_d; and especially, it is not about deciding between competing solutions to P_d. So, Creationism *simpliciter* will play no further rôle in this discussion (though some related theories will still have to be considered). Gould (1988, Essay 17, p. 234) has previously made essentially this same point.

I shall present the proposed solution to Darwin's problem based on the abstract formulation of Darwinism given in (McMullin 1992) (I rely particularly on the notion of a *Similarity-lineage*, or *S-lineage,* introduced there; the discussion which follows will be almost unintelligible except by reference to that concept). I claim that this is essentially the solution offered by Darwin himself. However, for reasons which should become clear, I do not proceed directly to this solution, but instead build up to it in a number of distinct steps.

3.2 Contemporary Selection (T_{cs})

> However, Darwin's own most important contribution to the theory of evolution, his theory of natural selection, is difficult to test. There are some tests, even some experimental tests; and in some cases, such as the famous phenomenon known as "industrial melanism", we can observe natural selection happening under our very eyes, as it were. Nevertheless, really severe tests of the theory of natural selection are hard to come by, much more so than tests of otherwise comparable theories in physics or chemistry.
>
> Popper (1978, p. 344)

Consider the following:

T_{cs} **(Contemporary Selection):** The *contemporary* biological world is such that, with organisms regarded as *actors*, these actors form coherent S-lineages, with well defined S-values,[7] which can (and do) selectively displace each other.

Informally, T_{cs} amounts to the claim that Darwinian (S-lineage) selection is now occurring in the biological world, with biological organisms as actors. For the moment I make no comment about the nature of S-creation (except that some such events occur). In particular I do not (yet) specify that it is *unjustified.* In my terms, this is the distinction between an actor *per se* and a *Darwinian* actor (D-actor). I shall return to this in section 3.4 below.

I stipulate certain things (which seem relatively uncontentious) without further discussion:

- Organisms do procreate—indeed, they provide the paradigmatic case of procreation.
- Organisms are capable, in suitable circumstances, of exhibiting the required (Malthusian) growth.
- Their numbers are, in fact, limited by resource availabilities (possibly among other things). It follows that there must be competition between, and especially within, species, for at least some resources.

In themselves, of course, these facts do not establish that S-lineage selection occurs. To begin with, we further require that there must exist S-classes—which is to say classifications of organisms such that:

- With "high" frequency, an offspring organism should be a member only of S-classes to which at least one of its parent(s) belong (S-founders should be relatively rare, but must occur occasionally). This guarantees the ongoing formation of coherent S-lineages, between which selection *may* (or may not) operate.

---

[6] We have here a typical illustration of Popper's schema for the growth of scientific knowledge (Popper 1972, Essay 3):

$$P_1 \rightarrow TT \rightarrow EE \rightarrow P_2$$

P_1 denotes some "initial" problem situation (in our case, P_b); TT denotes the tentative theories invented (somehow) as candidate solutions (T_t, T_c, and, in fact, the entire continuum between them); EE denotes the process of criticism, testing, and error elimination (whereby we concluded that T_t was, for the time being at least, decisively better corroborated than any of its competitors); finally, this results in the establishment of a *new* problem situation P_2 (identified here with the emergence, and recognition, of P_d).

[7] S-value is my term for what is conventionally called *fitness* in population genetics. It is a quantitative measure of the "reproductive success" of an S-lineage, given particular environmental constraints. It must have the property that, when all other conditions for S-lineage selection are satisfied, the difference in S-values of the competing S-lineages is predictive of both the direction and the rate of selective displacement. The term is originally defined in (McMullin 1992).

- At least some S-classes should be such that corresponding S-lineages will selectively displace each other, in a manner which is at least approximately deterministic (in given environmental conditions); which is to say that selection will, in fact, operate.

The satisfaction of the first of these conditions depends crucially on the nature of inheritance in biological organisms. The modern theory of biological inheritance, grounded in molecular biology, is elaborate and complex. For my immediate purposes I shall consider just the following relatively simple form:

T_i **(Inheritance):** Organisms exhibit at least some characteristics (termed *heritable characters*) which have the following properties: with "high" frequency, an organism will exhibit one of these characters *only* if at least one of its parent(s) were similarly characterised; however, with "low" frequency an organism may exhibit a completely novel character, i.e. one not exhibited by any of its parent(s) (but which can, nonetheless, be inherited by its offspring).

In this form, T_i consists in little more than the assertion that there do, in fact, exist characteristics of biological organisms which, if used to classify the organisms, will satisfy the *inheritance* condition for S-classification, and the *novelty* condition for continuing evolution of a (D-)lineage. However there are some subtle ideas concealed behind the apparent simplicity of this formulation of T_i.

First note carefully that there *is* some worthwhile empirical content in T_i: it need not be true (even though it is "almost" inherent in our very conception of what a biological organism is). For example, we can (just about) conceive of a world in which there are living organisms, and in which the organisms procreate, but where there are *no* characteristics whose inheritance is consistently (with "high frequency") constrained in the way prescribed by T_i—for example, if "elephants" gave "birth" to "mice" or "bacteria" (or arbitrary other "classes" of organism) more or less as commonly as to other "elephants". Somewhat more plausibly, we can imagine that there could be characteristics whose inheritance is constrained as required by T_i (which is to say, there would exist coherent S-lineages), but where no *new* characteristics of this sort ever come into existence. This would effectively amount to the existence of static or immutable biological species, contrary to T_t; to this extent, T_i can be viewed as being, at least partially, a consequence of T_t.

Secondly, note that T_i takes the form of a *prohibition* on the nature of inheritance, rather that a positive affirmation. That is, we do not explicitly say that heritable characters *will* be inherited by offspring (not even if *both* parents share a particular character); rather we say that novelty (of heritable characters) is rare—that, almost always, an organism will *not* exhibit a character unless at least one parent already exhibited it. Of course, given that, in practise, an offspring has to exhibit *some* complement of heritable characters (in order to qualify as an organism—a member of a particular class of biological actors—at all), it will be the case that, almost always, an offspring will exhibit some (perhaps even "most") parental characters; the point is simply that T_i does not require that any *particular* character be inherited by any *particular* offspring.

Indeed, there could, in principle, be good, heritable, characters, satisfying T_i, which are nonetheless "rarely" inherited by offspring; this would simply mean that an S-lineage identified by such a character would have relatively low S-fecundity—which may mean that it is not viable at all (i.e. be incapable of exhibiting Malthusian growth, in *any* circumstances), or that it will be displaced relatively easily by competitors. However, none of this would detract from the heritable nature of such a character.

Finally, note that I am deliberately using the term *character* here, and not the modern term *gene* which might seem to be essentially equivalent. This is a fine terminological distinction, but an important one.

"Character" is the term introduced in Mendel's original paper (or, more accurately, introduced in the English translation of that paper, made by the Royal Horticultural Society of London, Mendel 1865). By this Mendel meant simply *any* characteristic of an organism which (*inter alia*) satisfies T_i (there are additional constraints on what I shall subsequently call *Mendelian* characters, but that will be clarified in due course). He did *not* speculate as to the underlying *mechanisms* whereby such characteristics are exhibited by organisms, or are inherited. Davern makes this point in his introduction to the reprinting of Mendel's paper, cited above:

> ... as Robert C. Olby has recently pointed out ... Mendel did not infer the existence of *determinants* of heredity that *conserved their integrity* in hybrids and successive generations. Instead, Mendel argued that the parental characteristics *per se* were conserved in the hybrid generations. While he apparently failed to dis-

> tinguish between the potential and actual when it came to accounting for the manifestation of parental characteristics, his discoverers in 1900 certainly made this distinction. They explained the phenomenon of dominance and the patterns of segregation for a contrasting characteristic pair in the hybrid generations in terms of genetic determinants or *genes*, thus clearly distinguishing between genotype and phenotype.
>
> Davern (1981b, p. 6, emphasis added)

Contrary to the confident tone struck here by Davern, I take the view that the step from overt characteristics of organisms to supposed "determinants" of these characteristics is distinctly problematical; it is associated, in Jensen's terms, with a dangerous (if not actually erroneous) "reification" of the gene (Jensen 1981). This is a complex topic which I shall not, however, pursue here; for the moment, it is sufficient to note that I make no claims (in T_i or otherwise) about the underlying mechanisms whereby heritable characters are exhibited and/or inherited.

It follows, incidentally, that I do not restrict "character" to denote the nucleotide sequence of some fragment of DNA (i.e. what is now typically regarded as a "gene"). My "character" *could* refer to a gene in this sense, but it could equally refer to at least some "phenotypic traits". The exhibition of a character, in my terms, could (for example) correspond to the presence of a particular protein or a particular organ, or the performance of a particular behaviour etc. In these cases there may be some simple correlation between what I call a character and a gene, but this is not necessarily the case. Thus, for example, karyotypic characteristics (such as chromosome number) could be heritable characters in my terms. I require *only* that it be rare for an organism to exhibit such a character unless at least one of its parents already exhibited it. This is sufficient for the establishment of S-lineages (though it does not, in itself, guarantee S-lineage *selection*).

I shall take it that T_i is well corroborated.

The next question is whether any S-lineages, based on heritable characters, can potentially "compete" with each other, such that one can competitively eliminate the other. This requires three conditions:

1. S-lineages should be limited in S-size by availability of some resources.

2. At least some S-lineages should be so limited by the *same* resources.

3. At least some such S-lineages should be constrained such that their memberships cannot (completely) converge. That is to say, their memberships are constrained not to be equal.

Together these three conditions ensure that there exist some S-lineages whose S-sizes are so related that, *if* one persistently increases, the other must, at some point, reduce and thus be ultimately eliminated—i.e. some S-lineages may selectively displace others. Note that these conditions are thus necessary but not sufficient for S-lineage selection.

I take it that the validity of the first two of these conditions is unproblematic in the biological world and does not require further discussion. However, the third condition requires much more careful consideration.

We can certainly say that if certain S-lineages are inherently disjoint (their intersections are empty) then, trivially, their memberships cannot be equal. Thus, for asexually reproducing organisms (i.e. where S-procreation is unimodal), where the S-lineages are, by definition, disjoint, we can say that condition 3 above will be satisfied. More generally, if the S-lineages consist of organisms from distinct (sexually reproducing) species (where distinct species are taken to be reproductively isolated, by definition), then, by a similar argument, they will be disjoint, and condition 3 will again be satisfied. I therefore take it that *at least* these kinds of exclusion between S-lineages can occur in the biological world (though, of course, these are precisely the cases where it may be difficult to establish that the S-lineages involved rely on the same resources).

The situation of S-lineages within a single (sexually reproducing) species is more complex, though it is (arguably) the more important. Certainly, such S-lineages will, almost by definition, share many of the same resources, which is to say that condition 2 above can be almost trivially satisfied. But since these S-lineages are not reproductively isolated from each other, they need not, and generally will not, be disjoint; so it is not at all clear that condition 3 can be satisfied in such cases.

If it were the case that heritable characters could, in general, be divided into sets such that the characters in any single set are mutually exclusive (i.e. any given organism could exhibit only one of the characteristics in the set), then S-classes defined by reference to distinct elements of any such set would be disjoint, and so also would be the corresponding S-lineages. Again, condition 3 would be satisfied. Unfortunately (?) it turns out that heritable characteristics do *not*, in general, satisfy such mutual exclusion conditions under sexual reproduction.[8]

However, it transpires that at least some heritable characters show a relationship which, although it is weaker than mutual exclusion, can still be shown to suffice for condition 3 to be satisfied (in suitable circumstances). I shall discuss only the simplest case, but emphasize that this is only an example: similar reasoning may well apply to different and/or more complex cases. This simplest case corresponds essentially to a pair of what I shall call strictly *Mendelian* characters, in diploid organisms. These are heritable characters which are associated or correlated with distinct (Mendelian) genes at a single genetic locus, in a species of sexually reproducing diploid organisms.

A pair of such Mendelian characters is distinguished by the following properties:

- Firstly, every organism (of the given species) exhibits at least one of the pair, and may exhibit both. I denote the two characters as A and B; thus any given organism can be labelled AA (exhibits A but not B), BB (exhibits B but not A) or AB (exhibits both A and B). In conventional genetic terms, AA and BB correspond to the two homozygotes, AB to the heterozygote. I assume the heterozygote to be distinguishable as such (in the case of "complete" dominance of one gene over the other, this may require a relatively elaborate process, such as a generation of controlled breeding, or DNA sequencing etc.; such methodological difficulties, though possibly severe, have no conceptual significance here).
- Secondly, the character(s) exhibited by an offspring are found to be randomly distributed, approximately according to table 1 (the distributions in this simplest case are symmetrical with respect to the two parents; redundant entries are therefore omitted from the table).

We see that, according to table 1, the characters A and B do satisfy T_i, namely that an offspring will generally exhibit A (or B) only if at least one parent exhibited A (or B). Again, note carefully that I do not assert that *all* heritable characters obey a table such as the above (are Mendelian); there are many possible complications which I shall not pursue. I merely claim that at least *some* heritable characters (namely, those called Mendelian) behave in this way.

PARENTS		OFFSPRING		
		AA	AB	BB
AA	AA	1.00	0.00	0.00
AA	AB	0.50	0.50	0.00
AA	BB	0.00	1.00	0.00
AB	AB	0.25	0.50	0.25
AB	BB	0.00	0.50	0.50
BB	BB	0.00	0.00	1.00

Table 1: Mendelian Inheritance

Formally, this table could be interpreted as a stronger or more explicit form of T_i, for the case of sexually reproducing diploid organisms. I shall take it that this augmentation of T_i is well corroborated and need not be critically examined here.[9]

I can now adopt the following (putative) S-classes:

S-class A: All organisms exhibiting A (i.e. all organisms labelled AA or AB).

S-class B: All organisms exhibiting B (i.e. all organisms labelled BB or AB).

By inspection of the table, we can see that these S-classes then satisfy the inheritance condition on S-classification, namely that an offspring will generally be a member only of S-class(es) of which at least one S-parent is a member (S-creation will be relatively rare). Note that this would not hold if, say, we considered AA, BB, and AB organisms to define distinct classes: for if AA mates with BB (for example) the outcome will be consistently (probability 1) AB i.e. of a different class to either parent relative to such a classification.

As anticipated, the proposed S-classes are not disjoint, and so the corresponding S-lineages will not, in general, be disjoint either. However, changes in the S-sizes of the two S-lineages *will* be more or less coupled, according to the restrictions implied by the table above. Our problem is to establish whether this coupling is such that condition 3 above for S-lineage selection (non convergence of S-lineages) can be satisfied.

---

[8] There are, of course, exceptions; for example, in species with fully differentiated sexes (*dioecious* species—Maynard Smith 1989, p. 236), the heritable characters "male" and "female" do generally satisfy a simple mutual exclusion condition, and do establish disjoint lineages. This is true even though, from a genetic point of view, the heterogametic (typically male) sex carries genes "for" both sexes.

[9] I note, in passing, that the results of Mendel's own original, pea breeding, experiments (Mendel 1865) actually provided almost "too" good a corroboration, to the extent that there has been some suspicion that Mendel adjusted some doubtful or intermediate results so as to yield just the expected outcome; but that, in any case, the theory has effectively withstood very severe subsequent, independent, tests (Hardy 1965, pp. 88–89; Gribbin 1985, pp. 33–34).

The exact effect of the coupling will be stochastic and will also depend on the mating pattern (e.g. are AA organisms more likely to mate with other AA organisms than AB or BB?). However, for my immediate purposes I require only to establish the *possibility* that this coupling can prevent convergence of the two S-lineages. This can be achieved by considering just a single (plausible?) case: random mating, in a population which is large enough that the proportions of offspring will be more or less deterministically related to the probabilities indicated on the table. In this case it can be shown that the *relative* proportions of organisms produced in each generation will consistently be in a so-called *Hardy-Weinberg* ratio (Maynard Smith 1989, Chapter 3); that is, the (relative) numbers of organisms labelled AA, AB and BB will consistently be in ratios which are of the form $p^2 : 2pq : q^2$ respectively, where p and q are parameters of the population which are potentially time varying but are always such that $0 \leq p, q \leq 1$ and $p + q = 1$. p, q would normally be interpreted as the total relative densities of the *genes*, with which the characters are (putatively) correlated, in the *gene-pool*.

Note that, as pointed out by Maynard Smith, the Hardy-Weinberg form of ratio will be achieved after one generation of random mating, even if the parents were not in this ratio. It will therefore be consistently maintained (as measured by the relative numbers of offspring *born*, as opposed to, say, surviving to maturity), even while p and q are changing—for example due to selection. My claim here—that a ratio of the Hardy-Weinberg *form* will be consistently maintained, even while selection is in progress—should not be confused with the closely related, but distinct, idea that the *specific* Hardy-Weinberg ratio will be stable (the Hardy-Weinberg *equilibrium*) if p and q remain constant.

I assume that there is a single S-lineage present of each of the two S-classes (A and B). Let S_A and S_B denote the relative S-sizes of these two S-lineages (i.e. as fractions of the total population of organisms; for simplicity, I assume this total population size to be fixed). Let R_{AA}, R_{AB}, R_{BB} denote the relative proportions of AA, AB and BB organisms respectively. Note that, because the two S-lineages are not disjoint, we *cannot* say that $S_A + S_B = 1$, and thus cannot (directly) conclude that an increase in one must necessarily be associated with a reduction in the other. Conversely, note that classes based on AA, AB, and BB *would* be mutually exclusive, and we *can* say that $R_{AA} + R_{AB} + R_{BB} = 1$; but, since these do not serve to label S-classes (i.e. do not give rise to coherent S-lineages) we cannot use this as our basis for establishing the possibility of S-lineage selection.

