Gary Hatfield probes another question at the loci of the traditional mind-body problem and the analysis of mental representation, though he is concerned to answer a question from philosophy of science. He wants to know what the direction of constraint is between psychology and neuroscience: What sort of theory or evidence trumps what, when does it, and why? His analysis helps to answer some of the considerations that Polger and Flanagan raise as well as to extend their analysis of visual perception.

A second way that considerations from biology and psychology can inform philosophy of science appears in Steve Downes's and Mark Bedau's chapters. Downes focuses on the debate concerning the putative parallel between scientific revolution and conceptual change within individuals. He uses the biological ideas of ontogeny and phylogeny to illuminate the psychological dimensions of the issue. Bedau also relies on biology in his explication of the notoriously problematic ceteris paribus laws in psychology. He uses recent models from the artificial-life community to help codify his notion of "supple" laws, which appear in both biology and psychology as well as in science in general.

Finally, though there are numerous parallels between debates in philosophy of biology and debates in philosophy of psychology, I have selected only three to illustrate this approach. Each enriches our current understanding of some important terms and trends in both biology and psychology.

Peter Godfrey-Smith uses discussions of mental representation to shed light on what it is that genes do. Do they transmit information or are they merely causal? How far into the environment does genetic information go? He speculates about how to approach answering these questions.

El-Hañia and Pereira look at biological causation through the lens of supervenience and emergence debates in philosophy of psychology. Like Godfrey-Smith, they try to bring advances in philosophy of mind to philosophy of biology. At the same time, they use considerations from biology to help solve the long-standing paradox of mental causation.

Rob Wilson explores the connections between biology and psychology in comparing Dawkins's extended phenotype with nonindividualistic or "wide" systems. He also explores the metaphors of causal powers and encoding and the notion of an individual as a locus of control.

Where biology meets psychology we find overlapping domains of study, larger philosophical implications, and even larger conceptual ties. Woven through these connections are shared concerns regarding the status of semantics, scientific law, evolution and adaptation, and cognition in general. One book can depict only a few examples of each of these connections and themes. But it is a beginning nonetheless.

I FUNCTIONS AND TELEOLOGY

1 Fitness and the Fate of Unicorns

Karen Neander

1 Is Teleosemantics Adaptationist?

Is teleosemantics adaptationist? People sometimes tell me so. But my own answer is that it all depends on what you count as adaptationist and on which version of teleosemantics you consider. Certainly, the general teleosemantic enterprise is not doomed to fail because of any commitments to dubious empirical assumptions of an adaptationst kind. However, I don't recommend that concerns about implausible adaptational assumptions be dismissed out of hand. There are particular versions of teleosemantics that have implausible adaptationist assumptions, and we need to be clearer about just what the adaptational commitments of teleosemantic theories are.[1]

The teleosemantic theories that I am interested in here are those in a class of theory in psychosemantics, as opposed to semantics proper. That is, they are an attempt to account for (what is often called) *original meaning*. A representation counts as having original meaning if it has meaning that does not depend on its (the representation's) being metarepresented by some further representation (e.g., in the intentions of someone using it to communicate with someone else). In this chapter I'll be assuming that original meaning is possessed by some brain states of evolved creatures such as ourselves. The open question is whether teleosemantic theories, which attempt to derive original meaning from a creature's evolutionary history, are bound to fail because they require implausible adaptationist assumptions.

In brief, teleosemantic theories attempt to move from historical facts about what some trait was selected for by natural selection to constitutive facts about what some (possibly other) trait (or state of a trait) represents. The attempted move is made via the teleological or teleonomic notion of a *function*, a notion that is involved in our, in some sense, normative talk about a biological system functioning properly or malfunctioning—a notion that entails a distinction between what a system is disposed to do and what it is supposed to do. Standardly, in teleosemantic theories, this notion of a proper function is understood by means of an etiological analysis, according to which the proper function of a trait is what it (i.e., that homologous type) was selected for by natural selection. And according to teleosemantic theories, the psychosemantic norms of original meaning are derived from these norms of proper functioning, although there is little consensus about how exactly this should be done. The extant theories are quite varied and there is scope for more variation still. For instance, extant theories tend to be versions of causal or informational theories, but teleonomy could also be used to add normativity to a conceptual role theory of mental content. All that teleosemantic theories need have in common is the idea that psychosemantic norms ultimately derive from the norms of teleonomic functions.

Given this, it is however true that all teleosemantic theories appeal to adaptational explanations, since these are explanations in terms of what something was selected for. So to this extent it is clearly appropriate that concerns about adaptationism arise. But not just any appeal to adaptational explanations counts as adaptationist, let alone as adaptationist in the pejorative sense of the term. There is a real problem for teleosemantics only if it too readily assumes that the adaptational explanations it requires are available. The open question is whether the most plausible teleosemantic theory or theories must do so.

Of course, teleosemantic theories are not in the business of providing such adaptational explanations themselves, for that's a task for the evolutionary biologist and not for the philosopher. Instead, the aim of teleosemantic theories is to explain how original meaning can in principle be derived from the natural facts, and in particular from those adaptational explanations that are true, whichever ones those are. If teleosemantics is doomed to fail because it has implausible adaptationist assumptions, then it must be the case that, whichever ones those are, they are insufficient as a basis for teleosemantics. I know of no good or even half good argument to this effect. At best, there are some good arguments with much more specific conclusions, to the effect that some particular teleosemantic theory (at least on the critic's construal) makes some unwarranted or even impossible adaptationist assumptions. At worst, there is a great deal of confusion and ignorance.

