Deconstructing Dennett's Darwin

JERRY FODOR

Abstract: Daniel Dennett's book, *Darwin's Dangerous Idea*, offers a naturalistic teleology and a theory of the intentionality of the mental. Both are grounded in a neo-Darwinian account of evolutionary adaptation. I argue that Dennett's empirical assumptions about the evolution of psychological phenotypes may well be unwarranted; and that, in any event, the intentionality of minds is quite different from, and not reducible to, the intensionality of selection.

> *Nobody likes me,*
> *Everybody hates me,*
> *I'm going to go out in the back yard and*
> *Dig up some worms and*
> *Eat them.*
>
> —*Song we used to sing around the campfire when we were*
> *very young.*[1]

Introduction

Dan Dennett's new book, *Darwin's Dangerous Idea*, is full of anecdotes, digressions, thought experiments, homilies, 'nifty' examples and other erudite displays; with such a density of trees per unit forest, it's easy to lose one's way. But I think there is discernibly a main line to Dennett's argument. I think it goes like this: Dennett offers a sketch of a metaphysical construction in which the (neo-) Darwinian account of evolutionary adaptation is supposed to ground[2] a theory of natural teleology; and this theory of natural teleology is in turn supposed to ground an account of the meaning of symbols and of the intentionality of minds. This programme is, of course, enor-

mously ambitious; much more so than adaptationism *per se*. So you can, if you wish, buy the Darwin stuff but not the teleology; or you can buy the adaptationism and the teleology but not the theory of intentionality; or you can buy all three but doubt that Dennett has got it right about how they relate to one another.

My own view is that adaptationism probably isn't true; and that, even if it is true, there probably isn't any notion of natural teleology that's worth having; and that, even if adaptationism is true and there is a notion of natural teleology that's worth having, the latter isn't grounded in the former; and that even if adaptation grounds a theory of natural teleology, natural teleology has nothing much to do with the metaphysics of meaning.

It might be fun to explore all of that, but life is short and I have tickets for the opera. So, here's what I propose to do instead. I'll say just a little, mostly pro forma, about natural selection and natural teleology and the connections between them; then I'll suppress my qualms and grant, for the sake of the argument, *a* sense of *selected for* and of *function* in which adaptations *can* be said to be selected for the functions they perform. (I think that what I'm thus proposing to concede, though it's pretty thin, is all the natural teleology that Dennett actually uses in his account of intentionality.) The main discussion will then concentrate on what I take Dennett to take to be the metaphysical connection between the teleological character of natural selection and the intentional character of thought. I'll argue, *contra* Dennett, that there isn't any such connection.

1. Adaptation

Dennett sometimes comes on pretty strong about the central status of the adaptationist paradigm in the overall scientific world view. 'Adaptationist reasoning is not optional: it is the heart and soul of evolutionary biology. Although it may be supplemented, and its flaws repaired, to think of displacing it from its central position in biology is to imagine not just the downfall of Darwinism but the collapse of modern biochemistry and all the life sciences and medicine' (p. 238). That sounds pretty fierce, alright; if, like me, you're inclined to think that just about everything in this part of the woods is 'optional', it should scare you half to death.

But it isn't clear that Dennett really means it. At one point, for example, he contemplates the possibility that Gould is right and there are 'hidden constraints' on geneotypic variability (and/or its phenotypic expression) which may substantially reduce the space of options that Natural Selection ever gets to choose from. 'The constraints of inherited form and developmental pathways may so channel any change that even though selection induces motion down permitted paths, the channel itself represents the primary determinant of evolutionary direction' (Gould, quoted by Dennett, p. 257). You might think that Dennett would reply that would make the sky fall; but, in fact, he doesn't. His rejoinder is merely methodological: '[Even if]

Address for correspondence: Department of Philosophy, Davison Hall, Rutgers University, New Brunswick, NJ 08903-0270, USA.
Email: fodor@ruccs.rutgers.edu.

[1] Readers who are puzzled by the epigraph should consult p. 17 of Dennett, 1995. All Dennett references are to that volume except as indicated.
[2] What's this 'grounding'? Search me; but philosophers, including Dennett, often like to talk that way. For present purposes, theory A 'grounds' theory B if (1) the ontology of A includes (e.g. reduces) the ontology of B, and (2) A is independently justified.

hidden constraints guarantee that there is a largely invisible set of maze walls ... in the space of apparent possibility ... we still can do no better in our exploration of this possibility than to play out our reverse-engineering strategies at every opportunity ...' (p. 261). Or, as he puts it in a summary passage 'Good adaptationist thinking is always on the lookout for hidden constraints, and in fact is the best method for uncovering them' (p. 261). That makes it look as though there is practically nothing that an adaptationist in good standing is required to believe about how evolution actually works; he's only required to buy into a methodological claim about how best to find out how it does. (Dennett doesn't, by the way, give any argument that assuming adaptationism is the 'the best method' for uncovering hidden constraints; or even that it's a good method (as compared, for example, to the direct investigation of the biophysics of genotypic variability). I don't blame him; it's notoriously hard to predict which way of doing a piece of science will work. The positivists were right about there not being a logic of discovery).