However, we *can* now argue as follows:

$$\begin{aligned}
R_{AA} &= p^2 \\
R_{AB} &= 2pq \\
&= 2p(1-p) \\
R_{BB} &= q^2 \\
&= (1-p)^2
\end{aligned}$$

$$\begin{aligned}
S_A &= R_{AA} + R_{AB} \\
&= p^2 + 2p(1-p) \\
&= 2p - p^2 \\
\Rightarrow \frac{d\,S_A}{dp} &= 2 - 2p
\end{aligned}$$

$$\begin{aligned}
S_B &= R_{AB} + R_{BB} \\
&= 2p(1-p) + (1-p)^2 \\
&= 1 - p^2 \\
\Rightarrow p^2 &= 1 - S_B \\
\Rightarrow 2p\frac{d\,p}{dS_B} &= -1 \\
\Rightarrow \frac{d\,p}{dS_B} &= -1/2p
\end{aligned}$$

$$\begin{aligned}
\frac{d\,S_A}{dS_B} &= \frac{d\,S_A}{dp} \cdot \frac{d\,p}{dS_B} \\
&= (2-2p)(-1/2p) \\
&= 1 - 1/p \\
&\leq 0 \;\; \forall\, p \leq 1
\end{aligned}$$

That is, for any pair of S-lineages whose relative S-sizes are (for example) consistently constrained by the Hardy-Weinberg ratio, a change in the S-size of either one will necessarily be accompanied by an opposite (though generally unequal) change in the S-size of the other. This means precisely that the two S-lineages, though they may intersect, cannot converge. Thus, for such S-lineages (at least), condition 3 above, for the possibility of S-lineage selection (non-convergence of the S-lineages) will be satisfied.

To review briefly: the discussion so far has established that, under our current, well corroborated, theory of inheritance (T_i) in biological organisms, there exists the *possibility* for the formation of coherent S-lineages whose sizes are so coupled that they *could* exhibit S-lineage selection. This is true in general of S-lineages whose organisms are not interbreeding (because they are asexual, or of different sexual species), and is true for at least *some* S-lineages within single interbreeding (sexual, single species) populations—namely where the characters are Mendelian; I emphasize again that I have not argued, and do not claim, that strictly Mendelian characters provide the *only* possibility for S-lineage selection within interbreeding populations.

But this is not yet sufficient to establish that S-lineage selection either can or does take place in the biological world. We further require that at least some heritable characters be sufficiently well correlated with the S-values of the corresponding S-lineages that the S-sizes of some will grow more or less deterministically (at the expense of others). That is, we must recognise that there could exist characteristics which are heritable (perhaps even Mendelian) but yet have no particular effect, one way or another, on the relative S-mortality or S-fecundity of the S-lineages which are so distinguished. Such S-lineages would, of course, still grow or shrink but these effects would be essentially capricious (so-called *random drift*), and could not be interpreted as examples of S-lineage *selection.*

This is a very difficult question to answer; and one must beware of accepting a simplistic answer which effectively avoids the problem. Thus, it has been established, by experiments involving artificial selection, that biological organisms exhibit a very wide range of heritable characters;[10] but this does not, *in itself,* guarantee that any of them are such as to result in selection.

Indeed, we should be wary of *any* argument based directly or exclusively on this kind of artificial selection experiment, *if* the experiment relies on amplifying some pre-existing variation in a population; for the whole point of selection (as *I* have used the word at any rate) is that, if it operates, it *eliminates* certain variations; so variations we now observe cannot *generally* be variations which have supported selection in the past. In evaluating the significance of such experiments we should strictly restrict attention to only those heritable characters which have been experimentally observed to arise "spontaneously" (i.e. S-creation events). Darwin himself implied something of this sort when he wrote:

> This preservation of favourable variations and the rejection of injurious variations, I call Natural Selection. Variations neither useful nor injurious would not be affected by natural selection, *and would be left a fluctuating element,* as perhaps we see in the species called polymorphic.
>
> Darwin
> (1859, Chapter IV, emphasis added)

However, the decisive corroboration of the *possibility* of S-lineage selection (in the biological world) does not, in fact, rely on artificial selection experiments of this sort. It relies instead on the actual observation of the phenomenon in *natural* biological populations. Various cases have now been very well corroborated; the examples most commonly cited are the spontaneous development of industrial melanism, and of insecticide resistance (e.g. Stebbins 1977, pp. 90–92; Maynard Smith 1989, pp. 42–45).

I take it, in conclusion, that the occurrence of S-lineage selection, based on S-classes identified by reference to heritable characters (in the sense of T_i), is very well corroborated; which is to say, that T_{cs} has withstood the most severe tests to which it has been (so far) subjected.

3.3 Historical Selection (T_{hs})

Consider now the following extension of T_{cs}:

T_{hs} **(Historical Selection):** The biological world has *always* been such that, with organisms regarded as *actors*, these actors have formed coherent S-lineages, with well defined S-values, which have selectively displaced displaced each other. All (inate) organismic attributes which have shown sustained growth in the phylogenetic tree, have done so by virtue of being positively correlated (directly or otherwise) with S-value.

T_{hs} thus tentatively extends the scope of T_{cs}, from its (relatively) well corroborated application in the present, into the indefinite past. This kind of extension of present causes into the past was a central element in the *uniformitarianism* of the geologist Charles Lyell and, subsequently, Darwin himself (see, for example, Gould 1983, Essay 9).

This step is, of course, essential if we are ultimately to address P_d, for that problem is explicitly concerned with the historical formation of the phylogenetic tree. Nonetheless, it is well to keep T_{cs} and T_{hs} clearly separated, for they require quite different approaches to their corroboration.

---

[10] I accept this fact, but caution against its exaggeration; thus, for example, I doubt that the following claim could be effectively defended: "The fact that artificial selection works almost every time it is attempted indicates that there is genetic variation in populations *for virtually every characteristic of the organism*" (Ayala 1978, emphasis added). I say this simply because a phrase like "virtually every characteristic of the organism" seems to me hardly meaningful.

I have termed T_{hs} an extension of T_{cs}, which is to say that T_{cs} is entailed by T_{hs}, and, were it to be refuted, T_{hs} would likewise also be refuted; but given that T_{cs} has, so far, withstood its tests, we must then consider to what extent T_{hs} can be *separately* tested or corroborated. That is, is there any specific means of testing, not just that selection is operating now, but that it has operated into the indefinite past? The answer to this rests on whether we can make any predictions about observable attributes of organisms (either living now, or fossilised) based solely on T_{hs}.

To achieve this, note that T_{hs} explicitly affords no means for the sustained, cumulative, growth of any organismic (or, more properly, S-lineage) attribute *except* that it be correlated with S-value. Or, to put it another way, if it could be established that any organismic attribute had shown sustained, cumulative, phylogenetic growth, but was either uncorrelated, or negatively correlated, with S-value, then this would, *ipso facto*, be a refutation of T_{hs}. As Darwin put it (where his "species" should, in this particular case at least, be read as equivalent to my "S-lineage"):

> If it could be proved that any part of the structure of any one species had been formed for the exclusive good of another species, it would annihilate my theory, for such could not have been produced by natural selection.
>
> Darwin (1859, Chapter VI, pp. 228–229)

In a similar vein, Maynard Smith has offered the following scenarios:

> . . . if someone discovers a deep-sea fish with varying numbers of luminous dots on its tail, the number at any one time having the property of being always a prime number, I should regard this as rather strong evidence against neo-Darwinism. And if the dots took up in turn the exact configuration of the various heavenly constellations, I should regard it as an adequate disproof.
>
> Maynard Smith (1969, p. 86)

So T_{hs} is, in principle at least, potentially refutable.

Noting that no organisms with the clearcut peculiarities suggested by Maynard Smith have, in fact, been reliably reported we may say that, to this extent, T_{hs} has successfully stood up to testing. Indeed, we can go somewhat further, and note that many organismic characteristics *can* plausibly be interpreted as being positively correlated with S-value. As Maynard Smith points out:

> . . . if it were not the case that most organs can readily be understood as contributing to survival [reduced S-mortality] or reproduction [increased S-fecundity], Darwinism would never have been accepted by biologists in the first place. . .
>
> Maynard Smith (1969, p. 87)

Alternatively, we can also approach this question from the other direction and ask whether there exist organismic attributes whose existence seems inexplicable *except* by an hypothesized correlation with S-value. This was a point emphasized by Darwin, presumably because he had in mind the possibility of competing theistic theories—which could easily account for the existence of "perfection" in the biological world, but could offer no explanation for evident "imperfection". Thus, we have his comment:

> Nor ought we to marvel if all the contrivances in nature be not, as far as we can judge, absolutely perfect; and if some of them be abhorrent to our ideas of fitness. We need not marvel at the sting of the bee causing the bee's own death; at drones being produced in such vast numbers for one single act, and being then slaughtered by their sterile sisters; at the astonishing waste of pollen by our fir-trees; at the instinctive hatred of the queen bee for her own fertile daughters; at ichneumonidae feeding within the live bodies of caterpillars; and at other such cases. The wonder indeed is, on the theory of natural selection, that more cases of the want of absolute perfection have not been observed.
>
> Darwin (1859, p. 445)

However, notwithstanding these various facts, we must be somewhat wary in evaluating the degree to which T_{hs} can be successfully corroborated in such a manner.

A first problem is that there *do* exist evolutionary phenomena which are *prima facie* refutations of T_{hs}, in the form of so-called *orthogenetic trends.* These are trends in organism lineages which seem to continue up to and beyond a point where they are "obviously" detrimental to the organisms (and thus the S-lineages) involved; if this is a correct interpretation then it is definitely inconsistent with T_{hs}.[11]

---

[11] I restrict the use of the term "orthogenesis" to these cases, where the trend supposedly persists to the point

Gould graphically describes the problem as follows:

> The theory of orthogenesis became a touchstone for anti-Darwinian paleontologists, for it claimed that evolution proceeded in straight lines that natural selection could not regulate. Certain trends, once started, could not be stopped even if they led to extinction. Thus certain oysters, it was said, coiled their valves upon each other until they sealed the animal permanently within; saber-toothed "tigers" could not stop growing their teeth or mammoths their tusks.
>
> But by far the most famous example of orthogenesis was the Irish Elk itself. The giant deer had evolved from small forms with even smaller antlers. Although the antlers were useful at first, their growth could not be contained and, like the sorceror's apprentice, the giant deer discovered only too late that even good things have their limits. Bowed by the weight of their cranial excrescences, caught in the trees or mired in the ponds, they died. What wiped out the Irish Elk? They themselves or, rather, their own antlers did.
>
> Gould (1978, Essay 9, pp. 84–85)

Gould is, of course, here acting as a Devil's advocate on behalf of orthogenesis. In the essay quoted above, he goes on to discuss two quite plausible alternatives to orthogenesis as a explanation of the undeniably enormous antlers of the Irish Elk.

Firstly, it may be an artefact of a correlation between body size and antler size: Gould established that antler size seems to increase as a superlinear function of body size—so that if selection favoured increased body size (which is not *im*plausible) then exaggerated antler size could result by conventional selection, as long as the (presumed) disadvantage of extreme antler size was outweighed by the advantage (whatever it may be) of increased body size. Such a correlation is known as *allometry.* Although (as discussed by Gould) this solution to the problem of orthogenesis (a problem in the context of T_{hs}) was formulated in the 1930s by Julian Huxley, Gould states that it had not actually been quantitatively tested—in the case of the Irish Elk at least—until he himself came to do so; but, in any case, the measurements he carried out did prove consistent with the allometric explanation.

Secondly, Gould goes on to question the very claim that the large antler size of the Elk was necessarily selectively disadvantageous. He points out that the antlers could have played a rôle in ritualised combat, where extremely large sizes could quite plausibly have been favoured by selection (the allometry argument *could* then run in reverse: there was not necessarily any selective advantage in large body size *per se* but increases in body size were correlated with the increases in antler size which *were* selectively favored). Gould is able to adduce some limited evidence in favour of this hypothesis from detailed anatomical features of the antlers.

In any case, a definitive decision between these two explanations is not necessary (and probably not even possible). The point is that the claim for an orthogenetic trend (not effectively controlled by selection) has clearly not been established, and T_{hs} emerges more or less unscathed. Gould has elsewhere similarly debunked the case of the coiled oysters (Gould 1983, Essay 30).

However, while this deals with cases where evolution seems positively to have run counter to selection, there is another class of cases which demonstrate a quite different kind of difficulty. These are organismic attributes which seem to demand an explanation in terms of T_{hs} (due, for example, to their complexity), but for which no consensus has yet emerged as to the nature of such an explanation.

The primary example of this is, perhaps, the existence of sex (in the sense of meiosis—not the existence of distinctive male and female forms). This is evidently a very complex process, which appears to have evolved over time; the *assumption* is that it has done so through natural selection; but a definitive explanation of why sex should have been thus favoured by natural selection is lacking. In 1975 Williams went as far as to claim, in this regard, that "there is a sort of crisis at hand in evolutionary biology" (Williams 1975, p. v). More recently, Dawkins has stated that: "The problem of what sex is good for is as tantalizing as ever" (Dawkins 1989b, p. 274).

This does not, of course, represent a refutation of T_{hs}—but it does perhaps indicate some degree of vulnerability.

---

of being detrimental to the organisms. However, a *general* concept of "trends" in organism lineages can also be called orthogenesis—see, for example, (Popper 1961). Such a weaker form of orthogenesis is certainly not inherently incompatible with T_{hs}; indeed Popper himself goes on to suggest a very general interpretation of the phenomenon which would be completely compatible with T_{hs} (provided always that the trend does not proceed to the point of being detrimental to the organism). I will not, therefore, refer further to this weaker concept.

A more subtle problem with the supposed corroboration of T_{hs} arises due to the ease with which more or less *arbitrary* organismic attributes can be argued to be positively correlated with S-value. Thus while I stated above that no *definitive* explanation of sex, in terms of correlation with S-value, has yet been forthcoming, there is no shortage of more or less plausible *hypotheses* (indeed, I shall comment further on one in particular, in section 3.4 below). Gould and Lewontin have discussed this general problem in the form of what they call the *adaptationist programme*:

> We wish to question a deeply ingrained habit of thinking among students of evolution. We call it the adaptationist programme, or the Panglossian paradigm ... This programme regards natural selection as so powerful and the constraints upon it so few that direct production of adaptation through its operation becomes the primary cause of nearly all organic form, function, and behaviour.
>
> Gould & Lewontin (1979, pp. 584–585)

> ... if it [the invocation of natural selection] could be dismissed after failing some explicit test, then alternatives would get their chance. Unfortunately, a common procedure among evolutionists does not allow such definable rejection for two reasons. First, the rejection of one adaptive story usually leads to its replacement by another, rather than to a suspicion that a different kind of explanation may be required. Since the range of adaptive stories is as wide as our minds are fertile, new stories can always be postulated. And if a story is not immediately available, one can always plead temporary ignorance and trust that it will be forthcoming ... Secondly, the criteria for acceptance of a story are so loose that many pass without proper confirmation. Often, evolutionists use *consistency* with natural selection as the sole criterion and consider their work done when they concoct a plausible story. But plausible stories can always be told.
>
> Gould & Lewontin (1979, pp. 587–588)

Following the adaptationist programme then, one could even, conceivably, immunise T_{hs} against the discovery of Maynard Smith's peculiar deep-sea fishes, whose luminous dots echoed stellar constellations. We might, for example, speculate on the existence of an ancestor which lived close to the surface, and camouflaged itself by imitating the night sky. This seems outrageous, but paleontological evidence would generally be far too fragmentary to definitively refute it; or, in the unlikely event that this *particular* story could actually be refuted, another modified version could then be generated which was immunised against the particular refutation. And so on.

Now it should be emphasized that Gould and Lewontin are not attacking T_{hs} as such. Indeed they explicitly accept the occurrence and importance of Darwinian selection ("... Darwin regarded selection as the most important of evolutionary mechanisms *(as do we)* ..."—Gould & Lewontin 1979, p. 589, emphasis added). Their primary objective is to deprecate a certain *style* of Darwinian theorizing—i.e. that which takes selection to be the necessary explanation of all attributes of all organisms. Nonetheless, whatever their motivation, their argument also serves to demonstrate that the corroboration of T_{hs} (separately from T_{cs}) is, at best, a somewhat fragile achievement.

I must repeat that I still consider T_{hs} to be definitely testable, or refutable, *in principle.* For example, we could conceive of conducting a controlled experiment in which the characteristics and life histories of all organisms within some delimited environment are recorded, so that the all episodes of S-lineage selection (if any) can be definitively established and recorded. It would then be possible that examples of the sustained, cumulative, growth of some inate organismic attributes could be observed which were *not* attributable to the (known) cumulative episodes of S-lineage selection. This would then represent a refutation of T_{hs}. Of course, to represent a severe test of T_{hs}, this experiment may have to be allowed run in an environment comparable to the entire planet Earth, and over a comparable (geological) timescale. Thus, although this is a "possible", severe, test in principle, it is entirely impractical.

The central difficulty in all of this is, of course, that the empirical phenomena with which T_{hs} deals are of such a scale in both time and space that they cannot generally be put into the form of practically repeatable experiments. As Popper incisively put it:

> ... we have to add that the phrase *in principle* is a very important restriction. Neither Darwin nor any Darwinian has so far given an actual causal explanation of the adaptive evolution of any single organism or any single organ. All that has been shown—and this is very much—is

> that such explanations might exist (that is to say, that they are logically possible).
>
> Popper (1961, p. 267)

In conclusion, T_{hs} has not been definitively refuted; but equally, it should not be considered to be highly corroborated (beyond what it inherits from the corroboration of T_{cs}) either. Nonetheless it seems to be better corroborated that any currently competing theory, and should be preferred (at least for the time being).

3.4 Organismic Darwinism (T_d)

Consider (finally) the following theory:

T_d (Organismic Darwinism): The biological world is, and has always been, a *D-system* (McMullin 1992, Section 7), with organisms functioning as D-actors. More particularly, S-creation in the biological world is a form of *unjustified* variation (does not involve anticipatory models predictive of the S-values of the resulting S-lineages).

T_d is, in effect, a still stronger form of T_{hs}. It adds the requirement that S-creation be unjustified, or unbiased, with respect to S-value. T_d claims not just that the biological world involves (and has always involved) the operation of S-lineage selection (with biological organisms as actors), but that it is a D-system proper. T_d thus entails T_{hs} but not *vice versa.*

It remains to consider the extent to which T_d can be tested, over and above the testing of T_{cs} and T_{hs}. The testing of T_d turns out to be rather difficult. We can certainly say that no counter example has been definitively demonstrated; but, as in the case of T_{hs}, given the infinitesimal scale of practical experiments compared to the reality of natural evolution, this cannot be ranked as especially strong corroboration.