What follows is in three main sections. Section 2 reviews the adaptationism debate and explains that the biologists who are known as anti-adaptationists never seriously meant to question the central importance of adaptational explanations. We can accept what the anti-adaptationists say (about constraints on selection, and so on) and yet still believe that natural teleology is pervasive in modern biology. Section 3 then explains why the modern version of the argument from design, namely the argument for selection, holds good for complex organized systems like the brain. No alternative kind of explanation is sufficient to explain the accumulation of coadapted traits that is required for this sort of organized complexity. I claim that we know enough to know that there is a correct naturalistic yet normative neurophysiological description of the brain's representational functions. Whether this is enough for teleosemantics then depends on the details. Rather than trying to survey extant teleosemantic theories, section 4 addresses some more detailed concerns that have been raised in the literature. To do so, it looks at how at least one (fabricated pseudo-Dretskean) teleosemantic theory has commitments to implausible (and even impossible) adaptational explanations. The question is, however, whether it is the teleonomy or some other aspect of the theory (its atomism or nativism, for instance) that should be dropped.

2 The Adaptationism Debate

Before we can say whether teleosemantics is unduly adaptationist, we need some clarification of the latter term. Not just any appeal to adaptationist explanations deserves to be called "adaptationist." Adaptationism is an ill-defined thesis about adaptational explanations, but however it is defined it is clear that both the so-called adaptationists and the so-called anti-adaptationists agree that some adaptational explanations are perfectly legitimate.

An *adaptational explanation* is a kind of causal explanation, one that explains a trait by appeal to the fact that past traits of that type had adaptive effects which caused the phenotype (or its underlying genotype) to be selected by natural selection. So, for example, the gloss on the leaves of the eucalyptus tree can probably be explained as an adaptation for reducing dehydration in a hot, dry climate. If so, the gloss was responsible for less water loss, which in turn was responsible for the underlying genes being replicated more than they would otherwise have been, which in turn was responsible for the trait being preserved in the population. These adaptational explanations are entirely respectable stock in trade of modern biology.

What is the adaptationist debate about, then, aside from a lot of breast beating? It is often said that it is about the relative importance of natural selection as compared to other evolutionary forces or factors in determining and explaining evolutionary outcomes. For reasons that should become clear by the end of this section, it is more accurate to say that it is about the relative importance of adaptational explanations as compared to other explanations in explaining evolutionary outcomes. Someone is accused of being an adaptationist if she is seen as ignoring or giving too little attention to other kinds of explanations. The accusers are the so-called anti-adaptationists. At first, the adaptationist label was virtually by definition pejorative, but now, as the rhetorical tide turns, we find some people describing themselves as adaptationists, presumably because they think that adaptational explanations have begun to be underrated. So put, the debate must seem a very vague one, and to a large extent it is. The main problem with this debate is that everyone is really somewhere in the middle ground, but almost no-one is calmly acknowledging that fact and trying to clarify the less radical differences that really exist.

For instance, Dan Dennett (1995) has signed on as an adaptationist, stoutly declaring in doing so that adaptational explanations have a central and essential role in modern biology. I am with him there, one hundred percent. Except that if this is enough to make one an adaptationist, then even the so-called anti-adaptationists qualify. Beyond their surface rhetoric, they've never really denied that adaptational explanations belong on center stage; they've only insisted that they shouldn't hog the limelight. They take themselves to be addressing people who too frequently talk as if every trait has an adaptational explanation and as if adaptational explanations are the only kind of explanation to be considered in explaining evolutionary outcomes. Perhaps it is understandable that in such a context the anti-adaptationists do not stress the importance of adaptational explanations. But even if they are at fault here, the charitable interpretation is not that they think them unimportant, but rather that they think their appreciation of their importance will be taken for granted. At worst, the so-called anti-adaptationist have probably been guilty of being carried away at times by their rhetoric, of making their position sound more radical and original than it really is. In any case, to my knowledge, no so-called anti-adaptationist has ever seriously tried to argue that adaptational explanations do not have a central and essential role in modern biology.

Lately we have seen several explicit denials of the radical reading (or misreading) of their position. For instance, Niles Eldredge tells us in no uncertain terms that:

> no rational evolutionary biologist feels that most change is not adaptive, or that adaptive change is not caused by natural selection. (1995, p. 56)

And Stephen Jay Gould (1997), in a review of Daniel Dennett (1996), says, in the weary tone of one forced to state the all too obvious:

> may I state for the record that I (along with all other Darwinian pluralists) do not deny the existence and central importance of adaptation, or the production of adaptation by natural selection. Yes, eyes are for seeing and feet are for moving. And yes, again, I know of no scientific mechanism other than natural selection with the proven power to build structures of such eminently workable design. (p. 35)

So, what then is it that distinguishes the adaptationists from the anti-adaptationists? Gould chooses to call himself a Darwinian pluralist, rather than an anti-adaptationist. This might seem a better name for the so-called anti-adaptationists, since it does not suggest that they deny the importance of adaptational explanations. However it does have the unfortunate implication that their opponents, the adaptationists, deny the importance of other forms of explanation, and this is just as misleading in a different direction.

One target of anti-adaptationist criticism is Richard Dawkins (see especially his 1986). He has been an influential advocate of the power of natural selection to create the wondrous and intricate designs for which God previously took all the credit. But this in itself does not make him an adaptationist. As we saw in the quotes from Gould and Eldredge, they will wearily agree that natural selection and natural selection alone has this power. But nor does it make Dawkins a Darwinian monist. He will, I am sure (and probably just as wearily) allow that, as every second year biology student should know, there are constraints on the adaptive power of natural selection. I am quite sure he knows that natural selection must tinker with the available materials, that there are spandrels, that there is allopatric speciation with relatively rapid change due to migration, reproductive isolation, and the increased power of drift (see, for example, his 1986, chapter 9) and so forth. Dawkins is a Darwinian pluralist too, if that just means that he acknowledges other mechanisms besides natural selection and that he appreciates that there are constraints on the adaptive power of natural selection.

Dawkins, like Gould, addresses a popular audience, but he is more concerned to correct a different way in which his popular audience can and does fall into error. More than fifty percent of university entry-level biology students in America are creationists. Among other things, Dawkins is trying to persuade them and their neighbors of the power of natural selection, to convince them of the fact that natural selection can do the job. Again, perhaps it is understandable that in this context he does not emphasize the limitations on the adaptive power of natural selection (although it is true that evidence of such limitations—evidence in the form of an imperfect design product—is also good evidence against creationism).