We can do somewhat better by considering the *mechanisms* underlying the realisation of heritable characters (in modern organisms at least). It seems that the origin of new characters relies, in an essential way, on modification of certain more or less isolable biochemical components—generally DNA molecules. I shall term such modifications *mutations*, though my usage will be somewhat more generalised than is conventional.

The claim that S-creation is a form of unjustified variation is a version of the standard Darwinian hypothesis that the occurrence of new, heritable, characters is "random" (it must be stressed here that this is a very special usage of the word "random" which has little, if anything, to do with conventional probability theory or stochastic processes). It has been generally supposed that our modern understanding of the chemical basis of heredity definitively rules out any involvement of anticipatory models in the occurrence of mutations. Thus, we have, for example, Maynard Smith:

> It is sometimes said—usually by critics of Darwinism—that mutation is random. Now 'random' is a notoriously difficult word to define. I think that most scientists, when they speak of an event being random, mean that it would not be efficient to enquire into its causes, either because they think that the cause is in principle unknowable, as in quantum theory, or because it would be too much trouble to discover. In this sense, mutation is certainly not random. A lot is known about the causes of mutation—that is of changes in DNA. However, most geneticists do hold two things to be true of mutation. First, there is no restriction on the kinds of changes in sequence of DNA molecules that can arise by mutation, any more than there is a restriction on the sequence of letters that can be produced by a typewriter. Second, if a mutation is caused by a particular agent—for example X-rays—it is not in general true that the effect of the mutation will be to make the organism more resistant to the causative agent: in brief, mutations are not adaptive.
>
> Maynard Smith (1986, p. 40)

Similarly, and explicitly citing the notion of anticipation, Dawkins has expressed this idea as follows:

> There is randomness and randomness, and many people confuse different meanings of the word. There are, in truth, many respects in which mutation is not random. All I would insist on is that these respects do *not* include anything equivalent to anticipation of what would make life better for the animal.[12]
>
> Dawkins (1986, p. 306)

However, this straightforward view has recently been challenged. To explain this I must first elaborate slightly on the mechanisms whereby mutations

---

[12] Of course, I would say "S-lineage" rather than "animal" here.

can take place. It seems that these can be classified as follows:

1. Spontaneous, and accidental, modification arising from uncontrolled effects. This could include accidental errors during replication.

2. Natural, systematic, modification of the genetic material, typically, but not exclusively, in the course of sexual reproduction. This particularly refers to recombination (during meiosis) and the union of gametes in fertilisation (syngamy). However, various other systematic modifications can occur, even in organisms which do not exhibit a sexual cycle as such (Maynard Smith 1989, Chapter 10).

3. Artificial, systematic, modification of the genetic material. By this I essentially mean the modern practise of *genetic engineering*.

The first of these, by its (hypothesized) unsystematic nature, precludes any dependence on anticipatory models. The third is included for completeness, but is evidently irrelevant to the corroboration (or otherwise) of T_d: (human) genetic engineers have not been operative over the requisite geological time scales.

The second case is much more difficult.

Firstly, it has been suggested that, in a certain sense, the genetic material of biological organisms may constitute a complex ecology *in its own right.* This is the implication, for example, of recent theorizing in relation to so-called *selfish DNA* (Doolittle & Sapienza 1980; Orgel & Crick 1980).[13]

If this viewpoint is taken seriously then it seems to at least open up the possibility that fragments of DNA may incorporate, or have associated with them, “inate knowledge” (anticipatory models) of the effects of certain genetic processes, and that this would then affect the participation of the DNA fragments in such processes. This sounds like an acceptance that “mutation” could be informed by anticipatory models and would therefore not be random or unbiased in the sense required by T_d.

However, such an interpretation would be premature.

While I admit the possibility that some such DNA level anticipatory models might conceivably exist, the scope of such models may be limited to the effects on the DNA fragments themselves—and *not* on the organisms they may happen to be embedded within. If this were the case then the genetic modifications which occur, be they ever so systematic, would still be “unjustified” in the sense required by T_d; thus we must concentrate only on the latter question, of the anticipatory relationship (if any) between mechanisms for genetic modification and the S-value of the resulting (organismic) S-lineages.

Now the general notion that only certain evolutionary changes, or mutations, are *possible* (relative to a given starting point), and that this strongly constrains the overall path of evolution, is a fairly common theme. For example, consider the following, from Waddington, introducing what he calls evolutionary “archetypes”:

> You don’t just get a “horse archetype,” a “dipteran archetype,” but you get a “horse family archetype,” *with inbuilt characteristics of directions in which evolutionary change can easily go.*
>
> Waddington
> (1967, p. 115, emphasis added)

Maynard Smith has more recently considered somewhat similar ideas under the name “structuralism”—the view that “only certain kinds of structural change are possible to particular organisms” (Maynard Smith 1986, p. 40).

But the gist of these ideas is that, relative to some given starting point, only certain other points will be mutationally accessible. It is true that, on such a view, depending *only* on the distribution and characteristics of these, and their relative accessibility, it may well happen that, relative to certain points, mutations leading to higher S-value may be more “likely” to occur than those leading to lower S-value. But this would not, in itself, indicate the operation of any anticipatory model in the mutational process, and would not impact on T_d as stated.

However, various workers have taken the *further* step of suggesting, more or less explicitly, that organismic evolution is informed by anticipatory models. Thus we have the following from D.T. Campbell:

> Bisexuality, heterozygosity, and meiotic cell division represent a secondary invention increasing the efficiency of the process [of organismic evolution] through increasing the range of variation and the rate of readjustment to novel environments.
>
> Campbell
> (1960, pp. 381–382, emphasis added)

---

[13] Selfish DNA should not be confused with Dawkins’ (1976) “selfish genes”; see the discussion by Gould, where he argues *inter alia* that “the theories of selfish genes and selfish DNA could not be more different in the structures of explanation that nurture them” (Gould 1983, Essay 13, p. 174).

In the terms I have used here, this amounts to the suggestion that sex (the machinery of genetic recombination etc.) incorporates anticipation of environmental *variability* (on a time scale significantly longer than that of any single organism). This idea has been more recently independently formulated and elaborated by Stanley (1975).

However, recombination is not conventionally recognised as a form of "mutation" as such, so it might be considered to represent a case for special pleading. It would be more significant if anticipatory models were more generally implicated in genetic processes.

Lenat (1983) has explicitly argued, on very general grounds, for just such a position, when he claims that the genetic machinery may incorporate evolutionary "heuristics" (which I take to be equivalent to my anticipatory models) which facilitate the subsequent progress of evolution . Furthermore Boden has also repeated this suggestion, in particularly clearcut terms:

> . . . some strategy of "Plausible-Generate-and-Test" is needed, *whereby mutations of a type likely to be adaptive become increasingly probable. . .*
>
> Boden (1984, p. 312, emphasis added)

Most recently, this concept has been independently advocated by two separate biologists, Dawkins and Wills. The case of Dawkins is, perhaps, especially significant because, as already noted above, he explicitly rejected the notion of anticipation, as a factor in mutation, as recently as 1986. Since that however, he has speculated as follows:

> Is the world filled with animal groups which not only are successful, as individuals, at the business of living, but which are also successful in throwing up new lines for future evolution?
>
> Dawkins (1989a, p. 219)

I should say however that Dawkins is still not completely clear as to what he is proposing, and my interpretation (that he is now advocating the possibility of a significant anticipatory component in mutation) may not be correct.

Wills, in discussing what he describes as "The Wisdom of the Genes", is somewhat more explicit:

> The theme of this book is that there is an accumulated wisdom of the genes that actually makes them better at evolving (and sometimes makes them better at not evolving) than were the genes of our distant ancestors.
>
> Wills (1991, p. 6)

To repeat, what I consider important in these various proposals, from a variety of workers, is not the idea that the range of accessible mutations may be merely "limited" or "constrained", *but that mutations leading to increased S-value should be preferentially and systematically generated.* There is a subtle difference between these two concepts, but I suggest that this difference amounts to the idea that anticipatory models(s) (predicting something about the relative S-value of organismic S-lineages) are implicated in (at least some) genetic processes.

There is no contemporary consensus on this issue, and I cannot offer any definitive conclusion. However, for my specific purposes, I think the following observations are salient.

Firstly, all of the workers identified above, who have suggested some rôle for anticipatory models in genetic processes, do so on the assumption that such anticipatory models, if they exist, are *themselves* products of a previous process of organismic S-lineage selection. That is, while it is argued that S-value is sometimes increased by a directed process (rather than "random" or "unjustified" variation), it is accepted that there is always a *residue* of cases in which a strictly Darwinian ("unjustified") variation must be recognised as acting. Thus, there is no fundamental disagreement with T_d, but there is a critical question mark over its *scope*: roughly speaking, we must ask what are the relative contributions of "unjustified" and "anticipatory" variations to the phylogenetic growth of any particular attribute of interest?

As regards this question of the scope of T_d I note that the rôles suggested for anticipatory models in mutation are limited to extremely generalised predictions—such as that the environment will be "variable" (on a long timescale compared to single organisms) and therefore there should be a more or less continuous, but controlled, probing of new genetic variations (via recombination for example). There is no suggestion that there is any detailed modelling of the implications of specific genetic modifications for the phenotypic properties of the resulting organism. Indeed, given our current understanding of the extreme complexity of this relationship (i.e. of epigenetic development, or embryology), it seems that if anticipatory models capable of such predictions existed within organisms then they would be rather prominent, and could hardly escape detection. On the contrary, there is no indication of any subsystems, *anywhere* accessible to contemporary organisms, capable of predicting the specific effects of particular genetic modifications *for the organism* (or its offspring).[14]

On this basis then, I conclude that while the scope of T_d may no longer include *all* growth in S-value of organismic S-lineages, it is still the preferred theory for very many cases of great interest; and that even for those cases where it may not be *directly* applicable, it is still indirectly applicable as a theory of the formation of anticipatory models of mutation. In one form or the other then, T_d still stands as the "ultimate" theory of *all* sustained, cumulative, phylogenetic growth in organismic attributes.

3.5 Darwin's Solution

I have still not explained how T_d, or its precursors, T_{cs} and T_{hs}, actually solve P_d. Indeed, it is not even clear how they can *contribute* to a solution—for they contain no explicit mention of adaptive complexity.

The solution to this lies in the fact that P_d itself states that there has, in fact, been a sustained, cumulative, growth of *adaptive complexity* of biological organisms; given this, *and* given T_{hs}, we can say that this growth of adaptive complexity must have occurred by virtue of its being positively correlated with S-value. This goes at least some way to solving P_d; but it is still not an adequate solution. At best, T_{hs} yields only an explanation for the *retention* of (some) increases in adaptive complexity, once given that they occur—it does not offer any explanation for how such increases may come about in the first place;[15] but, with the additional claim, embodied in T_d, that S-creation is a process of *unjustified* variation, we now achieve something which, if accepted, might be considered a more or less *adequate* solution of P_d, for it finally addresses the source of increased adaptive complexity. Of course, it still does not answer this question completely for it does not describe the process of the origination of variation in any detail; but by stipulating that the process is unbiased or unjustified, it removes the need for any sentient or conscious "designer". To the extent that the objective is to provide an alternative to Paley's theistic solution to the problem, T_d is now adequate.

It is crucial to note here that T_d does *not* entail that adaptive complexity be *always* correlated with S-value, or (therefore) that adaptive complexity should *necessarily* grow. Maynard Smith makes the point thus:

> There is nothing in neo-Darwinism which enables us to predict a long-term increase in complexity. All one can say is that since the first living organisms were presumably very simple, then if any large change in complexity has occurred in any evolutionary lineage, it must have been in the direction of increasing complexity; as Thomas Hood might have said, 'Nowhere to go but up'. But why should there have been any striking change in complexity? It is conceivable that the first living thing, although simple, was more complex than was strictly necessary to survive in the primitive soup, and that evolution of greater fitness meant the evolution of still simpler forms.
>
> Maynard Smith (1969, pp. 88–89)

That is, T_d solves P_d only in the special sense that it *permits* a spontaneous growth of adaptive complexity; it does not compel or predict such growth. As discussed previously (McMullin 1992, Section 9), I adopt the position that this is the best that can be achieved.

Finally, I should add that an attempt to realise the growth of adaptive complexity, via (genuinely) Darwinian processes, in *artificial* systems, could, in itself, represent a further kind of test of T_d. While such tests could not provide a strong *refutation* of T_d (due to the necessarily limited scale of artificial systems compared to biological evolution) the *successful* demonstration of significant, spontaneous, growth in adaptive complexity in artificial (D-)systems, were it to be achieved, would still represent a significant *corroboration* for T_d.

4 Competitors

> . . . I do not think that Darwinian natural selection is the only thing we need to understand to understand evolution. However, I do think that Darwin's theory is correct, and that it is the only adequate explanation for what is for me the most characteristic feature of living organisms. This feature is the way in which their structure and behaviour adapts them to survive and reproduce in a specific environment.
>
> Maynard Smith (1986, p. 40)

---

[14] I am, of course, continuing to discount the exceptional case currently presented by the genetic engineers of species *homo sapiens*.

[15] This, precisely, is the sense of the following remark, originally attributed to Samuel Butler: 'To me it seems that the "Origin of Variation", whatever it is, is the only true "Origin of Species".' (quoted by Fisher 1958, p. 1).

I agree with the sentiments expressed here by Maynard Smith. The application of Darwinism to the biological world, embodied specifically by T_d, is surely not a complete or final solution to P_d—in fact, I propose to consider some detailed, technical, difficulties with it in a future work (McMullin Forthcoming). Nonetheless, I consider that there is currently no significant *competitor* to it. To establish this, I must consider a number of supposed competitors, and identify their deficiencies relative to T_d.

In general, any competitor of T_d must hold that the biological world *fails* to be a D-system (or has so failed in the past) in some significant, specified, respects. The comparison between T_d and each competitor must therefore focus on the particular respects in which this failure is claimed.

4.1 Saltationism (Revisited)

I consider first (because I think it a very special case) the idea of *saltationism.* As already introduced in section 2.2.2, saltation (properly so-called) involves "large" evolutionary changes (up to and including the establishment of new species), which may involve "large" increases in adaptive complexity, occurring in a single generation (which is to say, a single episode of S-creation).

In section 2.2.2 I argued only that saltationism is not an alternative or competitor to T_t. The question which I now wish to address is whether, *within* T_t, it might be a competitor or alternative to T_d.

The point of saltationism (relative to P_d) is, roughly speaking, that if such "large" increases in adaptive complexity occur in a single generation, then the "truly" creative aspect of evolution *must* lie in these changes, with selection playing, at best, a secondary rôle of weeding out grossly deformed organisms.

Darwin was adamant that the existence of saltations would be utterly incompatible with his theory: "If it could be demonstrated that any complex organ existed which could not possibly have been formed by numerous, successive, slight modifications, my theory would absolutely break down" (Darwin 1859, Chapter VI, p. 219). Darwin here presents, in opposition to saltationism, what I have termed *gradualism*: the doctrine that "large" evolutionary changes (or, at least, "large" increases in adaptive complexity) can only occur by accumulation of a "large" number of individually "small" changes, each being retained via natural selection (which requires some significant number of generations for each such episode).

In some sense then, Darwin considered gradualism to be an inherent part of T_d, and saltationism to therefore represent a significant competitor to T_d. Popper seems (at one time at least) to have shared this view of Darwin's: "Gradualness is thus, from a logical point of view, the central prediction of the theory. (It seems to me that it is its only prediction.)" (Popper 1974a, pp. 137–138).

However, this view was opposed by Thomas Henry Huxley even on the very eve of publication of *The Origin*, when he wrote to Darwin that: "You have loaded yourself with an unnecessary difficulty in adopting *Natura non facit saltum* so unreservedly" (quoted by Gould 1980, Essay 17).

I shall here adopt Huxley's position precisely. Saltationism is not a competitor of Darwinism—or, at least, not necessarily so; rather, saltationism and gradualism are competing *elaborations* of T_d. T_d says nothing about the "size" of the variations which have been the subject of S-lineage selection; and, *pace* Popper, I do not see that any "prediction" relating to this can be wrung from it.[16] In fact, as noted in section 3.1, I do not accept that adaptive complexity (whatever about S-value) admits of a metric which would allow "size" to be meaningfully introduced (at the level of T_d) at all.

I hold that this is a strictly correct statement of the logical status of saltationism (and gradualism) relative to Darwinism. However, lest there be any confusion, I should stress that I am not commenting (*either* way) on the relative merits of saltationism versus gradualism *qua* elaborations (rather than competitors) of T_d. In fact, any conclusion on this requires a definite commitment to some particular interpretation of "large" versus "small" variations. A comprehensive discussion would take me too far afield here, but I should like to consider, in detail, just one particular approach to this question; this may allow me to isolate just what Darwin was "really" concerned with, in his stipulation of gradualist evolution.

Fisher (1958, pp. 41–44) has formulated a detailed, in principle, argument *against* a particular version of saltationism (though Fisher himself does not use this word). I shall first paraphrase and elaborate his argument, though using my own terminology.

Suppose that S-class can be identified with a point in some Euclidean n-space, where n is "large". Roughly speaking, we suppose that some set of po-

---

[16] I should note however that Popper's formulation of Darwinian theory, on which he bases his prediction of gradualness (Popper 1974a, pp. 135–136), is not quite identical with my T_d.

tential organisms (or, more properly, potential S-lineages) of interest (i.e. which can potentially compete with, or selectively eliminate, each other) can "match" their environment more or less well in each of n distinct traits or characters. Any given S-lineage can be represented by a point in this space. A D-lineage ("species" in Fisher's terms) would be represented by a set of such points—a somewhat amorphous "cloud" in the S-class space. With each S-creation event a new point appears; every time an S-lineage is selectively eliminated (or dies for any other reason) a point disappears. Thus, over time, the set of points making up the D-lineage could "move" through the space.

Now, the environment itself will also be represented by a point in this space—in the sense that some particular S-lineage, or some particular set of characteristics, would be "best matched" to the environment. Fisher identifies the distance between any other point in the space, and this "optimal" point as a measure of the "degree of adaptation" (strictly one should say the reciprocal of this distance is a measure of degree of adaptation). Thus, to the extent that the points making up a D-lineage lie "close" to this optimal point, then the D-lineage is said to be "well adapted".

While Fisher does not expressly say so (as far as I can see) it seems that we must assume that, to the extent that "adaptation" in this sense occurs by Darwinian selection, the degree of adaptation must be well correlated with S-value. That is, if we imagine adding another dimension to the space to represent S-value, and thus establish a "surface" of S-value, it follows that the optimum point will be at the (global) peak of this surface. Again, it seems implicit in Fisher's discussion that this must be the *only* peak in the surface, in general (though I shall return to critique this idea below). That being the case, the evolution of a D-lineage can be envisaged as a process of *hill-climbing*; S-creation events result in new points being tested; where any of these are "uphill" relative to the existing points they will selectively displace old points etc. Evolution would be represented by a trajectory whereby a point (or set of points) representing the S-lineage(s) of a D-lineage approaches more or less closely to this peak (and, indeed, could track it if it moves slowly enough—which is to say the environment changes slowly enough).[17]

Fisher supposes that S-creation (mutation) might be modelled, at least approximately, in this space by the appearance of a new point (S-lineage) which is displaced from some existing point (corresponding to the parent(s)) in a random direction. The word "random" *is* used here in its technical, stochastic, sense: i.e. the direction of displacement is a random variable uniformly distributed over all possible directions. The interpretation of "possible directions" depends, of course, on the dimensionality of the space (n). The size of the displacement may also vary, of course, presumably in some "random" manner. For our purposes we need not worry about the details of this. For the moment, we can consider just the set of all possible displacements of a "fixed" size r; subsequently, we shall go on to consider the expected effects of different such sets (i.e. having different values of r).

Let d denote the diameter of the circle centered on the adaptive peak, and passing through the "parental" point (S-lineage); that is, $d/2$ is the distance from this point to the peak (and d^{-1} is a formal measure of "degree of adaptation"). r denotes the (fixed) size of a randomly directed displacement from this point, identifying a new point (S-lineage) resulting from an S-creation event.

It is now (trivially?) clear that, $\forall\ r > d$, and regardless of the dimensionality of the space (n), the probability of the new point being closer to the peak must be zero (even if it were in the "perfect" direction—pointed straight at the peak—it would still overshoot by more that the original distance, $d/2$, on the other side). To put this another way, evolutionary improvements in adaptation can occur *only* through displacements $r < d$. If we define a "saltation" as any displacement for which $r > d$ (and acknowledging that this definition will quantitatively vary with the "instantaneous" degree of adaptation) then Fisher's result (at least in a first version) is that, in this particular sense, Darwinian evolution cannot proceed by saltations.

In fact, Fisher's result is significantly stronger than this. While it is true that, regardless of n, the probability of improvement (in Fisher's model) is zero $\forall\ r > d$, this does not, in itself, say what happens for $r < d$. It turns out that this is moderately complicated, and depends crucially on the dimensionality of the space. Informally, the greater the dimensionality, the more extra ways there are of getting a displacement in the "wrong" direction, and therefore the lower the probability of getting an

---

[17] It seems clear from Fisher's description of his space in terms of a *match* along each dimension to the organism's (or better, S-lineage's) environment, that it is, in effect, a space of *phenotypic* characteristics. The qualitative characteristics of this space are also just those described by Lewontin (1978); but Lewontin identifies his space as being one of *gene frequencies,* rather than phenotypic characteristics, which would be a different kind of thing entirely. Provine (1986, Chapter 9) provides a much more detailed discussion of these issues.

improvement. Fisher's quantitative version of this is that, for "large" n, the probability of improvement approaches zero for all r much greater than $d/\sqrt{n}$. So we can now refine and strengthen the overall result by defining a "saltation" as any displacement for which $r \gtrsim d/\sqrt{n}$. In other words, improvements in adaptation can feasibly come about *only* through displacements $r \lesssim d/\sqrt{n}$, which, for "large" n, means $r \ll d$. And the greater the number of dimensions, the stronger this result becomes.[18]

This argument seems quite convincing—*in its own terms*; but the question is whether the situation envisaged by Fisher is a good model of S-creation.

A first problem is the (implicit) assumption that, in the relevant space, the S-value surface has a single peak or optimum. This seems to be quite unjustified; and of course, if there are multiple peaks then "large" displacements in the space ("saltation") may well allow a D-lineage to move to a separate (and "better") peak. Indeed, this may be actually *necessary* for continued evolution (to "escape" from local optima). This point seems to have been a central aspect of the disagreement between Wright and Fisher about the general nature of evolutionary change; Wright has long argued, in his *shifting balance* theory, that what we might call "peak jumping" of some sort (albeit not necessarily by single mutational events) is, at the very least, an important evolutionary process (Wright 1982). Turner (1985, pp. 172–173) has also considered this issue of multiple adaptive peaks, including a reconsideration of an analogy cited by both Fisher and Dawkins (see Appendix A below), and concludes that Fisher's analysis, as an argument against the *possibility* of saltationism, simply does not work (although he does hold that saltations cannot suffice to explain *all* increases in adaptive complexity).

But quite aside from this question as to the existence of multiple "adaptive peaks", there is an even deeper problem here. We must ask whether S-class should be modelled by a point in a "fixed" Euclidean n-space (discrete or continuous) *at all*. It might seem that, in modern terms, this could surely be achieved by defining S-class directly on the basis of DNA sequence (compare, in particular, the "genetic (hyper-)space" of Dawkins 1986, p. 73). There are various possible problems with this—for example there is no way to represent the evolution of the "genetic system" itself in such a scheme. Furthermore, as noted in section 3.4, it is now considered that there may be significant systematic aspects to the modification of DNA sequence in mutation. But more seriously, it is clear that no *single* notion of S-value would be applicable throughout such a space (see also the more detailed discussion in McMullin 1992, Section 8).

In any case, to conclude this discussion of saltationism: I take the view that, as long as S-creation (saltationary or otherwise) is *not* taken to be informed by anticipatory models (and this is certainly the case for the version considered by Fisher) then the question of what "size" of variations will be evolutionarily significant may be an interesting one—but it lies entirely *within* the scope of T_d. Saltationism, in this sense, is not a competitor to T_d. If, on the other hand, one means to *define* "saltations" as variations whose "size" is such that they could not "plausibly" arise by unjustified variation (i.e. without the involvement of some kind of anticipatory model or knowledge) then saltationism does indeed become a competitor to T_d. This *seems* to be the essence (though not the detail) of Fisher's argument. But, in that case, saltationism becomes equivalent to what I call *mutationism*—and I will discuss *that*, its own right, in section 4.3 below.

4.2 Theism *et al.*

A valid competitor of T_d is, of course, Paley's own theory—that the growth of adaptive complexity is due to the intervention of some supernatural agent, of more or less unspecified nature. While Paley probably thought in terms of special creation, the exact nature of the intervention is not important here—it could be by way of creation, biasing of selection, or biasing of variation, or any combination of these.

Paley's argument, briefly, was that the mere existence of adaptive complexity *demanded* the existence of a "creator" in some sense. Further, since this argument can be applied recursively, it demands an *immortal* creator, which is to say *God* (by any other name).

The crux of the argument (and the point of the "watchmaker" analogy) is the claim that, in effect, adaptive complexity *cannot* spontaneously grow, except by the intervention of an intelligent and/or conscious agent. The crux of T_d is, of course, the converse claim, precisely that adaptive complexity *can* grow spontaneously.

---

[18] I note that Dawkins has also recently presented an alternative version of this argument of Fisher's (Dawkins 1986, pp. 230–236); however, Dawkins' discussion is limited to the weaker version of Fisher's result (he does not mention the effect or significance of the dimensionality of the space), and also harbours a minor technical error which is potentially confusing (at least, it confused *me*). This is not a substantive point, but, for completeness, I include a more detailed review in Appendix A.

Now it has been admitted that there are serious limits to the extent that T_d can be severely tested. Nonetheless, it is the case that T_d has effectively withstood such tests as have been attempted. Furthermore, the *corroborability* of T_d is, of course, greater than Paley's theism; in fact, it is greater than *any* competitor which fails to stipulate a definite mechanism for the growth of adaptive complexity—i.e. which fails to make testable predictions. No matter what evolutionary, or even non-evolutionary, pattern were observed in the biological world, it could not refute a generalised theistic, or metaphysical, explanation. Thus, T_d can and should be rationally preferred to such theories.

I may say that clinging to a more or less theistic solution of P_d, merely on the basis that T_d is too "difficult" to believe (or perhaps I should say simply too *unpalatable*) is no substitute for a serious attempt to formulate and carry out severe tests of T_d. Dawkins has rather caustically characterised the former approach as mere "argument from personal incredulity" (Dawkins 1986, p. 38).

Finally, I should stress that, while the (hypothesized) existence of God cannot seriously compete with T_d (in the current state of testing) as a solution of P_d, the acceptance of T_d in its turn does not, in any way, preclude a general belief in the existence of some God; it merely, tentatively, precludes Her involvement in the growth of adaptive complexity…

4.3 Mutationism

> It is hard for us to comprehend now but, in the early years of this century when the phenomenon of mutation was first named, it was regarded not as a necessary part of Darwinian theory but as an *alternative* theory of evolution!
>
> Dawkins (1986, p. 305)

Saltationism, as I have described it, was concerned only with the "sizes" of mutation which are significant in the evolutionary process—it accepts or presupposes that these changes are unjustified or unbiased with respect to S-value. *Mutationism*, on the other hand, is not particularly, or necessarily, concerned with the relative sizes of mutations, but rather holds that mutations are significantly *biased* in some sense. While the nature of this bias need not be fixed, or universal, if mutationism is to address P_d we must suppose that, in some evolutionary lineages at least, there has been a persistent bias toward increased adaptive complexity. Given that such biases exist, we may conclude that major features of evolution, particularly including growth in adaptive complexity, are primarily (if not exclusively) due to the directional bias of mutation, and not to natural selection.[19]

Unlike saltationism, mutationism is a genuine competitor to T_d—it offers a different, rather than a more detailed, solution to P_d. Unlike theism, it is not metaphysical; while it may be vague about the mechanisms for mutational biases, it stipulates that such biases exist—especially a bias toward increased adaptive complexity. Such biases should therefore manifest themselves if we look for them under natural and/or experimental conditions. Indeed, a significant factor in the original impetus of the mutationist theory may be traced to de Vries' interpretation of certain spontaneous, inheritable, variations he observed in wild plants; de Vries interpreted these events (wrongly as it subsequently transpired) as the spontaneous formation of new species (Hardy 1965, p. 82; Wills 1991, pp. 112–113). Indeed, it was de Vries who coined the term *mutations* for such phenomena. Similarly, the (supposed) phenomenon of orthogenesis, discussed in section 3.3 above, appeared to support a mutationist view of evolution.

The original mutationist school did not address the question of how the required mutational biases could come about. Dawkins describes the situation as follows:

> The problem with mutation as the sole evolutionary force is simply stated: how on earth is mutation supposed to *'know'* what will be good for the animal and what will not? Of all possible changes that might occur to an existing complex mechanism like an organ, the vast majority will make it worse. Only a tiny minority of changes will make it better. Anybody who wants to argue that mutation, without selection, is the driving force of evolution, must explain how it comes about that mutations tend to be for the better. By what mysterious, built-in wisdom does the body choose to mutate in the direction of getting better, rather than worse? … The mutationists, needless to say, never answered it. The odd thing is that the question hardly seems to have occurred to them.
>
> Dawkins (1986, pp. 305–306)

---

[19] Assuming that the *operation* of selection is still accepted (T_{hs}), then, to make this argument fully coherent, we must further suppose that, among the spectrum of mutants of increased adaptive complexity, there are consistently some which *also* represent increased S-value.

However, Dawkins is perhaps benefiting somewhat from hindsight here: as long as there seemed to exist empirical phenomena (spontaneous speciation in a single generation, orthogenetic trends) which supported mutationism against T_d, it was not necessary to stipulate a mechanism for directed mutation; obviously the elucidation of such mechanisms would be a significant element of any mutationist research programme—but it would not be *un*reasonable to defer this until, say, mutationism was accepted over T_d.

In any case, this original mutationist school withered away, as the empirical basis for it was found to be flawed, and it became clear that the particulate view of inheritance (implied by mutationism), which had initially been supposed to be incompatible with T_d, was, in fact, complementary (or even essential) to it. In the terms I am using here, tests which were thought to be decisive between mutationism and T_d were not, in fact, decisive at all, for the evidence was actually compatible with both theories; and since the information content, or corroborability, of T_d was greater (since it does not postulate any particular biases, of unspecified origin, in the mutational process) it was clearly preferable.

Subsequently, our increased knowledge of genetic processes at the level of molecular biology seemed to decisively rule out the existence of any mechanism for the persistently biased mutations required by mutationism, and the theory was considered to be more or less definitively refuted.

However, as discussed in section 3.4, the notion of directed mutation has recently been resurrected, in the form of Wills' "wisdom of the genes" or Dawkins' "evolution of evolvability". While these recent authors do not refer to themselves as mutationists (perhaps wisely) the logical structure of their theories is at least similar to the original mutationists. Nonetheless, there are two decisive differences in this *neo*-mutationism:

- Directed mutation is *not* considered to be responsible for detailed evolutionary changes, typical of increased adaptive complexity—because that would require the existence of an anticipatory model capable of predicting the detailed phenotypic effects of given genetic modifications, and there is no evidence for the existence of such models.

- In any case, insofar as mutation is considered to be directed at all, the mechanism for this direction is itself considered to be the outcome of earlier Darwinian selection, in the sense of T_d.

In conclusion: mutationism, in its original form, as a competitor of T_d (in solving P_d) has been effectively refuted; and in its modern form it is no longer a competitor to T_d, but more like a possible elaboration (and even then having strictly limited scope).

4.4 Lamarckism

> The Lamarckian theory seems to have great emotional appeal, for certain types of intellectual as well as for laymen. I was once approached by a colleague, a celebrated Marxist historian and a most cultivated and well-read man. He understood, he said, that the facts all seemed to be against the Lamarckian theory, but was there really no hope that it might be true? I told him that in my opinion there was none, and he accepted this with sincere regret, saying that for ideological reasons he had wanted Lamarckism to be true. It seemed to offer such positive hopes for the betterment of humanity.
>
> Dawkins (1986, p. 291)

The term *Lamarckism* has been used to denote a wide variety of evolutionary theories, not all of which are fairly attributable to the original Chevalier de Lamarck. Lamarck[20] appears to have accepted evolutionary descent or transmutation of species (T_t—see, for example, Dawkins 1986, p. 289), and thus was attempting to confront more or less the same problem as Darwin, the growth of adaptive complexity (P_d). However, his attempt long predated Darwin's work, and as a result is not easily compared to it.

Lamarck's central idea, to explain the evolution of adaptive complexity, seems to have involved ascribing evolutionary force to the wishes, or perceived needs, of organisms themselves (Wills 1991, p. 67; Dawkins 1986, p. 289). Roughly speaking the theory goes like this:

1. Organisms have "desires", or "aims", or "objectives" (these need not be conscious).

2. Among these is a generalised desire for "improvement" or "perfection".

---

[20] According to Wills (1991, pp. 65–66), the title *Chevalier de Lamarck* was a somewhat dubious affectation (Lamarck's "real" name being Jean-Baptiste Antoine de Monet). However, convention now decrees that he is known simply as Lamarck, and I shall conform to this.

3. In response to these desires, organisms are, in fact, successful (to a greater or lesser extent) in improving themselves during their lifetimes. The effect of "use and disuse" on morphological characters is commonly cited; examples include the development of muscles through exercise, darkening of skin exposed to sunlight etc. However, the theory does not hinge on use and disuse as a *sole* mechanism for improvement; it suffices that improvement is possible, howsoever it occurs.

4. *The improvements which organisms actually achieve during their lifetimes are (preferentially) transmitted to their offspring*—i.e. organisms can somehow control which of their characteristics are inherited and which are not. This is the celebrated "inheritance of acquired characteristics"—though the latter phrase is potentially misleading, as I shall discuss below.

5. In this way, an evolutionary lineage can show progressive improvement of its members as a cumulative result of their individual, incremental, success in improving themselves.

I may say that, in this form, Lamarckism seems eminently reasonable. In particular, the truth of points 1, 2, and 3 seems quite clear; 4 requires support, but is not *inherently* implausible; and 5 (which is to say the solution of P_d) is then simply the logical consequence of what has gone before. Furthermore, in ascribing a central rôle to the efforts of organisms themselves toward their self-improvement, it seems positively attractive (as compared to T_d) in restoring some kind of purpose to (human) life. It is perhaps not surprising, therefore, that Lamarckism continues to be a kind of holy grail of evolutionary theorizing, pursued despite the continuing privations of the quest, as reflected in Dawkins' earlier anecdote.

The modern genetic theory of inheritance was not, of course, available to Lamarck. But his theory can be described in modern terms as a particular elaboration of mutationism. What Lamarckism adds to mutationism is a specific mechanism for distinguishing between "good" mutations and "bad" mutations: good mutations are precisely those which transmit characters that an organism has tried and found beneficial in its own lifetime. This is, at first sight at least, rather a good proposal, and, if mutationism were successful as a *general* theory, then this specific elaboration would certainly bear careful consideration.

Unfortunately, as previously discussed, mutationism itself, as a general theory of evolution, has been found severely wanting. This is primarily because it seems to rely on the existence of anticipatory models of the (phenotypic) effects of specific genetic transformations. Such models do not seem to exist (except in the most rudimentary and limited sense); and given the complexity of the mapping from genotype to phenotype, it seems unlikely that any detailed models of this sort could feasibly be constructed.