Fodor complains of Dennett that he concedes too much to the anti-adaptationists to be an adaptationist in good standing, for Dennett concedes that the power of natural selection is seriously constrained in various ways. "Good adaptationist thinking is always on the lookout for hidden constraints, and in fact is the best method for uncovering them," he quotes Dennett (1995, p. 261) as saying. Fodor objects, "That makes it look as though there is practically nothing that an adaptationist in good standing is required to believe about how evolution works; he's only required to buy into a methodological claim about how best to find out how it does" (Fodor, 1996, p. 248). But Fodor's complaint is inappropriate. We may as well complain of Eldredge and Gould that they concede too much to the adaptationists to be anti-adaptationists in good standing. Either way, it's a silly exercise. As far as I can see, no one in the debate has seriously denied that adaptational explanations are of central importance *or* that the adaptive power of natural selection is seriously constrained, leaving aside any fictional caricatures created to lampoon one's enemies.[2] What real differences there are between adaptationists and the so-called anti-adaptationists concern *degrees* of importance. The debate is over how important certain influences on evolutionary outcomes are. These differences are hard to quantify precisely and the more visible side of the debate is far removed from any genuine attempt to do so.

I am not trying to suggest that there are no interesting and substantial issues in the vicinity of the adaptationism debate. But they are not, I think, what people have in mind when they accuse teleosemantics of dubious adaptationist assumptions. For instance, John Maynard-Smith (1978) believes that optimization theory is useful and he might be described as an adaptationist because of this. In optimization theory, an evolutionary trend with respect to some trait is modeled as if the adaptive power of natural selection were unconstrained. That is, it is modeled as if drift were absent, as if the relevant traits can be selected for (or against) independently, and so on. Maynard-Smith argues that this style of modeling is useful, not because nature is this way, but because deviations from the predictions provided by the model give us some measure of the effect of factors other than natural selection. So-called adaptationists and so-called anti-adaptationists can disagree about the usefulness of such models and the disagreement sometimes seems to depend on beliefs about the degree to which the model can be treated as an accurate picture of reality (even though both sides agree that it is far from exact). Whether or not such models are useful is an interesting issue on which sensible biologists can disagree, but it is hard to see exactly how it concerns teleosemantics. No one has yet tried to argue that it does, so I'll assume that this methodological issue is not what we are concerned with here.

The critics of teleosemantics seem to have a cruder notion of adaptationism in mind. Their adaptationist is the one targeted by Stephen Jay Gould and Richard Lewontin (1979) in their classic spandrels paper. This adaptationist is supposed to believe three things: (1) that virtually every trait that we can identify has an independent adaptational explanation; (2) that the form of these traits is perfectly fashioned for their function; and (3) that in modeling, explaining and predicting evolutionary outcomes, we can, in practice, ignore everything but natural selection, because natural selection is powerful enough, at least in the long run, to swamp the effects of other forces. For such an adaptationist, natural selection has more or less slipped right into the role vacated by God in guiding the design of living creatures. Although it lacks foresight and is therefore far from omniscient, it is pretty well omnipotent and omnipresent. This is the position Gould and Lewontin label 'Panglossian'. It is an extreme that no real biologist has avowed, although we might accuse some of too nearly approaching it in practice. The Panglossian sin is to err too far in this direction.

3 Teleonomy and Panglossian Sin

Does teleosemantics commit the Panglossian sin? This question will take a while to answer, but we can make a start on it in this section by noting that it does not do so merely by being committed to natural teleology. In this case, natural teleology, or teleonomy, is a teleological notion of proper functioning that can be cashed out in naturalistic terms, and more specifically, in terms of natural selection (as opposed to the purposeful intentions of a designer). Many readers will think it too obvious to need saying, but judging from what I have heard in discussions, it does need saying: natural teleology does require some adaptational explanations but it is not in itself Panglossian. To think otherwise is simply confused.

The standard means of introducing the idea of natural teleology is by the now familiar distinction between selection *of* and selection *for*. A trait with two properties *P* and *Q* counts as having been selected for *P* and not for *Q* if its being *P* contributed causally to its being selected and if its being *Q* did not contribute causally to its being selected, either because *Q* was neutral or because *Q* worked against selection of the trait.[3] Along the same lines, a gene sequence that's responsible for two traits, *R* and *S*, counts as having been selected for *R* and not for *S* if its being responsible for *R* contributed causally to its being selected whereas its being responsible for *S* did not contribute causally to its being selected, either because *S* was neutral or because *S* worked against selection of the gene sequence. If the gene sequence is selected, there is selection of *S*, even though there is no selection for *S*, only selection for *R*. In such circumstances, *S* is said to be a piggy-back trait because it rides on the back of some other trait that is adaptive and which does the work involved in getting them selected. Some piggy-back traits have been nicknamed "spandrels," for like the spandrels of cathedrals they are the inevitable architectural outcomes of selection for other traits (e.g., the human chin is the outcome of selection for the jaw). Other piggy-back traits are the result of pleitropy (where two potentially separate phenotypic traits are caused by the same gene sequence) and gene-linkage (where two potentially separate phenotypic traits are caused by gene sequences that are close together on the same chromosome). The possibility of piggy-back traits is thus standardly recognized in introductions to teleonomy for the purpose of explaining teleosemantic theories.

As it is normally understood in teleosemantic theories, a (or the) teleonomic function of a trait (type) is what it did that it was selected for.[4] This obviously does not imply that everything has an adaptational explanation, or even that most things do, or even, strictly speaking, that anything does. It entails that if something has such a teleonomic function then it has an adaptational explanation, an explanation in terms of what it was selected for.