This argument is equally, if not more effective against Lamarckism; for Lamarckism demands not only *some* anticipatory model (of embryology), but a model that operates in the direction from phenotype to genotype. This relationship is, if anything, vastly more complex that the mapping from genotype to phenotype. Dawkins has dubbed the virtual impossibility of carrying out a mapping from phenotype back to genotype the *Central Dogma of Embryology* (Dawkins 1982, p. 174). This was apparently inspired by the Central Dogma of *Molecular Genetics*. The latter states that DNA sequence may be translated into protein sequence, but not *vice versa.* While this seems to hold in practise, there is no deep reason why it should not be contravened: the coding from DNA to protein is relatively straightforward and easily reversible in principle. The Central Dogma of Embryology, on the other hand, is much stronger—for the mapping in question is, to say the least, extremely complex. These two "Central Dogmas" should therefore be kept clearly distinct, and not confused with each other. The significance of the distinction is the following: the Central Dogma of Molecular Genetics precludes Lamarckian inheritance—but only very weakly so. If this were the only barrier, then we could easily envisage that Lamarckian inheritance, even if it now no longer occurs, could well have played a significant rôle in the past. But the Central Dogma of Embryology seems to rule out any rôle for Lamarckian inheritance for as long as life has employed an embryology remotely as complex as now exists on Earth. This is a much stronger refutation.

It should be noted that while Lamarckian evolution entails the existence of Lamarckian inheritance—the inheritance of "acquired" characteristics—the converse does not hold. Lamarckian *inheritance* is not synonymous with Lamarckian *evolution* (or Lamarck-*ism*, as I am using the term). That is, even if it should one day be discovered that at least some "acquired" characteristics could be translated back into genetic modifications which would preserve them indefinitely, this would still not *entail* Lamarckian evolution. For the latter, we require that "favourable" (acquired) characteristics be *preferentially* transmitted. This is a more important distinction than may at first appear, as we shall see.

The refutation of Lamarckism, based on the Central Dogma of Embryology, is logically compelling—but it is still vaguely unsatisfactory. One is still left with the feeling that, even if T_d is the only effective solution to the growth of adaptive complexity in the (terrestrial) biological world, Lamarckism still offers a better solution "in principle". Thus, we might suppose that Lamarckian evolution may well occur on some other planet—or *could be stimulated in an artificial evolutionary system.* The attraction of this is that it seems inherently much more "efficient" than Darwinian evolution—the experiences gained by an organism in its lifetime need not be disgarded when the organism dies, but can be retained, accumulated, and capitalised on. If this reasoning is correct, then it would seem that, despite the failure of Lamarckism as a *biological* theory, it might still offer a short cut to success in artificial evolution. This possibility needs, therefore, to be carefully analysed.

The proposal then is as follows. We assume that we have some natural or artificial system, which exhibits sustained growth in adaptive complexity. We suppose (for the sake of argument) that Lamarckian evolution, in the terms described above, *does* occur in this system. We then ask whether this could be a satisfactory explanation of the observed growth in adaptive complexity.

Lamarckian evolution implies two things—an organism must develop new or modified characteristics (not possessed by its parent(s)), in somatic time, and it must selectively transmit those found to be "favourable" to its offspring. Two distinct, but closely related problems arise:

1. How *can* an organism develop "favourable", "new" characteristics relative to its parent(s)?

2. How can an organism (consciously or otherwise) distinguish "favourable" from "unfavourable" acquired characteristics?

In many ways we have just recast the original problem, but now localised in the lifetime of a single organism. That is, we must either assume that there is a bias in the generation of new characteristics (so that "favourable" ones are preferentially generated), or we must assume some mechanism for "selecting" favourable from unfavourable ones (or some combination of these two).

Let us suppose there *is* a favourable bias in the generation of new characteristics. This is just the original mutationism dressed up slightly. It is made possible, or plausible, by the assumption that Lamarckian *inheritance* is possible. But it begs the question of the source of the bias. The only possible answer seems to be the use of effective anticipatory models, to anticipate that certain modifications will be advantageous before they are actually carried out. It is Boden's "Plausible-Generate-and-Test" again (section 3.4 above—see Boden 1984). It could well be a significant evolutionary process: but we should then have to ask where the anticipatory models which support it came from in the first place. Ultimately *(if Popper's position on induction is accepted)* this must lead back to a reliance on *unjustified* variation again.

Well then, let us restrict attention to unjustified variations generated in the lifetime of the organism. We can still eschew T_d provided that the organism itself can select or distinguish between favourable and unfavourable variations (i.e. the organism(s) carry out the selection of variations which are transmitted and retained, rather than "natural" selection). But this then begs the question of the origin for the selection *criteria* used by an organism. Related questions are, for example, how does an organism (or perhaps we should say "lineage" in this Lamarckian context) come to classify certain stimuli as "painful", or other stimuli as "rewarding" etc.? This line of argument seems to lead inevitably back to natural, S-lineage, selection as the *ultimate* evolutionary force.

The conclusion from all this is that Lamarckian evolution, if it were practical (i.e. if the Central Dogma of Embryology could be somehow rescinded), might well be a very significant and important evolutionary process; but it could never, *in principle* be a complete solution to the problem of the growth of adaptive complexity. The most we can say seems to be that Lamarckism might be a useful extension or enhancement of T_d: but it could never compete with or replace it—it simply does not solve the problem at hand. Dawkins has previously made much the same arguments, concluding that:

> ... the Lamarckian theory can explain adaptive improvement in evolution only by, as it were, riding on the back of the Darwinian theory. Given that Darwinian selection is there in the background, to ensure that some acquired characteristics are advantageous, and to provide a mechanism for discriminating the advantageous from the disadvantageous acquisitions, the inheritance of acquired characteristics might, conceivably, lead to some evolutionary improvement. But the *improvement*, such as it is, is all due to the Darwinian underpin-

> ning. We are forced back to Darwinism to explain the adaptive aspect of evolution.
>
> Dawkins (1986, p. 300)

While I consider that these arguments do, finally and decisively, eliminate Lamarckism as a *competitor* of Darwinism, there is yet an epilogue. I have perhaps been slightly disingenuous in this discussion. Roughly, the argument was that Lamarckian evolution is not *practical*, and even if it were it would be *inadequate*. I hold to this conclusion; but it can easily be read as "Lamarckian evolution is not *possible,* and even if it were it would be *insignificant*". I do not intend these much stronger claims.

In fact, it can be argued that a form of Lamarckian evolution is not only possible, but well known, and has been crucially significant in the evolution of at least one species. This is so-called *cultural* evolution, particularly in the human species. Humans acquire cultural characteristics ("knowledge") selected by their parents (among others) from their own accumulated wisdom. Humans, in turn, generate new characteristics ("knowledge") not possessed or transmitted from their parents. This is Lamarckian evolution *par excellence*, and is undoubtedly a "significant" mechanism for the evolution of the unique characteristics of this extraordinary animal.

I accept this, in a limited sense which I shall detail below; but emphasize again that it cannot be considered a *solution* to the problem of the growth of adaptive complexity in general, and the growth of human knowledge in particular. This is worth clarifying because at least one biologist has implied not only that Lamarckian evolution is a *significant* mechanism in human, cultural, evolution (which I can accept) but that it is in some sense an *explanation* of it. The claim is made by Gould, and seems worth quoting at length:

> *Homo sapiens* arose at least 50,000 years ago, and we have not a shred of evidence for any genetic improvement since then. I suspect that the average Cro-Magnon, properly trained, could have handled computers with the best of us (for what its worth, they had slightly larger brains than we do). All that we have accomplished, for better or for worse, is a result of cultural evolution. And we have done it at rates unmatched by orders of magnitude in all the previous history of life. Geologists cannot measure a few hundred or a few thousand years in the context of our planet's history. Yet, in this millimicrosecond, we have transformed the surface of our planet through the influence of one unaltered biological invention—self-consciousness. From perhaps one hundred thousand people with axes to more than four billion with bombs, rocket ships, cities, televisions, and computers—and all without substantial genetic change.
>
> Cultural evolution has progressed at rates that Darwinian processes cannot begin to approach. Darwinian evolution continues in *Homo sapiens*, but at rates so slow that it no longer has much impact on our history. *This crux in the earth's history has been reached because Lamarckian processes have finally been unleashed upon it.* Human cultural evolution, in strong opposition to our biological history, is Lamarckian in character. What we learn in one generation, we transmit directly by teaching and writing. Acquired characters are inherited in technology and culture. *Lamarckian evolution is rapid and accumulative.* It *explains* the cardinal difference between our past, purely biological mode of change, and our current, maddening acceleration toward something new and liberating—or toward the abyss.
>
> Gould (1980, pp. 70–71, emphasis added)

There is certainly an underlying terminological difference here between myself and Gould. In this context he is restricting "Darwinism" to mean what I have called "Organismic Darwinism". So part of his point is, presumably, that cultural evolution is different because it does not rely on the (organismic) actors of (organismic) Darwinism. I quite agree with this—I freely admit that cultural evolution involves the emergence of a new kind of *Darwinian* actor (the linguistically formulated hypothesis) and new selective forces or environments (preeminently, the environment of *science* which seeks to select "true" theories—or, at least, to reject "false" ones). But it does seem that Gould intends something much stronger than this—not just a change in the actors, but a change in the evolutionary process or mechanism. A change, in fact, away from anything that might be called Darwinism, even by analogy. I reject entirely this implication that *Lamarckian* processes impart to cultural evolution its crucial distinguishing features. Rather I argue that the crucial processes, which *might* constitute an explanation of how cultural evolution can show a growth of adaptive complexity (human knowledge), are still essentially Darwinian rather than Lamarckian.

In more detail, this Darwinian underpinning to cultural evolution is manifested in two distinct ways:

- The organismic pre-requisites (whatever they are) to support cultural evolution must first have evolved by conventional organismic Darwinism (though there may have been some interaction between these two processes at a late stage).
- Cultural evolution itself relies, ultimately, on unjustified variation and selection. The variations are not recorded in organismic genetic material, but rather in books, pictures, computer files etc. Selection is (typically) carried out by argument and/or testing rather than via relative S-value of competing organismic S-lineages. But the *process* is still an essentially Darwinian one, even if the actors are no longer sensibly regarded as organisms.

This kind of view has been advocated (though in more circumscribed form) by Dawkins in his theory of the (Darwinian) evolution of *memes* (Dawkins 1989b, Chapter 11). In its strongest form, the picture I have presented here is just the evolutionary epistemology of Campbell (1974) or the Popperian growth of knowledge by *conjectures and refutations* (Popper 1963).

5 On the Scientific Status of Organismic Darwinism

I have presented and analysed Organismic Darwinism as an essentially straightforward theory—whose scientific status is not problematic. Indeed, this was a crucial, if implicit, assumption in my comparison between T_d and metaphysical competitors such as Paley's theism. I stand over this claim for the scientific status of T_d, but must admit that this has not been a universally accepted position.

In fact, the status of Organismic Darwinism is a very vexing question, which has received considerable attention in both the biological and philosophical literature; the debate has been positively acrimonious on occasion. The issue is clearly a matter of some importance for my stated purposes here—for if the status of Darwinism is unclear even within biology, its extension into other fields must necessarily be suspect.

I shall not presume to offer any definitive "resolution" to this question, but will simply attempt to make my own position as clear as possible.

I consider two general questions: whether Darwinism may be a mere *tautology*, or, failing that, whether it may be *irrefutable* (and thus *metaphysical*). These two issues are commonly conflated—for example by Maynard Smith (1969). It is of course true that all tautologies are untestable and therefore (in a trivial sense) metaphysical, but the converse does not hold—theories may well be metaphysical yet not tautologous. Thus, given that Darwinism is *not* actually tautological then its testability becomes entirely moot, and must be considered in its own right.

5.1 Tautology...

> My dear Darwin,—I have been so repeatedly struck by the utter inability of numbers of intelligent persons to see clearly, or at all, the self-acting and necessary effects of Natural Selection, that I am led to conclude that the term itself, and your mode of illustrating it, however clear and beautiful to many of us, are yet not the best adapted to impress it on the general naturalist public ... I wish, therefore, to suggest to you the possibility of entirely avoiding this source of misconception in your great work (if not now too late), and also in any future editions of the 'Origin', and I think it may be done without difficulty and very effectually by adopting Spencer's term (which he generally uses in preference to Natural Selection), viz. 'Survival of the Fittest'.
>
> Wallace (1866)
> (Quoted by Dawkins 1982, pp. 179–180)

I believe that the charge of tautology levelled against Organismic Darwinism is quite mistaken. However, contrary to, for example, Dawkins (1982, pp. 180–181), I also consider that the confusion is deep seated and subtle. In particular, I believe that there are (at least) three quite different *kinds* of misconception involved.

There are a number of more or less authoritative replies to the tautology charge already available in the literature (some of which I shall comment upon below); but these have, in general, only recognised one or another of the various possible misconceptions. I am not aware of any previous discussion which has clearly distinguished all three kinds of misconception which I identify here. I suggest that this may explain why this debate has sometimes appeared interminable: the participants have frequently been talking at cross purposes. In itself, this would justify the somewhat lengthy discussion given here; but the discussion is also justifiable in its own terms for it illuminates some quite important aspects of Darwinian theory which might not otherwise be explicitly dealt with.

The discussion will (not surprisingly) be quite complex and potentially confusing. Therefore let me outline the general structure in advance.

The arguments all revolve around Spencer's unfortunate phrase "the survival of the fittest". I shall initially show how the phrase can be interpreted as an (approximately) correct, non-tautologous, statement of T_{cs} (or, more generally, T_{hs})—which is, presumably, the interpretation Spencer intended. The first misconception which I consider is such that the phrase is still "correctly" interpreted (i.e. as a statement of T_{cs}), but, in a certain peculiar and austere sense, is labelled as tautologous *anyway.* Under the second misconception, there are two distinct ways in which the phrase can more or less correctly be interpreted as a *definition* of "fitness"; both of these interpretations are, of course, tautologous, but they are not statements of T_{cs} and do not impinge upon its status. Under the third (and final) misconception, there is an interpretation of the phrase which is not strictly tautologous, but is not equivalent to T_{cs} (or T_{hs}) either (and is, in fact, mistaken). This last error does involve an element of circular reasoning, and might therefore be still said to "smack" of tautology; in many ways it is the most pernicious misconception of all.

In presenting this analysis I do not claim that my taxonomy is complete or unique. In practice, various combinations and permutations of the errors identified below may well be simultaneously present in any single author's treatment; and, of course, there may be new errors which I am unwittingly originating, in carrying out this very analysis.

5.1.1 On "survival"

To analyse the arguments effectively, it is first necessary to distinguish two quite separate notions of "survival":

> **Survival$_1$:** The survival of individual organisms (or D-actors).

Survival$_1$ obviously does not refer to survival in any absolute sense: individual organisms are essentially mortal, and have a finite lifetime—it is hardly meaningful to speak of individual organisms "surviving" without qualification. However, some organisms do survive *longer* relative to others. More generally, there is potentially a valid notion of "mean survival$_1$"—i.e. that a certain "kind" (S-class?) of organism may, in given environmental conditions, show a consistent distribution of values for survival$_1$. Furthermore, if (and only if) there exist *heritable* characteristic(s) distinguishing such different "kinds" then they can serve to differentiate S-classes (and, thus, S-lineages). In that case coherent S-lineages, exhibiting distinctive (statistical) distributions of survival$_1$, can be formed. This distribution of survival$_1$ would be a characteristic of an S-lineage, and its mean value would be essentially equivalent to the reciprocal of what has previously been called S-mortality. In fact, (mean) survival$_1$ can be thought of as a crude measure of S-value (crude because the latter depends, at least, on S-fecundity as well as S-mortality); in particular it will generally be true that the greater the value of (mean) survival$_1$ then the greater the S-value of the corresponding S-lineage.

I emphasize that (mean) survival$_1$ (as with S-value proper) is defined as an objective characteristic of an S-lineage in given environmental conditions, which can (in principle) be evaluated independently of any knowledge of the success or otherwise of the S-lineage in competition with other S-lineages.

> **Survival$_2$:** The survival of organism lineages (in particular, S-lineages).

Survival$_2$ *does* (potentially) refer to "absolute" survival—in the sense that organism lineages can (apparently) survive indefinitely long.

Survival$_2$ may also be related to S-value, but not in the relatively direct way which holds for survival$_1$. Let us suppose that, for *independent* reasons, we believe that two competing S-lineages have significantly different S-values. Then we can infer that whichever S-lineage is the eventual survivor$_2$ must have the greater value of S-value. Note carefully that this inference is valid if and only if we *already* know that we are dealing with a case of S-lineage selection.

It should be clear that survival$_1$ and survival$_2$ are not the same thing. While they *may* be related (ultimately via T_{cs} as we shall see), this relationship is a contingent one; it would not hold if, for example, there were no inheritable characteristics which were well correlated with survival$_1$; or if the S-lineages distinguished by different values of (mean) survival$_1$ were not actually in competition with each other etc.

5.1.2 What Spencer Meant To Say?

I suggest that the only correct interpretation of Spencer's phrase (which is not *necessarily* the interpretation Spencer himself intended) is the following:

we interpret "survival" as survival$_2$ (i.e. survival of (S-)lineages); we interpret "fitness" as *S-value*; "the survival of the fittest" is then at once the assertion that:

- There exist at least some lineages which are fitter than others (which implicitly requires fitness to be hereditary to some extent). Lineages so distinguished are, precisely, S-lineages.
- At least some of these S-lineages, of differing fitness, are in competition with each other.
- As a result, S-lineage selection occurs and, in any such episode, the surviving S-lineage will be the fitter (i.e. having the greater S-value).

As it stands this is simply T_{cs}. Assuming (by default) the extension of this process indefinitely back in evolutionary time, we get T_{hs}. It is not tautological—any more than T_{cs} or T_{hs} is.

Again, it must be emphasized that this non-tautologous formulation relies on the fact that S-value (fitness) is not *defined* by lineage "survival"—it is, in principle, something that can be assessed of an S-lineage isolated from competition (it is, in fact, a prerequisite for the operation of selection). However, confusion might arise in cases where, for *independent* reasons, one *already* believes that the displacement of one S-lineage by another is, in fact, a case of selection: in that scenario one can validly, and non-tautologically, infer that the surviving S-lineage must have had the greater S-value.

Essentially this (correct) interpretation has been commonly identified in the literature—for example by Medawar (see Moorhead & Kaplan 1967, p. 12), and Hodge (1983, p. 58).

5.1.3 Misconception 1: Logic

As detailed in the previous section, once the premises for natural selection are granted (and this is an empirical question) then the outcome—"the survival of the fittest"—is assured. That is, if we adopt these premises as *axioms*, in the sense of a formal logical system (i.e. they are taken to be "true" by definition), then "the survival of the fittest" becomes a theorem of the system, which is to say, in the strict terminology of formal logic, a *tautology.*

This is, of course, a technically valid point; but it can hardly be called a *criticism.* It amounts to interpreting Spencer's phrase *only* as the (necessarily "tautological") conclusion of a certain deduction—rather than as an implicit assertion of the truth of the premises which lead to that conclusion. This is at best pedantic, at worst misleading. It is equivalent to saying that $E = mc^2$ (say) is a tautology—given the relevant properties of E, m and c. As Maynard Smith has put it,

> Of course Darwinism contains tautological features: any scientific theory containing two lines of algebra does so...
>
> (Maynard Smith 1969, p. 85)

This is such a peculiar misconception that it seems difficult to believe that it should genuinely arise. In practise I suggest that it does not normally arise in isolation, but may be combined with one of the other distinct misconceptions yet to be described. Having said that, there *is* at least one case where this misconception seems to have been uniquely involved:

> The notion of natural selection depends on the empirically verifiable observation that offspring on the average resemble their parents more closely than they do the other members of the population, that individuals are not all the same; that all environments are not the same. Concepts such as natural selection by the survival of the fittest are tautologous; that is, they simply restate the fact that only the properties of organisms which survive to produce offspring, or to produce more offspring than their cohorts, will appear in succeeding generations.
>
> Eden (1967, p. 5)

Eden does seem to use the correct interpretation of Spencer's phrase (notwithstanding the fact that he immediately goes on to use "survive" in the sense of survival$_1$ rather than survival$_2$); but insofar as he describes it as a tautology he merely seems to mean that *any* valid deduction ("restatement") from true premises is a "tautology". While formally correct, the observation does not add anything except, possibly, confusion.

Consider also, the following comment from Dawkins:

> Biologists thought they needed a word for that hypothetical quantity that tends to be maximized as a result of natural selection. They could have chosen 'selective potential', or 'survivability', or 'W' but in fact they lit upon 'fitness'. They did the equivalent of recognizing that the definition they were seeking must be 'whatever it takes to

> make the survival of the fittest into a tautology'. They redefined fitness accordingly.
>
> Dawkins (1982, pp. 181–182)

I suggest that what Dawkins means here is that fitness can be (indeed has been) defined as whatever it takes to make "the survival$_2$ of the fittest" into a logical consequence of the existence of heritable fitness variations and competition. I should emphasize that, in context, Dawkins is *not* suggesting that this "tautology" can be translated into any *criticism* of Darwinism; but, as with Eden's version, I still find the reference to "tautology" to be confusing and gratuitous.

5.1.4 Misconception 2: Words

Spencer's phrase does reduce strictly to a tautology if "fitness" is equated with "survival"; that is, the phrase is read as a *definition* of fitness (and definitions are, of course, a paradigmatic case of tautology). This formulation of the tautology argument is the most common; a good example would be that of Popper (1965, pp. 241–242).[21] As Dawkins puts it, this kind of argument is "a remarkable example of the elevation of words above their station" (Dawkins 1982, p. 181).

This misconception leads to tautology regardless of whether we take "survival" to mean survival$_1$ or survival$_2$. Both cases are conceptually possible, although they have quite different flavours. In general, writers suffering from this misconception are not clear about which sense they intend. In fact, the most likely scenario may be a failure to distinguish that there *are* two possible, but distinct, strictly tautologous interpretations—for if that fact is once recognised, the possibility of a non-tautologous (and correct) interpretation more or less immediately presents itself.[22]

The two distinct cases of this misconception are as follows:

Case 1: *the survival$_1$ of the survivors$_1$*

This amounts to defining fitness as (mean) survival$_1$.

Now this interpretation is not "incorrect". As already discussed above, survival$_1$ is indeed a possible, though extremely crude, measure of S-value; and "fitness" can be (and commonly is) interpreted as synonymous with S-value. So it is not entirely unreasonable to define fitness as (mean) survival$_1$ (other things, particularly S-fecundity, being equal).

But of course, under this interpretation, the phrase is no longer an expression of (the theory of) natural selection—it is merely a preliminary definition. Pointing out its tautologous nature cannot impinge at all on the status of T_{cs} or T_{hs}.

Case 2: *the survival$_2$ of the survivors$_2$*

This amounts to defining fitness as survival$_2$.

Unlike case 1, this is hardly even a coherent definition. Since survival$_2$ is (at any given time) a binary valued quantity, it is at least peculiar to equate it with "fitness" which, on any common sense interpretation, should be continuous valued.

But let us stretch this point, for the time being. We *could* consider the phrase as equating fitness with survival$_2$, *regardless of whether natural selection is known to be operational.* But, as far as I am aware, this would be a usage of "fitness" which has never been seriously proposed, is counterintuitive, and would be of no apparent utility.

This leaves only the possibility that we consider the phrase as a definition of fitness *only in cases where we have prior, independent, knowledge of the operation of natural selection.*

Well, in this case we cannot say it is positively incorrect—the operation of selection guarantees precisely that survival$_2$ *will* be related to S-value (the "normal" meaning of fitness), as already discussed for the correct, non-tautologous, interpretation of Spencer's phrase. At this point we are back to a similar situation to that obtaining with Case 1: the interpretation can "reasonably" be adopted, but it is no longer an expression of (the theory of) natural selection—it is merely a preliminary definition (and a rather confusing one at that). Adopting this definition, we would then have to introduce some additional term other than fitness (S-value perhaps?), whose definition would *not* be already contingent on the operation of selection, before we could even formulate T_{cs} properly. But, in any case, we again conclude that

---

[21] However note that Popper has since repudiated this analysis, as I shall discuss in section 5.2 below.

[22] There may be a lingering misconception that all tautologies are equivalent, so that the possibility of "distinct" tautologies cannot arise; this is not the case. All tautologies have the same *truth* value (namely, unconditionally "true"), so that they are *logically* equivalent—but this is not at all the same thing. "Cats are a kind of domesticated feline" and "π is the ratio of the circumference to the diameter of a circle" are (*qua* definitions) two *different* tautologies.

pointing out the tautologous nature of this interpretation cannot impinge at all on the status of T_{cs}, or T_{hs}.

5.1.5 Misconception 3: Adaptation

The final misconception which arises does not strictly involve a tautology, and is only incidentally inspired or supported by Spencer's phrase. However, it does involve a degree of "circular" reasoning, is sometimes *said* to be tautologous, and is commonly associated with the strictly tautologous misconception(s) of the previous section. It is therefore appropriate to consider it here.

I shall present the "argument" in what I consider to be its plainest form, but caution that it is rarely if ever expressed in such explicit terms:

1. T_d is proposed as a solution of P_d—i.e. as an explanation for the growth of adaptive complexity in the phylogenetic tree.
2. For T_d to successfully solve P_d would require that it *predict* (statistically or otherwise) the growth of adaptive complexity.
3. At best, the only thing that T_d actually predicts is a growth of S-value in certain lineages.
4. Thus T_d can be said to solve P_d only by (re-)defining adaptive complexity to be the same as S-value (fitness). This effectively uses a circular definition, which is just such that the problem to be solved (originally P_d) becomes simply "whatever problem *can* be solved (by T_d)".

Arguments of this sort are tacitly involved whenever there is a debate about the relationship between "fitness" and "adaptation". Darwinism is first taken to be a theory of the growth of "adaptation" (informally conceived of in terms of adaptive *complexity*); but it is then noticed that Darwinism *per se* can actually only explain the growth of "fitness" (in the sense of S-value); so it seems that it can work as an explanation only if adaptation is *defined* as equivalent to fitness (and we forget our original informal notion of adaptive complexity, or adaptation to an environment). This does not turn Darwinism into a tautology in any strict sense (though it does involve a kind of circularity); but if this step is taken then Darwinism loses its ability to solve the kind of problem we were originally interested in—for the terms of the problem no longer appear in the theory. It is thus greatly diminished in scope and significance.

Popper raised essentially this problem in what was (as far as I am aware) his earliest consideration of the status of Organismic Darwinism—his Herbert Spencer lecture, *Evolution and the Tree of Knowledge*:

> ... survival, or success in the sense of an increase in numbers, may be due to either of two distinguishable circumstances. A species may succeed or prosper because it has managed, say, to improve its speed, or its teeth, or its skill, or its intelligence; or it may succeed or prosper merely because it has managed to increase its fecundity. It is clear that a sufficient increase in fecundity depending fundamentally on genetical factors, or a shortening of the period of immaturity, may have the same survival value as, or even a greater survival value than, say, an increase in skill or in intelligence.
>
> ... But be this as it may, it should be possible, I think, to [measure] the *success in the adaptation of the individual organisms* of a species ...
>
> Without some distinction such as this ... we are liable to lose sight of the original problems of Lamarck and Darwin, and especially of the explanatory power of Darwin's theory ...
>
> Popper (1961, pp. 271–272)

Popper is essentially pointing out that the temptation to equate adaptation (or adaptive complexity) with fitness (or S-value) must be resisted at all costs, for otherwise we lose contact with the problems we wish to solve. Lewontin (1978) has discussed this problem in very similar terms. Hull has also recently made much the same point, concluding that the requirement to identify adaptive complexity (which he actually calls "fitness") independently of S-value (which he terms "differential perpetuation") cannot be circumvented "without evolutionary theory degenerating into an empirically empty formalism" (Hull 1980, pp. 318–319).

Now I have, indeed, been careful not to define adaptive complexity in terms of S-value (or fitness). Granted, I have not attempted any formal or detailed definition of adaptive complexity; and I have particularly eschewed any attempt to establish a metric for it. But, as already discussed in section 3.1, this vagueness is not untypical in the field; and I would argue that my general formulation in terms of *inborn knowledge* (Popper 1961, pp. 258–259) or *anticipatory systems* (Rosen 1985a) is still a more definite ontological commitment that is usual.

It must be repeated that misconception 3 does not strictly involve a tautology. It is in this light that we must read Hodge's (1983, p. 58) claim that it is "a mistake to defend natural selection against the tautology objection by proposing criteria of fitness independent of reproductive success". I suggest that Hodge is here referring only to the *strictly* tautologous arguments already discussed (particularly the two cases of misconception 2), and *not* to misconception 3, which has quite a separate character. Dunbar (1982, p. 10), on the other hand, *rejects* the argument that T_{cs} cannot be tautological because it is empirically testable (i.e. the kind of argument offered by Hodge), saying that "this claim misses the point entirely". I contend that Dunbar is effectively taking up a position precisely complementary to that of Hodge, confining his attention exclusively to misconception 3 and ignoring or dismissing misconceptions 1 and/or 2. My position is that, despite the apparent contradiction between Hodge and Dunbar, they are actually both correct, so far as they go—but they are discussing different problems (I shall consider Dunbar's analysis in more detail below).

At this point my argument is that misconception 3 is mistaken in concluding that adaptation or adaptive complexity should, or must, be defined in terms of S-value (fitness). I therefore insist on retaining essentially independent definitions of adaptive complexity and fitness (and thus retain P_d in its interesting form). What then are we to make of the original criticism—that T_d cannot actually solve this problem, for it does not predict the growth of adaptive complexity (so-defined)?

This brings us to the nub of the problem, which is to ask how much we can sensibly ask of a proposed solution of P_d. The error is in supposing that a solution must (or even can) take the form of some general theory which predicts a growth in adaptive complexity. This derives in part from a fundamental misunderstanding of what P_d actually says. P_d recognises that there has been a sustained, and progressive, growth in adaptive complexity, in at least some evolutionary lineages, and asks for an explanation of how this could be (preferably one which does not assume the pre-existence of an even more complex creator). It does *not* say that growth in adaptive complexity *must* occur (in general or in particular lineages); or that it had to occur in the particular way which it did; or that it must continue occurring. *But only the latter kinds of problem would call for a solution which incorporates a general prediction of growth of adaptive complexity.*

We might characterise the general difficulty here as a supposition that P_d entails some kind of guaranteed, monotonic, "progress" (in adaptive complexity). It is a harking back to the "great chain of being". As Gould puts it: "The familiar iconographies of evolution are all directed—sometimes crudely, sometimes subtly—toward reinforcing a comfortable view of human inevitability and superiority" (Gould 1989, p. 28). The idea of necessary progress in evolution is so deeply entrenched that it is very difficult to free oneself of it. Not even Darwin himself was completely immune. Although, as discussed in section 3.3 above, he explicitly emphasized that the *absence* of "perfection" in biological organisms should be interpreted as positive evidence in favour of the operation of natural selection (at least as compared with a theistic theory), we still find him remarking, in the concluding pages of *The Origin*, that "… as natural selection works solely by and for the good of each being, *all corporeal and mental endowments will tend to progress towards perfection*" (Darwin 1859, Chapter XIV, emphasis added).[23]

Once, however, we free ourselves from the preconception that organismic evolution is a story of necessary or inevitable "progress"—and excise this idea from P_d—then the way is open to consider T_d as offering at least a partial solution; indeed, as I have already argued (section 4.4), it may offer as good a (general) solution as it is possible to give *in principle.*

Briefly, as has been commented upon several times, T_d cannot and does not predict the growth of adaptive complexity. However, given that adaptive complexity *has* grown, T_d can retrospectively offer a generalised explanation for it—namely that the growth of adaptive complexity "happens" to have been correlated with the growth of S-value—and, crucially, T_d can do this without postulating the existence of any "designer" (anticipatory system) directing the overall course of evolution.

T_d in itself, does not explain why adaptive complexity should be correlated with S-value either in general, or in any particular case; not does it explain how (heritable) adaptive complexity can increase at all, in general or in any particular case. In addressing P_d it *does* entail that some (heritable) increases in adaptive complexity have occurred, and that, of these, at least some have been correlated with S-value; but it does not require increases in adaptive complexity to have been "designed" or "anticipated" (i.e. preferentially generated), or that they should have been *universally* correlated with S-value.

Now, of course, T_d *does* predict more or less monotonic "progress" in one very specialised sense: in-

---

[23] See (Gould 1978, Essay 4) for a more detailed discussion of whether Darwin "really" endorsed the idea of necessary progress in evolution.

crease in S-value (within some lineages). But that much is true of all D-systems, and does not impinge at all on the central problem of the growth of adaptive complexity. No doubt we could conduct a certain amount of evolutionary theorizing without ever referring to adaptive complexity: but we can certainly never solve P_d. The crucial extra step, which is rarely made explicit is to say that, given T_d, adaptive complexity can grow in evolutionary lineages if and only if at least some increases in adaptive complexity occur, and, of these, at least some are correlated with a net increase in S-value (i.e. are selectively retained).

Of course, the reasoning given here only works at all if it is accepted that an increase in adaptive complexity *might* be correlated with greater S-value. If I had defined adaptive complexity in some (strange) way which was intrinsically opposed to S-value (say, necessarily involving unconditional altruism on the part of the S-lineage) then the argument could not go through. But my actual definition—which corresponds to something like "inate knowledge"—does precisely have the characteristic that, *ceteris paribus*, it may be expected to be correlated with higher S-values. But the explicit inclusion of the *ceteris paribus* clause is crucial here; to ignore or omit it would effectively mean a reversion to equating adaptive complexity with S-value, and thus defining away the real problem.

None of this is to say that the growth of adaptive complexity cannot be explained (or even predicted); it simply claims that there cannot be any *general* theory of it (as always, this is just another way of denying the existence of a logic of induction). T_d asserts precisely that, for all *particular* historical cases of an increase in adaptive complexity, there is a *particular* explanation, involving an undirected or unbiased variation in adaptive complexity (which happened to be an increase) which was selectively favoured. It may or may not be possible to organise these particular, historical, explanations into a smaller number of more general cases: but I take the view that it will not be possible to translate these particular explanations, nor generalisations of them, into predictions for continuing growth of adaptive complexity into the future. The aggregation of all the particular cases (if such could be individually established) would then be the complete (historical) "explanation" of all growth of adaptive complexity in the biological world. But T_d does not assert that such growth had to occur, nor that it will continue into the future, nor even that it would necessarily recur on another "similar" planet.

The confusions and misconceptions discussed here have centred on the distinction between "adaptation" and "fitness". For this reason I have tried to avoid these terms in my own general presentation of Darwinism: the only lingering remnant is the "adaptive" in "adaptive complexity". I have retained this in deference to the existing biological literature, but I suggest that it might actually be better to eliminate even this concession. In speaking of the "adaptation" of biological organisms it seems almost impossible not to think in terms specifically relating to their success in living and procreating—which is to positively invite a reversion to the relatively sterile concept of S-value.

A particularly "good" example of this is Dunbar's analysis (Dunbar 1982). He is very clear about the *need* to distinguish "adaptation" and "fitness". Furthermore, he seems to adopt much the same kind of distinction as I have suggested above, interpreting adaptation in terms of "problem-solving" (Dunbar 1982, p. 11) (following Lewontin 1978, among others). He seems to recognise the essential *independence* of the two concepts when he cites Dobzhansky (apparently favourably) as saying that "we cannot draw inferences about fitness from a knowledge of adaptation, nor of adaptation from a knowledge of fitness" (Dunbar 1982, p. 14). However, he then goes on to explicitly deal with the alleged circularity of Darwinian explanations, in the following terms:

> The relationship between the concepts of adaptation and fitness might seem to confirm the worst fears of the anti-Darwinians. Each appears to depend on the other in a way which makes them virtually inseparable. It is, however, crucial to appreciate that they are not *definitionally* interdependent: adaptation is not *defined* in terms of fitness, nor vice versa. Adaptation (and hence reproductive success) is defined with reference to *individuals*, whereas fitness is defined with reference to genes and is thus a characteristic of populations.
>
> (Dunbar 1982, p. 16–17)

Dunbar here seems to suggest that the distinction between adaptation and fitness is (merely?) a distinction between properties of organisms and (consequent) properties of lineages. The best interpretation I can offer of this is as a somewhat tortuous reference to what I have called misconception 2 above—effectively a failure to distinguish survival$_1$ and survival$_2$. Having thus retreated from the real issue—misconception 3—Dunbar finds that he must admit that a significant circularity may still remain in Darwinian theory. It is perhaps not surprising that he then resorts to the philosophical relativism of Kuhn and Feyerabend as his final defence of Darwinism against circularity. I, of course, take the view that such a conclusion is unsatisfactory and unnecessary—that the conceptual independence of fitness and adaptation (S-value and adaptive complexity) can and should be recognised, and this can be done without depriving Darwinism of its power as an explanatory schema. However, it does underline the point that the terminology of "fitness" and "adaptation" may be critically flawed. Thus, if one confines oneself to discussion in terms of (inate) knowledge, or (inate) anticipatory models, instead of adaptation or adaptive complexity, it may be easier to remember that there is no *necessary* connection between these things and their retention or growth under natural selection. It should then be clear that any connection which may exist will have to be individually argued for in each particular case.