Further, there is no implication from this definition that if something has a teleonomic function, and therefore has an adaptational explanation, then it is perfectly fitted for its function. Nor is it implied that its adaptational explanation is its complete evolutionary explanation. The Panda's "thumb" was selected for stripping leaves off bamboo, and that therefore is its teleonomic function. But it does not follow that the Panda has the best bamboo leaf-stripper it could possibly have, or that natural selection worked toward an adaptive outcome in this case in the absence of all constraints. Far from it (Gould 1980). While the Panda's thumb has a clear teleonomic function, it is a wonderful illustration of the fact that natural selection is a tinkerer and a satisficer, heavily constrained by the past and by the alternatives that are presently available. The Panda's thumb—in fact, an elongated wrist bone—is an imperfect thing from a design engineering point of view. To explain its imperfections we need to consider the limitations on the adaptive power of natural selection. An explanation in terms of these limitations is entirely compatible with and complimentary to the adaptational explanation. There is nothing remotely Panglossian about this combination of explanations and teleonomy is in no way committed to the Panglossian error.

As long as there is variation, there can be selection for one variation over another. Even if there is just one barely workable variation available, one workable variation among others that are worse than useless, that one workable variation can be selected for working as well as it does. Selection for a trait can occur as long as that trait increases fitness relative to alternative alleles, and that is consistent with the fitter trait falling well short of perfection. There is no incompatibility in the following two theses: that most traits have teleonomic functions and that the adaptive power of natural selection is severely constrained.

Actually, it is worth being clear at this point just in what sense the adaptive power of natural selection is constrained. People often speak of the various so-called constraints as constraints on natural selection, but they are really aspects of natural selection, and they are not sensibly viewed as opposing forces, or as other evolutionary mechanisms. I am speaking here of what are often referred to as developmental, phylogenetic and architectural constraints.

Consider the so-called developmental constraints first. These concern the fact that mutations that alter processes early on in embryonic development are less frequently beneficial than those that alter processes that occur later in development. This is because random changes to complex organized systems are more likely to be damaging if they are large and widespread than if they are small and localized, and the earlier the developmental process is, the larger and more widespread the ramifications of any changes to it are likely to be. But it can be misleading to see this as a constraint on natural selection, unless one also understands that this is itself an aspect of natural selection. When a mutation is selected against because it causes widespread changes that are on balance disadvantageous, this is just part of the selection process. Natural selection is just doing its thing, selecting against an unfit allele. We describe this aspect of selection as a constraint, I suppose, because it limits the kind of changes that natural selection can, given its nature, bring about. These facts about development explain why it is hard for natural selection to alter the fundamental design of a system, why it tends to be somewhat conservative in its designs, and why its designs will often be imperfect from an engineering point of view, because it cannot start from scratch when new design demands arise. But the conservative maintenance of early developmental pathways itself has an adaptational explanation. Or, to put it another way, to explain something as due to developmental constraints is to give a special form of adaptational explanation. Perhaps this is not what people generally have in mind when they think of adaptational explanations. But it is a matter of the maintenance of already established adaptation, especially the internal, delicate and massively complicated coadaptation of traits required for the functioning organism.

The so-called phylogenetic constraints concern macroevolutionary outcomes and explain why natural selection has ignored large areas in the space of apparent design possibilities. Natural selection cannot explore all possible kinds of workable designs because it has to get from where it is to where it is going via pathways consisting of actual functional creatures, choosing the most adaptive alternatives at each step along the way. As it happens there are lots of some kinds of creatures (mammals, for example) and none of some other possible kinds of creatures (I'll leave them to the reader's imagination) and so we get modifications on existing kinds and not modifications on kinds that don't exist. But once again, that's just the way natural selection works, and there is no separate force involved. The phylogenetic constraints are not like a dam holding back a river's natural flow. They are instead like deep river beds, eroded over vast time, created by as well as channeling the river's natural flow.

The so-called architectural constraints do not even limit design optimization in the way that the other two so-called constraints might be claimed to do, since architectural constraints just are elements of good design. Just as large office buildings need strong skeletal support and good circulation, so does something the size of an elephant, and that's one reason mosquitoes will never grow to an elephantine size. While a human architect can design unworkable buildings despite these architectural constraints, what is unworkable cannot be selected by natural selection. So I suppose we describe the architectural constraints as constraints because they limit possible designs, even though they do not limit optimization. But once again, of course, this is just the way natural selection works.

None of these so-called constraints on natural selection is an outside force, separate from or external to natural selection. It is senseless to ask whether they are more important than natural selection in determining evolutionary outcomes, since they are themselves aspects of natural selection. That there are such constraints is reason not to be a Panglossian, but it is no reason at all to think that adaptational explanations are not pervasive. Maybe one needs to be a little in the grip of the Panglossian perspective to even view these so-called constraints as constraints. They prevent natural selection from being an omnipotent designer (from being able to start from scratch, summon materials from thin air, explore all possible workable designs, and even unworkable ones). But to think that natural selection would otherwise be an omnipotent designer is definitely to err too far in the Panglossian direction. In the absence of these so-called constraints, natural selection would not be natural selection. It would be God.

There are mechanisms of change involved in evolution that can be counted as other than or external to selection itself. Mutation, migration, geographic isolation and drift are perhaps the most obvious ones. Of course, natural selection can only select from the available variation and the variation available is randomly generated by mutations and chromosomal cross-overs, so what occurs in this way both provides the raw materials for selection and limits its potential: it is both empowering and constraining. It limits optimization of design for the obvious reason that unless the optimal allele arises it can’t be selected. But once again, if this were not so, we would not be talking about natural selection. And while the dependence of natural selection on the availability of mutations is a good reason not to be Panglossian it is no reason at all to think that adaptational explanations are not pervasive. If we want to explain why no mammal has green fur, it might have something to do with the fact that some mutation that could have occurred has not occurred (I don’t know). But suppose that, in a given case, green fur would have been selected had the alternative been available. And suppose that, given that it was not available, brown fur was instead selected because it provided better camouflage than the alternatives that were available (orange, black and white stripes, etc.). It remains true that the brown fur was selected for camouflage. Teleonomy does not require perfection. It is enough that there was selection for a trait over its actual competition.

Migration, geographic isolation and drift are thought by many biologists to be of great importance in speciation. Drift is usually thought to include two components (i) deviation from the statistical norm in the random fertilization of gametes, and (ii) inequalities in parent sampling due to accident rather than variations in fitness.[5] Drift tends to tend in a different direction to natural selection because there are far more damaging mutations than beneficial ones, and while natural selection tends to favor the beneficial ones, drift is indiscriminate. Drift therefore competes with selection. It can also play a decisive role in determining the direction of evolution by eliminating certain variations when their representation is small, as is the case of new mutations, and in allopatric speciation where a small part of population separates from the rest. It can tend toward conservatism, as it probably does when it eliminates new mutations, but it can also tend toward change and phylogenetic diversity, as it probably does in allopatric speciation. When the overall numbers are small, relatively small fluctuations can eliminate a gene sequence or drive it to fixation. But just because drift (et al.) were decisive in determining an evolutionary trajectory does not mean that natural selection was not similarly decisive. As long as drift leaves more than one variation available for selection, then selection can still occur. And, to use an oft used metaphor, while drift (et al.) can force the train onto a different track it cannot drive the engine.

In brief, even if we assume that adaptational explanations are in principle available for most morphological change, there is nothing in the least Panglossian about such an assumption. Let us now turn to the somewhat more specialized question of whether the same holds true when we are considering the evolution of perceptual and cognitive capacities, for this is where we are concerned with specifically representational functions.

4 Teleosemantics

Yes, eyes are for seeing . . . (yawn). The eye is almost the paradigm case of intricate, organized complexity in need of an adaptational explanation. Yet it is a relatively simple thing compared to the visual cortex that processes its input. The human brain is sometimes said to have the highest degree of organized complexity of anything in the known universe. Perhaps so. It is anyway, of all known things, one of the thing most in need of an adaptational explanation. And teleosemantics could hardly be based on a more solid empirical assumption insofar as it assumes that there is in principle available a rich adaptational explanation of the representational capacities of the brain. I'll say why in this section, starting with a few thoughts about teleosemantics by way of a little motivation and elaboration.

A central part of the puzzle about original meaning is how it is possible to move from ordinary descriptive facts about cognitive systems to psychosemantic norms (for correctness of representation, truth of beliefs, and so on). In general, if something is a representation, then it can, in principle, *mis*represent. So to know the content of a representation is to know something about which circumstances would count as correct applications and which would count as incorrect applications of the representation. Traditionally, this move from the natural to the normative has seemed daunting and one major attraction of teleonomy as a basis from which to develop a naturalized theory of meaning is that teleonomy is already both natural and normative. Moreover, there is a significant sense in which both talk of natural functions and talk of original meaning are normative in the same sense. To attribute a natural function or to attribute original meaning to something is to attribute a certain kind of normative property to the thing. That is, it is to attribute an evaluative standard to it that it could fail to meet, even chronically (i.e., systematically and persistently and even under ideal external circumstances). I call it an evaluative standard since intuitively it is a standard that the thing is in some sense supposed to meet. In the case of natural functions we are speaking of a standard of proper functioning and in the case of original meaning we are speaking of a standard of correct representation. According to an etiological theory of functions, the intuitive sense in which a trait is supposed to perform its function is cashed out naturalistically in terms of what the trait was selected for. While these biological norms of proper functioning cannot be simply equated with psychosemantic norms, the hope is that they can be deployed to determine what states of the brain are supposed to represent.

Philosophers have for far too long now spoken as if our puzzle is how to understand the relation between the intentional properties of the mind, on the one hand, and the nonintentional physical, chemical and biological properties of the brain, on the other hand. Some philosophers even recommend the elimination of intentional properties in favor of making do with neurophysiology. But neurophysiology is already thoroughly intentional, at least in the broader sense, in which something counts as intentional if it is representational. Once we are above the level of microbiology, neurophysiology is steeped in descriptions of the representational functions of neural states and processes. Physiology is the relating of form to function, and the function of the brain is largely representational, so it is hard to even make sense of the idea of a complete but nonrepresentational neurophysiology. The brain is a biological organ selected by natural selection for various representational functions; for representing variable states of a creature's environment, for representing the creature's own body and its place in this environment, for processing information about these things, and for mediating between them and the creature's behavior.[6]

My hunch is that this natural and normative description of the brain's representational functions will form the foundation for a future psychosemantic theory adequate for the purposes of a mature cognitive science. In fact, my hunch is that this will happen and is already beginning to happen whether we philosophers approve of it or not. But that is more than I an attempting to establish in this chapter. My goal here is simply to defend teleosemantics from those who argue that it is doomed because it is committed to dubious adaptationist assumptions.

We have already seen that a commitment to teleonomy, and even a commitment to pervasive teleonomy, are not at all unduly adaptationist (or Panglossian). However, most teleosemantic theories require something more specific; that there be a fairly rich teleonomic description of natural representational systems, as such.[7] That is, they require that there be in principle available a fairly rich adaptational explanation of cognition and perception. Do we have good reason to believe that such an adaptational explanation is in principle available, given that we know little about the evolution of the brain or even (relatively speaking) about how the brain represents?[8]

We do. We have very good reason to believe that adaptational explanations are important in explaining cognition and perception. One reason we understand so little about how brains represent is the sheer complexity of brains. From this, and from what we do know about the brain functioning (and malfunctioning), and from our general knowledge of the mechanisms of evolution we can infer with moral certainty that there is in principle available a rich adaptational explanation of perception and cognition (or rather, a rich complex of such explanations). I am appealing here to the modern version of the argument from design; an argument that I'll call the argument for selection. I am of course making no claims to originality in offering such an argument. This argument is as old as Charles Darwin's appreciation of William Paley.[9] But I think that some philosophers have forgotten just how powerful this argument is.

By far the best explanation available for the existence of organized complexity—that is, of a system with heterogeneous parts that are harmoniously coordinated toward the production of some complexly achieved overall activity—is an adaptational explanation. William Paley (1802) and others used what has been known as the argument from design to support the existence of a Divine Designer on the basis of an analogous appeal to the best explanation. (It would have been more appropriate to call it the argument *for* design since it was an argument for design from the existence of organized complexity.) There is disagreement about just how powerful Paley's argument was in the absence of the alternative Darwinian explanation. Was God the best explanation available at the time? Contemporary philosophers and historians disagree. It can be argued that the appeal to a Divine Designer only shifted the explanatory burden, or that even the mechanistic explanations that were available at the time were more plausible, at least in the light of the scientific knowledge of the time. But none of these worries about the power of Paley's argument in the least trouble its descendent, the argument for selection. It is an immensely powerful argument and no new developments in biology, anti-adaptationist or otherwise, have done anything to challenge this fact.