5.2 ... or Metaphysics?

As with the idea that Darwinism may be tautological, the assertion that it is essentially metaphysical has been more or less independently suggested (and criticised) by a number of different authors.

The criticism of the so-called *adaptationist programme* by Gould and Lewontin is an example of this kind of argument (Gould & Lewontin 1979, p. 589). Indeed, Lewontin has explicitly claimed that "the adaptationist programme makes of adaptation a metaphysical postulate, not only incapable of refutation, but necessarily confirmed by every observation" (Lewontin 1977, as quoted by Maynard Smith 1978, p. 38).

However, as already discussed at length in section 3.3 above, the claim put forward by Gould and Lewontin is not strictly that Darwinism *per se* is metaphysical. Indeed, they accept the reality of natural selection (i.e. T_{hs}); their argument is about which evolutionary phenomena are explicable in terms of selection. The adaptationist programme, which they criticise, presumes that *all* evolutionary phenomena (or organismic attributes) are the direct outcome of selection.[24] Conversely, it seems clear that Gould and Lewontin accept that what I have called complex adaptation, where it exists, does demand an explanation in terms of selection; and, while such explanations may, individually, be almost impossible to test in practise, they are testable in principle, and are not therefore metaphysical. Thus, *in terms of the problem of complex adaptation (P_d),* Gould and Lewontin appear to accept that T_d is not metaphysical, and, indeed, that it is the best theory currently available.

Concerns of a similar sort to those raised by Gould and Lewontin have been independently discussed by various other workers. For example, they were prominent in the discussions at the *Wistar* symposium on "Mathematical Challenges to the neo-Darwinian Interpretation of Evolution" (Moorhead & Kaplan 1967). Although Popper was not present at this symposium, he was repeatedly cited (directly or indirectly) as the source for such concerns. Thus, for example, Medawar makes the following comment in introducing the symposium:

> Then there are philosophical or methodological objections to evolutionary theory. They have been very well voiced by Professor Karl Popper—that the current neo-Darwinian Theory has the methodological defect of explaining too much. It is too difficult to imagine or envisage an evolutionary episode which could *not* be explained by the formulae of neo-Darwinism.

Unfortunately, there are no detailed citations to original sources, and I am not aware of Popper ever having published *exactly* this criticism of Darwinism. In any case, my response to this argument is essentially as already discussed in the case of Gould and Lewontin: while T_d is, undoubtedly, difficult to test in respect of the evolution of specific complex adaptations, it is still testable in principle (i.e. is not metaphysical) and is the best theory currently available.

However, Popper *has* published a slightly different argument for regarding Darwinism as metaphysical—or, more precisely, as a *metaphysical research programme* (Popper 1974a, Section 37). This arose (at least partly) because, as already noted in section 5.1.4, Popper's earliest considerations of the status of Darwinian theory suffered from a form of misconception 2, and he felt that the theory was therefore "almost tautological"; yet he also felt that, *despite this,* the theory had considerable explanatory power. Popper seems to have thought that this apparent contradiction might be resolved by regarding Darwinism as a metaphysical research programme. While I think his interpretation of Darwinism as tautologous was mistaken, I actually agree that, in a certain limited sense, it *can* usefully be regarded as metaphysical.

Firstly, as discussed in (McMullin 1992), I consider that it is not unreasonable to describe the

---

[24] In this respect, the term "adaptationist" is unfortunate, as it invites a form of the tautology misconception 3; "selectionist" might be less prejudicial.

abstract form of Darwinian theory, presented in that essay, as a metaphysical research programme in Popper's sense. It is not a scientific theory until the primitive entities (especially D-actors) are given some specific empirical interpretation. It must be emphasized that to view this admission (of the metaphysical nature of the abstract theory) as a *criticism* of any particular interpretation of the theory (such as Organismic Darwinism) would be to indulge again in a form of the tautology misconception 1. This, for example, is the only way in which I can understand one of Peters' purported criticisms of Organismic Darwinism (Peters 1976, p. 4), which is apparently based on its being a specific interpretation of the axiomatic Darwinism of Williams (1970).

Having said that, it must be recognised that the abstract or axiomatic form of Darwinism is metaphysical in a deeper or more profound sense than the conventional abstract theories underlying all science. The general kind of problem which abstract Darwinism seeks to solve is the growth of knowledge; and its mechanism of solution entails a refusal to make predictions—it "works" (in the face of the impossibility of a logic of induction) precisely by *declining* to predict the growth of knowledge. This is a quite unique kind of abstract theory. It follows that, *even* when the abstract theory is interpreted in specific empirical terms (such as in the form of Organismic Darwinism) it can never predict the *future* growth of knowledge. We must say that, as long as such a particular interpretation of Darwinism is viewed as an *historical* theory of the *past* growth of knowledge it is perfectly testable (in terms of "retrodictions") and scientific; but if it is mistaken for a "universal law" of the growth of knowledge, then, since it cannot predict such growth, it must be treated as metaphysical. The (metaphysical) position adopted here is, of course, that no "universal law" of the growth of knowledge exists.

It can be seen that this argument for viewing even Organismic Darwinism (as opposed to Darwinism in the abstract) as metaphysical hinges on its (in)ability to predict the future growth of knowledge, or adaptive complexity; thus it is closely related to what I have labelled tautology misconception 3. This is brought out clearly by considering Popper's most substantive presentation of this viewpoint:

> . . . assume that we find life on Mars consisting of exactly three species of bacteria with a genetic outfit similar to that of three terrestrial species. Is [organismic] Darwinism refuted? By no means. We shall say that these three species were the only forms among the many mutants which were sufficiently well adjusted to survive. And we shall say the same if there is only one species (or none). Thus Darwinism does not really *predict* the evolution of variety. It therefore cannot really *explain* it. At best, it can predict the evolution of variety under "favourable conditions". But it is hardly possible to describe in general terms what favourable conditions are—except that, in their presence, a variety of forms will emerge.
>
> And yet I believe I have taken the theory almost at its best—almost in its most testable form. One might say that it "almost predicts" a great variety of forms of life. In other fields, its predictive or explanatory power is still more disappointing. Take "adaptation". At first sight natural selection appears to explain it, and in a way it does; but hardly in a scientific way. To say that a species now living is adapted to its environment is, in fact, almost tautological. Indeed we use the terms "adaptation" and "selection" in such a way that we can say that, if a species were not adapted, it would have been eliminated by natural selection. Similarly, if a species has been eliminated it must have been ill adapted to the conditions. Adaptation or fitness is *defined* by modern evolutionists as survival value, and can be measured by actual success in survival: there is hardly any possibility of testing a theory as feeble as this.
>
> Popper (1974a, pp. 136–137)

As already mentioned, Popper had earlier (Popper 1961) recognised the danger of misconception 3, and the consequent need to keep adaptation and fitness (selection) clearly separated; but in this later consideration of the problem he seems to have decided that such separation cannot be achieved. Viewed as a putative theory of the growth of adaptive complexity (i.e. of the evolution of a "rich variety" of more or less "well adapted" forms), Darwinism then becomes irrefutable (metaphysical), for, no matter how little the variety or adaptation we observe, it *could* still result from Darwinian processes.

Clearly, I agree with the essence of Popper's argument; the difference is that instead of disgarding Darwinism, I disgard the idea that adaptation and fitness be defined in terms of one another. Granted, Darwinism cannot then "predict" the growth of adaptation or adaptive complexity.

But, once adaptation is interpreted in terms of *knowledge* this becomes precisely consistent with Popper's general evolutionary epistemology, and is seen as the strongest kind of theory we can expect. And, as a theory of the *historical* growth of adaptive complexity in the biological world, it is perfectly scientific.

I have expended some effort in considering Popper's criticism of the scientific status of Darwinism because I think it relates to some difficult and important issues, which are relevant objectives of this essay as a whole. However, it must be added that Popper himself has, in any case, since modified his views significantly (Popper 1978).

In particular, Popper has now recognised and corrected the error implicit in tautology misconception 2, and accepted that (Organismic) Darwinism can be so formulated that it is definitely not tautologous. Unfortunately, he then goes on to say that, in such a form, it is literally false (has been refuted). At first sight this is an even worse accusation than the original assertion that the theory was (almost) tautologous and/or metaphysical. However, Popper's reformulation is the strong one that *all* aspects of the phylogenetic tree are the outcome of cumulative natural selection; such a strong claim, which is essentially equivalent to the adaptationist programme criticised by Gould and Lewontin, is, indeed, false. Popper does not explicitly consider the lesser (but still non-tautologous) claim that all instances of the growth of adaptive complexity (i.e. my P_d) are the outcome of cumulative natural selection (i.e. my T_d), and certainly does not argue that *this* formulation has been refuted; so there is, in fact, no conflict with the views I have expressed.

In conclusion, let me reiterate that I consider T_d, viewed as an historical theory of the growth of adaptive, organismic, complexity, to qualify as a good scientific theory—*despite* the fact that actually testing it in specific cases is enormously difficult. More specifically, T_d qualifies as scientific according to *Popper's* criteria. I emphasize this last point because, even though Popper might be called a "naïve" falsificationist with respect to the *logic* of (scientific) theories, he has always been a critical falsificationist with respect to the *methodology* of actually carrying out scientific research. This point has, apparently, been commonly misunderstood or misrepresented (Magee 1973, pp. 23–24; Popper 1974b, pp. 981–984). The relevance of the distinction here is that it can be a perfectly consistent Popperian position to assert that Darwinism is scientific by virtue of its (logical) falsifiability, *even* if such falsifiability is *methodologically* almost impossible to exploit (i.e. tests which could falsify the theory may be perfectly conceivable, yet wholly impractical).

In this section I have given a detailed discussion of Popper's assertions that Darwinism may be tautological and/or metaphysical. I have noted that these assertions were, at the least, confused, and that Popper himself has subsequently clarified and corrected his position. I believe that this puts the substantive issues to rest. In closing, however, I note that Popper's excursions into evolutionary biology seem, for whatever reason, to have also occasionally provoked some quite mistaken and/or irrelevant responses from professional biologists; by definition, these do not affect the important issues here, and therefore I relegate even my brief discussion of them to Appendix B.

6 Conclusion

This essay has been concerned with the nature and status of Darwinian theory in its conventional biological setting. This is important because there are persistent suggestions that, in some sense, Organismic Darwinism is significantly or even fatally flawed. Clearly, if such doubts were justified, then it would hardly be appropriate to consider the extension of Darwinism to any other field.

I have suggested that such misgivings are unfounded: that a careful analysis of Organismic Darwinism reveals it to be perfectly scientific, and currently the best available theory to explain the growth of adaptive organismic complexity. On the other hand, I also hope to have established that this result is by no means straightforward. Darwinian theory is structurally complex, and is, furthermore, extremely difficult to test *in its entirety.* While I have argued that it is better corroborated than its competitors, these competitors do deserve serious consideration, and the limitations in the *practical* corroborability of Darwinism need to be acknowledged. Similarly, great care must be taken to avoid a variety of more or less subtle misconceptions about the nature of Organismic Darwinism.

The most significant single result claimed in this essay is that Organismic Darwinism is necessarily, and perhaps irredeemably, incomplete: heritable variation and natural selection do not and cannot, in themselves, guarantee any growth in adaptive complexity.

Finally: this essay has been concerned exclusively with establishing that Darwinism in general, and Organismic Darwinism in particular, provide a more or less satisfactory *overall* framework for the discussion of evolutionary biology; granting that this has now been established, it must still be admitted that there is considerable on-going debate, *within* the scope of T_d, as to its detailed workings and interpretation. These are inner issues in evolutionary biology, which do not challenge the *essential* correctness of T_d; nonetheless they represent substantive disagreements which would necessarily colour any practical attempt to realise artificial Darwinism. I therefore intend, in a separate essay (McMullin Forthcoming), to at least establish what these disagreements are, if not to actually resolve them.

This is an internal Technical Report, and I rely on the informality of that medium to excuse the many rough edges remaining. In any case, I would greatly appreciate comments and criticism of any sort.

Acknowledgements

This essay is part of an ongoing attack on the problem of realising the spontaneous growth of Artificial Knowledge by Darwinian (or any other!) means. In this pursuit, I have benefited greatly from discussions with colleagues, particularly Noel Murphy in DCU, and John Kelly of University College Dublin. Paul McKevitt also made useful comments on an earlier draft of the essay. I am indebted to the School of Electronic Engineering in DCU (particularly through the agency of its Head, Charles McCorkell) for continuing encouragement, not to mention material support. All errors remain, of course, my own responsibility.

Appendices

A Dawkins' Microscope

Fisher's (1958, pp. 41–44) argument against saltationism (discussed in section 4.1 above) centers on the (im-)probability that a random movement in a Euclidean n-space will bring a given point closer to a specified target ("optimal") point; he gives exact results for the cases $n = 3$ and $n \to \infty$, and a qualitative discussion based on these. Dawkins (1986, pp. 230–236), on the other hand, in a generally laudable attempt to simplify Fisher's original argument, chooses to work with a system in which $n = 1$. Now this *is* actually derived from an hypothetical analogy given by Fisher:

> The conformity of these statistical requirements with common experience will be perceived by comparison with the mechanical adaptation of an instrument such as a microscope, when adjusted for distinct vision. If we imagine a derangement of the system by moving a little each of the lenses, either longitudinally or transversely, or by twisting through an angle, by altering the refractive index and transparency of the different components, or the curvature, or the polish of the interfaces, it is sufficiently obvious that any large derangement will have a very small probability of improving the adjustment, while in the case of alterations much less than the smallest of those intentionally effected by the maker or operator, the chance of improvement should be almost exactly one half.
>
> Fisher (1958, p. 44)

Note carefully, that, as originally stated by Fisher, this example involves multiple degrees of freedom (n is certainly more than 3) and is therefore a perfectly correct illustration of Fisher's abstract analysis. However, Dawkins' version of the example is rather different:

> Suppose that the *(sic)* lens is slightly lower than it ought to be for perfect focus, say a tenth of an inch too close to the slide. Now if we move it a small amount, say a hundredth of an inch, in a random direction, what are the odds that the focus will improve?
>
> Dawkins (1986, p. 232)

By talking specifically in terms of the alteration of a single real-valued parameter (the position of "the" lens) Dawkins implicitly shifts from Fisher's multiple degrees of freedom ("by moving a little *each* of the lenses" etc.) to just a single degree of freedom, which yields a mutational space of just one dimension ($n = 1$). This is a case which was not explicitly analysed by Fisher at all; and as it happens, the behaviour in this case is rather distinctive. Dawkins seems (at first at least) to have overlooked this for his qualitative discussion of the outcome still directly echoes Fisher's discussion of the case $n \to \infty$:

> I think it really will now be sufficiently obvious [!] that the smaller we make the move, the closer we shall approach the extreme case in which the odds of an improvement are one-half; and the larger we make the move, the closer we shall approach the extreme case in which the odds of an improvement are zero.
>
> Dawkins (1986, p. 232)

The technically correct statement for the $n = 1$ case is actually as follows (with r and d as defined by Fisher—see the discussion in section 4.1 above). The adjustment is $\pm r$ with equal probability of each of the two directions. Then $\forall\ r \geq d$ the probability of improvement is exactly 0 (as already noted, this holds regardless of n); and $\forall\ r < d$ the probability of improvement is exactly 0.5 (this is specific to the case $n = 1$). Note that the probability does *not* (in this special case) change continuously with r, approaching 0.5 only in the limit as $r \to 0$, as described by Dawkins; rather, it changes discontinuously at $r = d$, and is otherwise constant. Strangely, Dawkins *does* (implicitly) use this correct analysis when he comes to consider a specific numerical example:

> If the microscope starts 2 inches out of focus, then a random change of 1 inch has a 50 per cent chance of being an improvement, *just as a random change of one-hundredth of an inch has.*
>
> Dawkins (1986, p. 232, emphasis added)

Dawkins further confuses matters (for me at least) by stating that Fisher's argument "depends on the initial assumption that the microscope was already pretty close to being in focus before we even started making random adjustments" (Dawkins 1986, p.232). In fact, Fisher makes no formal or explicit statement of such an assumption (presumably Dawkins takes it as implicit in the reference, at the start of the paragraph quoted above, to the microscope being "adjusted for distinct vision"); and, more to the point, it is certainly not *necessary* to Fisher's argument. His result is, in effect, scaled relative to d, and therefore holds *regardless of how large d might be.* To be fair, Dawkins actually goes on to say this (more or less)—the point he wishes to make about the "initial" distance from the peak ($d/2$) being "small" seems to be quite separate (though I confess I am still not quite clear what this separate point is); but his conflation of the two issues seems to me to be another unnecessary distraction.

But to conclude on this issue, while Dawkins' discussion is technically slightly inconsistent, and therefore somewhat confusing, it would not merit discussion for that in itself. My real reason for reviewing Dawkins' version of Fisher's argument is to point out that, as stated, it yields only the *weak* form of the result—that adaptation (in this model of evolution at least) can be improved only by displacements of size $r < d$; whereas, Fisher's original argument does not even explicitly mention this weaker result, but emphasizes instead the case of "large" n ($n \to \infty$), which yields the distinct, and stronger, result that adaptation can be feasibly improved only by much smaller displacements than this (we must actually have $r \lesssim d/\sqrt{n}$, which is to say, $r \ll d$).