Let me make it clear that I am not maintaining that organized complexity of the kind that brains have is strictly nomologically impossible without selection for it, but I am saying that there is no other available explanation that is plausible and that makes organized complexity of this magnitude anything other than extremely improbable, improbable almost to the point of impossible. The probability that adaptational explanations are not important in explaining cognition and perception is so tiny it can be ignored for all practical purposes.

To properly understand this point, we should note that there are two notions of "complexity" in currency these days. One is a simple notion of complexity as mere heterogeneity (e.g., see Peter Godfrey-Smith 1996). On this notion, the more the parts of a system vary, and the more various the kinds of things they do, the more complex the system is. A high degree of complexity, in this sense of the term, is compatible with a high degree of *dis*organization. In contrast, organized complexity, requires organization as well as heterogeneity. The parts are described as communicating, coordinating, and cooperating with each other. In organized complexity, the order of the parts and their interactions as well as their variety matters. And it is this latter notion of organized complexity that is relevant to the argument for selection. Mere heterogeneity can easily enough result from nonselectional processes (a tornado, for example).

The adaptationism debate has served to highlight the importance of factors other than natural selection in accounting for evolutionary outcomes: mutation, drift, migration, pleitropy, genetic linkage, heterozygous advantage, and so on. But these other phenomena are none of them candidates for explaining organized complexity. The first three tend away from adaptation, and hence do not bias change toward the accumulation of coadapted traits. The next three just prevent independent selection of traits, and only act to preserve the present or hinder the future coadaptation of traits.

Mutations are arbitrary, not in the sense that a mutation is as likely to occur as not, or that all mutations are equally likely to occur, but in the sense that whether a mutation occurs is insensitive to whether its occurrence is adaptive. Since most variation produced at random in a complex organized system will tend to disrupt its organization rather than improve it, mutation without selection tends away from, not toward, the accumulation of coadapted traits. On its own, it is a force for change, but it moves things in the wrong direction. Complex systems that cease to be positively selected for, such as the eyes of fish that take up residence in dark caves, tend to lose their organized complexity. Some people, not well educated in biology, can be overly impressed by the fact that a mutation or two sometimes results in profound phenotypic changes. They are tempted to think that saltation, where new complex organization arises suddenly without cumulative selection, is a common enough phenomenon. They might be impressed by the fact, for example, that two point mutations (*bithorax* and *postbithorax*) in *drosophila* can produce two whole extra wings, turning a two-winged fly into a four-winged fly. But this is not an example of saltation; the "design-work" has already been done by natural selection in cases like this. The genes responsible for the new wings are the ones that are responsible for the old wings, and they are the result of extensive cumulative selection for flight. What has been altered by the two point mutations is just some positional information controlling the expression of the relevant gene sequence.[10]

Drift is also random and for the same reason it also tends to favor disorganization, not organization. That is why it can be decisive in determining the evolutionary trajectory and yet cannot power the accumulation of coadapted traits. Drift increases the probability of complexity only in the sense of mere heterogeneity. It is, for example, heavily implicated in the variation in junk DNA (DNA that is either not transcribed or not translated into amino acids). But this is so precisely because junk DNA has no phenotypic outcome and hence no effect on the physiology and complex organization of the organism.

Anyone who denies that we can be confident that adaptational explanations will be important in explaining cognition and perception owes us a response to the argument for selection. These phenomena are the products of a system of immense organized complexity. Cumulative selection is the only thing that can in practice explain such complexity. And while this fact seems to have become blurred behind all the rhetoric over recent radical, or anyway supposedly radical, revisions to (neo-) Darwinian biology, this fact remains.

There is the in-principle possibility that a complex system could be selected for one function and then fortuitously be used to do something else instead or in addition. Such things happen. To borrow an example from Fred Dretske, a scale for measuring weights can, using fixed weights, be used for measuring altitude instead. So it doesn't *follow* from the fact that cognition is performed by a complex system that the complex system was selected *for cognition*, and therefore it doesn't follow that the adaptational explanation of the brain is one that will bestow representational functions upon it. However, while the in-principle point is certainly correct, its application is absurdly far-fetched in this case. For example, what else might the visual cortex have been selected for that just happened to have the right complex structure for visual perception? The ancients believed that brains were an organic radiator with the function of dissipating heat. But even if this were a function of the brain, its selection for the dissipation of heat could only explain at most a few of its features: for example, the large surface to volume ratio and the high concentration of capillaries near the surface. It would not begin to explain, for example, why the input to the retina was mapped on to V1 and from there on to V2 and V3, and so on and so forth. The suggestion that brains were not selected for representing variable features of the environment and for processing information about them is about as plausible as the suggestion that a car's radiator was really designed to be a portable navigational computer. In the second case, the radiator does not have the kind of structural complexity required for it to be the plausible outcome of such a process, and in the first case, the proposed process could not plausibly explain the outcome in need of explanation, namely the specific structural complexity of the brain.

There's no room for reasonable doubt that there is a rich and detailed adaptational story to be told about the organized complexity responsible for the representational capacities of our brains. Of course, not every mental trait will have an adaptational explanation. But the presence of spandrels, piggy-back traits, vestigial traits, and so on, does not prevent teleonomy from setting standards of proper functioning in ordinary somatic physiology. So we need a special argument to show that the presence of such things in psychological systems prevents teleonomy from setting standards of correct representation for cognitive science, specifically with respect to the representational capacities that are our natural endowment. No good argument so far has been forthcoming.

The general teleosemantic enterprise does have some empirical assumptions, but insofar as it just requires that there be a rich and detailed adaptational explanation of our innate representational capacities, it is on about as firm a footing as it can possibly be.

5 Specific Teleosemantic Theories

It is, however, easy to think up particular versions of teleosemantics that require utterly implausible adaptationist assumptions. I'll describe one in a moment.

The general idea of teleosemantics is very abstract. As I've already mentioned, all teleosemantic theories attempt to derive content from the notion of a (teleonomic) function, but this can be attempted in many different ways, and there is at present little consensus concerning how it should be done. I will not attempt to survey the variety of possible teleosemantic theories here. What I want to do is address some objections that have been made to specific versions of teleosemantics. However, I have no wish to defend these particular versions of teleosemantics. I want to avoid getting bogged down in exegetical details altogether by fabricating a simple pseudo-Dretskean teleosemantic theory. The point of doing so is to acknowledge that a version of teleosemantics can have unduly adaptationist assumptions while at the same time illustrating the fact that it is not the teleonomy alone that is responsible for the implausible adaptationist assumptions. In this case, at least, it is the combination of teleonomy with radical atomism and radical nativism. And in this case, in my view, these are the more obvious candidates for elimination.

My sacrificial theory is a single-factor theory of content for perceptual representations. On this theory, representations of a type, *R*, refer to some type of feature, *F*, iff instances of *R* have the function of indicating instances of *F*. An *R* indicates an *F* iff (reliably) if there is an instance of an *R* then there is an instance of an *F*. And *R*s have the function of indicating *F*s iff *R*s were selected for indicating *F*s (by natural selection) because past instances of *R*s caused a characteristic movement, *M*, and doing *M* in the presence of *F*s was (often enough) fitness enhancing. Furthermore, according to the sacrificial theory, radical atomism and radical nativism are true.

According to radical atomism, every (or virtually every) simple lexeme of a natural language has a corresponding Mentalese lexeme (i.e., a representation in the language of thought, or in whatever system of representations the brain employs) that is semantically simple. One way to express this is by saying that the meanings of such mental representations are not molecular, which is to say that they are not constructed out of the meanings of other simpler or more basic representations. For example, the Mentalese equivalent of "bachelor" does not have its meaning constituted out of the meanings of the Mentalese semantic equivalents of "male" and "married" and "not," and so on.

Atomism is not the same as nativism. The latter concerns whether or not we learn new concepts. Perhaps there is no uncontroversial way to spell out what a concept is or what it would be to learn one, but it may be enough here to note that there should be nothing in the definition of the relevant terms to rule out the possibility that a Mentalese lexeme can be simple and yet learned. For example, it could be learned by classical conditioning and have its meaning constitutively determined by that classical conditioning (e.g., by what it was recruited to indicate during conditioning, as in Dretske's actual psychosemantic theory). According to radical nativism, virtually all Mentalese concepts are somehow acquired or possessed without their having to be learned. The idea is that they come about as the result of maturation or are triggered by particular experiences, but that none of this counts as learning a new concept. Given the addition of radical atomism and radical nativism to the pseudo-Dretskean formula, the version of teleosemantics we are considering entails that every, or virtually every, Mentalese term-type has its content determined independently by its own individual evolutionary history. More specifically, it entails that its content is determined by past occasions in which instances of it occurred in ancestral creatures in the presence of instances of the things that belong in its extension.

This pseudo-Dretskean theory has many problems. Just one of them is that there will be many missing concepts on this theory. This is because it implies that virtually everything we can think about has been detected by and has had an impact on the fitness of our ancestors, which is of course nonsense. As Peacocke and Fodor have argued, content reaches beyond fitness. Nonexistent things, fictional things, impossible things, things that exist outside our light cone, minutely small and extremely distant things, things that only exist in the present or in the future; we can think about all of these kinds of things and our ability to do so is inexplicable on the pseudo-Dretskean theory we are considering. There are no true adaptational explanations in which such things appropriately figure in the ways specified by the theory.

Of course, thoughts about unicorns could certainly affect our fitness. A hunter who only hunts unicorns will go hungry. And someone might misperceive a small deer as a unicorn and succeed in spearing it anyway and so be happily well fed. But the pseudo-Dretskean theory requires members of the relevant represented kind themselves to impact upon our fitness. We cannot employ the pseudo-Dretskean theory to provide appropriate content for nonreferring but purportedly referring terms. No type of Mentalese lexeme was ever selected because it indicated unicorns and because it caused some characteristic movement that was fitness enhancing in the presence of unicorns, for the simple and obvious reason that there never were any unicorns. Equally obvious problems also arise for the representations of many actually instantiated kinds. We can be quite sure that no innate Mentalese lexeme was ever selected because in our ancestors it indicated quasars and caused some characteristic movement that was fitness enhancing in the presence of quasars. For one thing because quasars are not detectable with our naked senses, nor are they even resolvable with a telescope. So our ancestors did not possess a quasar detector, and so our ancestors could not have had a Mentalese indicator of quasars. For another thing, although quasars are real, and the most energetic objects known in the universe, and an encounter with one would definitely be fatal to all of life on eath, they are luckily very distant from us and must have always been so.

So here is a teleosemantic theory that is definitely committed to implausible (and even impossible) adaptational explanations. But the interesting question is, what should we learn from this? One has to be very eager to see the end of teleosemantics to see its doom forecast in such flimsy tea leaves as these. Does anyone in their right mind seriously think that we have innate and simple concepts of quasars and unicorns? (No, seriously. Not even Fodor does.) It should be obvious that there are other elements of the pseudo-Dretskean theory that can be dropped instead of the teleonomy.

Fodor (1996) argues that teleosemantics won't work because it attempts to base intentionality on natural selection's selection *for*, and this won't work, he argues, because adaptational contexts, unlike intentional contexts, guarantee existential quantification. Now it is, of course, true that adaptational explanations guarantee existential quantification. That's just to say that if, for example, the frog's optic fibers were selected for helping the frog to feed on flies (or small moving black things), then flies (or small moving black things) must obviously have existed and must have fed the frog. Natural selection can only select on the basis of *actual* past causal contributions to fitness, and no kind of thing can causally contribute if it does not exist. But it doesn't even begin to follow from this truism that a teleological theory of content cannot account for the failure to guarantee existential quantification in intentional contexts.