B Three Red Herrings

In the body of the report I detailed my views of the scientific status of Organismic Darwinism, including an extended consideration of Karl Popper's various comments on this issue. If the only response to Popper's analysis had been to criticise and highlight his error, then that could be the end of this discussion. Unfortunately, it is the case that, for whatever reason, Popper's analysis has provoked, *inter alia*, some responses which actually tend to further confuse the issues rather than clarify them, and these should therefore be considered here, even if only briefly.

Firstly, consider the following comment by Rose:

> ...But more than a century ago Darwin pointed out quite bluntly that there was no such thing as fact-collecting in a vacuum; facts are always collected *for* or *against* a particular hypothesis. Popper, of course, has made his career as a philosopher of science elaborating Darwin's insight (though without, I believe, ever crediting it, perhaps because of the reservations he has had, at least over a long part of his philosophical career, concerning the scientific status of Darwinian evolutionary theory).
>
> Rose (1986, p. 4)

This really seems a very cheap shot. If Rose had any valid criticism to offer, either of Popper personally (for failing to credit his sources appropriately), or of Popper's view of Darwinism, then he could have expressed these forthrightly and substantively—rather than confounding the two, in the form of a purely offhand remark, as he has. In any case, one would have thought that, before making such an attack, he would have taken the elementary precaution of being factually correct. As it happens, Popper *has* explicitly credited Darwin with the insight that "all observation must be for or against some view", (Popper 1961, p. 259). In fact, Popper gives a detailed citation, which is rather more than Rose does. More pertinently, this acknowledgment appears in the lecture *Evolution and the Tree of Knowledge*, which was discussed in section 5.1.5 above—i.e. one of the primary documents which Rose would necessarily be familiar with *if* he had ever given any serious consideration to Popper's "reservations about the scientific status of Darwinism". In any case, the important point which I wish to make is that innuendo is hardly a substitute for rational criticism.

Consider next, the discussion by Halstead (1980). Halstead is described as a paleontologist, and the proximal target of his criticism seems indeed to be Popper's analysis of (Organismic) Darwinism. However, he does not in fact attempt any direct response to these views of Popper's, but chooses instead to launch a peculiarly speculative and personalised attack on Popper's entire philosophy of science. Popper has provided his own response to this (Popper 1980); Bondi has also independently made similar criticisms of Halstead's position (Bondi 1980). Halstead's primary error is to suppose that Popper has denied scientific status to (Organismic) Darwinism *because* its phenomena are essentially historical. Halstead asserted this particular idea without any definite citation of Popper's work. I am not aware of Popper ever having made such a claim,[25] and, in his own response to Halstead, he stipulates that historical disciplines may indeed be perfectly scientific according to his criteria.

Finally, consider Gribbin (1985, p. 347). He firstly asserts (possibly working from secondary sources) that Popper's comments on (Organismic) Darwin-

---

[25] Halstead's criticism *might* validly have been leveled at Peters, who certainly comes close to denying scientific status to Darwinism on the basis of its historical nature (Peters 1976). However, Halstead does not cite Peters, and, in any case, Peters describes his "criteria for the acceptance of a scientific theory" as deriving from the logical positivists—which, in itself, would automatically redirect any criticism of Peters' view away from Popper.

ism are “half-baked”; but he then offers, as his primary refutation of Popper’s views, the work of Penny *et al.* (1982) on establishing the structure of the phylogenetic tree from protein sequencing data. Gribbin actually describes this as “... the test I like best, which seems utterly to refute both Popper’s original claim and the spurious arguments of those who invoke Popper as an ally in their attacks on evolutionary theorizing”; but, as discussed in section 2.1 above, this work relates entirely to T_t and (therefore) does not bear on T_d at all!

In fact Popper has, so far as I am aware, always accepted the reality of evolutionary descent. Thus, we have, for example, his comment:

> I have always been extremely interested in the theory of evolution, and very ready to accept evolution as a fact.
>
> Popper (1974a, p. 133)

There is thus no reasonable sense in which the work of Penny *et al.* could be interpreted as being contrary to (much less as “utterly refuting”) anything proposed by Popper.

In conclusion, the interesting point in all of this is that, even though Popper was, in fact, mistaken in his analysis of Darwinism, and even though there are good arguments against his views, all three of these writers have chosen instead to attribute to Popper views which he has never actually expressed or defended! It seems to me that such stratagems, whether conscious or otherwise, have only served to further confuse the issues at hand.

References

Adams, Douglas. 1979. *The Hitch Hiker's Guide to the Universe.* London: Pan Books Ltd.

Ayala, Francisco Jose. 1978. The Mechanisms of Evolution. *Scientific American*, **239**(3), 48–61.

Bendall, D. S. (ed). 1983. *Evolution from Molecules to Men.* Cambridge: Cambridge University Press.

Beurton, P. 1981. Organismic Evolution and Subject-Object Dialectics. *Chap. 2, pages 45–60 of:* (Jensen & Harré 1981).

Boden, Margaret A. 1984. Failure is not the Spur. *Pages 305–315 of:* Selfridge, Oliver G., Rissland, Edwina L., & Arbib, Michael A. (eds), *Adaptive Control of Ill Defined Systems.* New York: Plenum Press. Proceedings of the NATO Advanced Research Institute on Adaptive Control of Ill Defined Systems, held June 21–26, 1981, in Moretonhampstead, Devon, England.

Bondi, Hermann. 1980. Letter to the Editor. *New Scientist*, **87**(21 August), 611. This is a response to Halstead (1980).

Burrow, J. W. (ed). 1968. *Charles Darwin: The Origin of Species.* London: Penguin Books. This is a reprinting of (Darwin 1859), together with some ancillary materials, edited with an introduction by Burrow.

Campbell, Donald T. 1960. Blind Variation and Selective Retention in Creative Thought as in Other Knowledge Processes. *Psychological Review*, **67**(6), 380–400.

Campbell, Donald T. 1974. Evolutionary Epistemology. *Pages 413–463 (Book I) of:* (Schilpp 1974).

Darwin, Charles. 1859. *The Origin of Species by Means of Natural Selection.* First edn. London: John Murray. Reprinted as *pages 65–460 of:* (Burrow 1968).

Davern, Cedric I. (ed). 1981a. *Genetics: Readings from Scientific American.* San Francisco: W. H. Freeman.

Davern, Cedric I. 1981b. Principles of Heredity: Introduction. *Pages 6–7 of:* (Davern 1981a).

Dawkins, Richard. 1976. *The Selfish Gene.* Oxford: Oxford University Press. See also Dawkins (1989b).

Dawkins, Richard. 1982. *The Extended Phenotype.* Oxford: W.H. Freeman and Company.

Dawkins, Richard. 1983. Universal Darwinism. *Chap. 20, pages 403–425 of:* (Bendall 1983).

Dawkins, Richard. 1986. *The Blind Watchmaker.* London: Penguin Books.

Dawkins, Richard. 1989a. The Evolution of Evolvability. *Pages 201–220 of:* Langton, Christopher G. (ed), *Artifical Life.* Series: Sante Fe Institute Studies in the Sciences of Complexity, vol. VI. California: Addison-Wesley Publishing Company, Inc. Proceedings of an interdisciplinary workshop on the synthesis and simulation of living systems held September, 1987, in Los Alamos, New Mexico.

Dawkins, Richard. 1989b. *The Selfish Gene.* New edn. Oxford: Oxford University Press. This is a significantly revised edition of (Dawkins 1976), with endnotes and two new chapters added.

Doolittle, W. Ford, & Sapienza, Carmen. 1980. Selfish Genes, the Phenotype Paradigm and Genome Evolution. *Nature*, **284**(17 April), 601–603.

Dunbar, R. I. M. 1982. Adaptation, Fitness and the Evolutionary Tautology. *Chap. 1, pages 9–28 of:* King's College Sociobiology Group (ed), *Current Problems in Sociobiology.* Cambridge: Cambridge University Press.

Eden, Murray. 1967. Inadequacies of Neo-Darwinian Evolution as a Scientific Theory. *Pages 5–19 of:* (Moorhead & Kaplan 1967).

Eldredge, Niles, & Gould, Stephen Jay. 1972. Punctuated Equilibria: An Alternative to Phyletic Gradualism. *Chap. 5, pages 82–115 of:* Schopf, Thomas J. M. (ed), *Models in Paleobiology.* California: Freeman, Cooper & Company.

Fisher, Ronald A. 1958. *The Genetical Theory of Natural Selection.* Second revised edn. New York: Dover Publications, Inc. This is a revised and enlarged edition of the work originally published in 1930 by the Oxford University Press.

Gould, Stephen Jay. 1978. *Ever Since Darwin.* London: Penguin Group.

Gould, Stephen Jay. 1980. *The Panda's Thumb.* London: Penguin Group.

Gould, Stephen Jay. 1983. *Hen's Teeth and Horse's Toes.* Harmondsworth, Middlesex: Penguin Books Ltd.

Gould, Stephen Jay. 1988. *An Urchin in the Storm.* London: Penguin Group.

Gould, Stephen Jay. 1989. *Wonderful Life.* London: Penguin Group.

Gould, Stephen Jay, & Eldredge, Niles. 1977. Punctuated Equilibria: The Tempo and Mode of Evolution Reconsidered. *Paleobiology*, **3**, 115–151.

Gould, Stephen Jay, & Lewontin, R. C. 1979. The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Program. *Proceedings of the Royal Society of London*, **B 205**, 581–598.

Gribbin, John. 1985. *In Search of the Double Helix.* London: Transworld Publishers.

Halstead, Beverly. 1980. Popper: Good Philosophy, Bad Science? *New Scientist*, **87**(17 July), 215–217.

Hardy, Alister. 1965. *The Living Stream.* London: Collins.

Hodge, M. J. S. 1983. The Development of Darwin's General Biological Theorizing. *Chap. 3, pages 43–62 of:* (Bendall 1983).

Hull, David L. 1980. Individuality and Selection. *Annual Review of Ecology and Systematics*, **11**, 311–332.

Jensen, U. J. 1981. Introduction: Preconditions for Evolutionary Thinking. *Pages 1–20 of:* (Jensen & Harré 1981).

Jensen, U. J., & Harré, R. (eds). 1981. *The Philosophy of Evolution.* The Harvester Press.

Lenat, Douglas B. 1983. The Role of Heuristics in Learning by Discovery: Three Case Studies. *Chap. 9, pages 243–306 of:* Michalski, Ryszard S., Carbonell, Jaime G., & Mitchell, Tom M. (eds), *Machine Learning: An Artificial Intelligence Approach.* Palo Alto, California: Tioga Publishing Company.

Lewontin, R. C. 1977. Adaptation. *In: The Encyclopedia Einaudi.* Torino: Giulio Einaudi Edition. As cited by Maynard Smith (1978, p. 55).

Lewontin, Richard C. 1978. Adaptation. *Scientific American*, **239**(3), 157–169. Reprinted as *Chap. 24, pages 252–263 of:* (Davern 1981a).

Little, John. 1980. Evolution: Myth, Metaphysics. or Science? *New Scientist*, **87**(4 September), 708–709.

Magee, Bryan. 1973. *Popper.* London: Fontana Paperbacks.

Maynard Smith, John. 1969. The Status of Neo-Darwinism. *Pages 82–89 of:* Waddington, C. H. (ed), *Toward a Theoretical Biology, 2: Sketches.* Edinburgh: Edinburgh University Press. This paper is also accompanied by various addenda and comments (pages 90–105 of the same work).

Maynard Smith, John. 1978. Optimization Theory in Evolution. *Annual Review of Ecology and Systematics*, **9**, 31–56.

Maynard Smith, John. 1986. Structuralism versus Selection—is Darwinism Enough? *Chap. 4, pages 39–46 of:* (Rose & Appignanesi 1986).

Maynard Smith, John. 1989. *Evolutionary Genetics.* Oxford: Oxford University Press.

McMullin, Barry. 1992 (Mar.). *Essays on Darwinism. 1: Ontological Foundations.* Technical Report `bmcm9201`. School of Electronic Engineering, Dublin City University, Dublin 9, IRELAND.

McMullin, Barry. Forthcoming. *Essays on Darwinism. 3: Genic and Organismic Selection.* Technical Report `bmcm9203`. School of Electronic Engineering, Dublin City University, Dublin 9, IRELAND.

Mendel, Gregor. 1865. Experiments in Plant Hybridization. *Chap. 1, pages 8–17 of:* (Davern 1981a). This is a translation by the Royal Horticultural Society of London, of the original paper, entitled "Vesuche über Pflanzen-Hybriden", published in the *Verb. naturf. Ver. in Brunn, Abandlungen,* iv. 1865, which appeared in 1866.

Moorhead, Paul S., & Kaplan, Martin M. (eds). 1967. *Mathematical Challenges to the Neo-Darwinian Interpretation of Evolution.* Philadelphia: The Wistar Institute Press, for The Wistar Institute of Anatomy and Biology.

Orgel, L. E., & Crick, F. H. C. 1980. Selfish DNA: The Ultimate Parasite. *Nature*, **284**(17 April), 604–607.

Penny, David, Foulds, L. R., & Hendy, M. D. 1982. Testing the Theory of Evolution by Comparing Phylogenetic Trees Constructed from Five

Different Protein Sequences. *Nature*, **297**(20 May), 197–200.

Peters, Robert Henry. 1976. Tautology in Evolution and Biology. *The American Naturalist*, **110**(971), 1–12.

Popper, Karl R. 1959. *The Logic of Scientific Discovery.* London: Hutchinson Education. Tenth impression (revised) February 1980, London: Unwin Hyman Ltd.

Popper, Karl R. 1961. Evolution and the Tree of Knowledge. *Pages 256–280 (Essay 7) of:* (Popper 1972). Based on the Herbert Spencer Lecture, delivered in Oxford on 30 Oct. 1961.

Popper, Karl R. 1963. *Conjectures and Refutations.* London: Routledge & Kegan Paul. Fifth (revised) edn. 1989, London: Routledge.

Popper, Karl R. 1965. Of Clouds and Clocks. *Pages 206–255 (Essay 6) of:* (Popper 1972). This was the second Arthur Holly Compton Memorial Lecture, presented at Washington University on 21 Apr. 1965.

Popper, Karl R. 1972. *Objective Knowledge: An Evolutionary Approach.* Oxford: Oxford University Press. Revised edition (reprinted with corrections and a new appendix 2) 1979.

Popper, Karl R. 1974a. Autobiography of Karl Popper. *Pages 1–181 (Book I) of:* (Schilpp 1974).

Popper, Karl R. 1974b. Replies to my Critics. *Pages 961–1197 (Book II) of:* (Schilpp 1974).

Popper, Karl R. 1978. Natural Selection and the Emergence of Mind. *Dialectica*, **32**(3–4), 339–355.

Popper, Karl R. 1980. Letter to the Editor. *New Scientist*, **87**(21 August), 611. This is a response to Halstead (1980).

Popper, Karl R. 1983. *Realism and the Aim of Science.* London: Hutchinson. From the *Postscript to the Logic of Scientific Discovery*, edited by W.W. Bartley, III.

Provine, William B. 1986. *Sewall Wright and Evolutionary Biology.* Series: Science and its Conceptual Foundations. Chicago: The University of Chicago Press. Series Editor: David L. Hull.

Ridley, Mark. 1981. Who Doubts Evolution? *New Scientist*, **90**(25 June), 830–832.

Rose, Steven. 1986. Introduction. *Chap. 1, pages 1–10 of:* (Rose & Appignanesi 1986).

Rose, Steven, & Appignanesi, Lisa (eds). 1986. *Science and Beyond.* Basil Blackwell.

Rosen, Robert. 1985a. *Anticipatory Systems.* IFSR International Series on Systems Science and Engineering, vol. 1. Oxford: Pergamon Press. Editor-in-Chief: George J. Klir.

Rosen, Robert. 1985b. Organisms as Causal Systems Which Are Not Mechanisms: An Essay into the Nature of Complexity. *Chap. 3, pages 165–203 of:* Rosen, Robert (ed), *Theoretical Biology and Complexity.* Orlando: Academic Press, Inc.

Schilpp, Paul Arthur (ed). 1974. *The Philosophy of Karl Popper.* The Library of Living Philosophers, vol. XIV. Illinois: Open Court.

Stanley, Steven M. 1975. Clades versus Clones in Evolution: Why We Have Sex. *Science*, **190**(24 October), 382–383.

Stebbins, G. Ledyard. 1977. *Processes of Organic Evolution.* Englewood Cliffs, New Jersey: Prentice-Hall, Inc.

Turner, John R. G. 1985. Fisher's Evolutionary Faith and the Challenge of Mimicry. *Pages 159–196 of:* Dawkins, R., & Ridley, M. (eds), *Oxford Surveys in Evolutionary Biology*, vol. 2. Oxford: Oxford University Press.

von Neumann, John. 1966. *Theory of Self-Reproducing Automata.* Urbana: University of Illinois Press. Edited and completed by Arthur W. Burks.

Waddington, C. H. 1967. The Principle of Archetypes in Evolution. *Pages 113–115 of:* (Moorhead & Kaplan 1967).

Wallace, A. R. 1866. Letter to Charles Darwin dated 2 July. *Pages 170–174 of:* Marchant, J. (ed), *Alfred Russel Wallace Letters and Reminiscences*, vol. I. London: Cassell. Collection published in 1916. As cited by Dawkins (1982, p. 280).

Williams, George C. 1975. *Sex and Evolution.* Princeton: Princeton University Press.

Williams, Mary B. 1970. Deducing the Consequences of Evolution: A Mathematical Model. *Journal of Theoretical Biology*, **29**, 343–385.

Wills, Christopher. 1991. *The Wisdom of the Genes.* Oxford: Oxford University Press.

Wright, Sewall. 1980. Genic and Organismic Selection. *Evolution*, **34**(5), 825–843.

Wright, Sewall. 1982. The Shifting Balance Theory and Macroevolution. *Annual Review of Genetics*, **16**, 1–19.