Actually, failure of existential quantification comes fairly cheaply (let Brentano turn in his grave, as he may). It is as cheap, at least, as a minimal capacity for *mis*-representation. Even our pseudo-Dretskean theory provides for some modest failure of existential generalization, assuming that it provides for the possibility of misrepresentation. (Fodor has vigorously denounced that latter assumption, or assumptions sufficiently like it, in earlier papers. But he intends this anti-adaptationist style of attack on teleosemantics to be a new and independent argument against it, so let's assume for the sake of the present discussion that the theory does permit the possibility of misrepresentation.) If *R* was selected for indicating *F*s, according to the formula specified, there must have been *F*s around during the selection process. But it doesn't follow from this that an *F* must be around every time an *R* is tokened. On the contrary, if misrepresentation is possible, then it is possible that an *R* be tokened in the absence of an *F*. And on such occasions, were they to occur, existential generalization would fail. Kermit's seeing something (or nothing) as a fly (or a small dark moving dot) before him does not guarantee that there really is a fly before him. He might be hallucinating.

However, the general critical point can be put this way. We humans can think about uninstantiated kinds of things, and about instantiated kinds of things that can have had no significant impact on our fitness or on the fitness of our ancestors. This is a valid objection to the pseudo-Dretskean theory (and perhaps to some actual versions of teleosemantics that have been seriously proposed).

Does this make the pseudo-Dretskean theory unduly adaptationist? Well, yes and no. Yes, it is certainly the case that the pseudo-Dretskean theory we are considering is committed to some extremely implausible adaptational assumptions (or alternatively to our having far fewer concepts than we thought we had). But to describe it as adaptationist is to do a severe disservice to any biologist who might be fairly thought of as an adaptationist (I doubt any of them were suggesting that we were really cavorting with unicorns or running away from quasars, or alternatively that natural selection could choose representations of such things for us, just in case a unicorn or a quasar happened along).

If we abandon radical atomism and radical nativism, however, we lose the implausible adaptational assumptions. A more modest teleosemantic theory (see, e.g., Sterelny 1990) is one that uses teleonomic functions to determine the content of a set of semantic simples, which are in turn used in different combinations to construct more complex concepts. This is the approach that I find most plausible. On such an approach, many of our concepts are learned or are aquired by a process involving (in some sense) construction out of simpler or more basic concepts (where simplicity is relative to a cognitive system). The idea is that ultimately all of our concepts are constructed out of semantic primitives. On a more modest theory of this sort, only the semantic primitives that purport to refer must refer to instantiated kinds that have had a significant causal impact on the fitness of our ancestors. The task of discovering what semantic simples humans and other creatures possess is a task for scientists, not for philosophers. To learn what the most plausible candidates are we should consult percepetual and other psychological theories. And I haven't heard of any that suggest that our concepts of unicorns or quasars might be primitive. The primitives of David Marr's (1982) theory of vision, for example, are such things as representations of edges, surface discontinuities, brightness, size, orientation, spatial arrangements, and so on. Here a more modest, revised pseudo-Dretskean theory will strike further problems, but the problem is no longer that the represented kinds are not instantiated or have not been detected by our ancestors or have not had a causal impact on our ancestors. (Now the problem is that there is no characteristic movement in response to these kinds of things, but this another aspect of the pseudo-Dretskean theory that I would drop.)

There are difficulties that the best of such modest proposals will have to face. But no extant theory of content is compatible with radical atomism and radical nativism, or not without costing us some concepts we thought we had.[11] The issues involved here are too large to be treated in the closing paragraphs of this chapter, but I will finish with some cursory comments on the difficulties to be faced.

One obstacle that is supposed to stand in the way of such proposals for molecular meaning is the traditional "problem of analysis": the problem of providing traditional philosophical analyses of concepts by means of providing a set of necessary and sufficient conditions that specify the conditions of application for the term under analysis. Good analyses of this kind are notoriously difficult to find; that's the problem. But these traditional philosophical analyses are not what is called for. The concepts employed in these analyses are a far cry from, for example, the primitives of the perceptual system that Marr proposed, and from which, he suggested, particular and canonical representations of objects are inferred and constructed. There is no reason to assume that the semantic primitives of Mentalese will have corresponding concepts in ordinary English (or in professional philosophical English either). Nor is there any reason to assume that there need be a set of necessary and sufficient conditions, expressed in terms of these primitives, that captures the meaning of complex concepts. Or at any rate, this is only a safe assumption if we are very lenient about what would qualify as such (e.g., if obtaining a certain outcome from the implementation of a complex program could count as satisfying a necessary and sufficient condition). Once again, we need to think outside of the philosophical tradition here. Consider, for example, Minskyean frames. Perhaps the concept of a "restaurant" is constituted by a complex data structure more nearly along these lines than along the lines of a philosophical analysis.

Another obstacle that is supposed to stand in the way of molecular meaning is the absence of an analytic/synthetic distinction. The problem is that without an analytic/synthetic distinction there is thought to be no distinction between meaning constitutive beliefs about things and other (non-meaning constitutive) beliefs about things. No distinction between the belief that bachelors are unmarried men and the belief that all bachelors have ears or that some bachelors are smelly. So, it is argued, there can be no molecular meaning, and we have to choose between meaning atomism and its opposite, meaning holism.

Of course, this is a hard problem, about which I can make only a few sketchy comments here. But I want to suggest that the problem is not as intractable as people sometimes suppose. For one thing, although traits that we can loosely refer to as innate traits need not be universal, they often are universal and will anyway generally be shared by a significant proportion of the population. This means that at the level of innate representations, there is no fast track from meaning holism (or rather, from a one-way meaning holism) to an absence of shared content. The idea needs more careful elaboration than I can give it here, but the idea is that concepts that are themselves innate, or that are formed fairly directly from concepts that are themselves innate, can be widely shared even if their meaning is determined by their place