So it's a bit unclear just what degree of commitment to adaptationism Dennett requires of biologists who wish to be in scientific good repute. It's also unclear exactly what doctrine it is that he requires them to endorse. Sometimes it seems pretty tepid; however it was we who got here from the primal protoplasmic slime, no miracles ('skyhooks') were required *en route*. One might indeed take a hard line on that—I am myself prepared to sign on—since it is asymptotically close to empty. If we did come across a miracle, we'd call it a basic law and swallow it down accordingly (cf. electricity, action at a distance and the like). Unlike many of my philosophical colleagues, I do think that there are substantial constraints that physicalism and reductionism impose on how we choose among special science theories. But these constraints are subtle and very abstract. It's unimaginable that they should have the force of an a priori endorsement (or prohibition) of any substantive evolutionary theory.

Sometimes, however, what Dennett means by adaptationism isn't merely methodological and really does have teeth; it includes not just the familiar Darwinian schema *incremental, random variability—▶natural selection for fitness—▶inheritance,* but such very tendentious views as Dawkins's 'selfish gene' story, according to which the Darwinian schema has primary applicability at the level of the individual gene (as opposed, say, to the unit of selection being the genotype, or the organism, or the breeding group, or the species). It's hard to imagine that Dennett really thinks that *that* theory is other than optional, or that the integrity of biology, or 'the life sciences and medicine', or 'secular humanism' (see p. 476) turns upon it. But so his text sometimes suggests.

I propose not to get involved in this; it will do, for my purposes, merely to stipulate. So, let an adaptationist be, at a minimum, somebody who thinks that the explanation of very many interesting phenotypic properties—and, in particular, of very many of such phenotypic properties as distinguish species—invokes a history in which:

(1) there is substantial and largely unsystematic genetic variation,
(2) in consequence of which, there is substantial and largely unsystematic variation in the corresponding phenotypes;
(3) natural selection operates differentially upon the phenotypic variants depending on their relative 'fitness';
(4) there is correspondingly differential inheritance of the genotype of the selected forms.

I suppose something like this to be about as relaxed as (neo-) Darwinism can get short of lapsing into mere vacuity. For example, this formulation is neutral about what 'fitness' consists in; it might be likelihood of individual survival, or likelihood of group survival, or likelihood of mating, or of contributing to the gene pool ... or whatever. Also, the doctrine I've gestured towards isn't, as it were, 'principled'. One could hold that something like what it calls for really does happen some of the time, but that it doesn't happen all the time; or that it's a part, but not the only part, of how evolution works; or that, though it's often true, it's wrong about the evolution of some interesting phenotypic properties of, for example, us. And so on.

I think that Dennett considerably overestimates the extent to which there is currently an adaptationist consensus in evolutionary theory (see the review by H.A. Orr, 1996). My own view is that Gould, Lewontin, Eldredge and that crowd have made a substantial case that even this sort of relaxed Darwinism is right about a good deal less than the whole nature of evolution. I don't, however, propose to argue that this is so. For one thing, I have no claim at all to expertise in this area; and, for another, there's every indication that those guys can fight their own battles (see especially Eldredge, 1995). I can't, however, resist a word on why I'm particularly dubious about the application of even a relaxed Darwinism to the phenotypic properties that I care about most, viz. human psychological traits and capacities.

Suppose that there is indeed a lot of relatively unsystematic genetic variability. It's truistic (and not in dispute) that natural selection can act to filter this variability only insofar as it is expressed by corresponding phenotypic variation; genetic variants are selected for their *phenotypic* 'fitness'. Now, there are all sorts of ways in which genetic variation might turn out to be invisible at whatever the level is that natural selection works on (individual phenotype; species phenotype; or what have you.) And there are likewise all sorts of ways in which the magnitude of a genetic variation might underestimate the consequent alteration of a phenotype. For one thing, as Dennett rightly stresses, the phenotypic expression of a genotypic change is mediated by processes of, for example, protein synthesis that are, to put it very mildly, complicated. It's thus open in principle, and also in empirical fact, how much genetic variability is damped or amplified on the route to building an organism or a species: it could turn out that many genotypes converge on much the same phenotypes; or that only slightly different genotypes get grossly different phenotypic expression; or that identical genotypes get different

phenotypic expressions in different contexts, etc. In short, it's perfectly possible that *even if genetic variation is unsystematic* (no 'hidden constraints' at the genetic level) the processes that 'read' it as phenotypic variation aren't. Thus Eldredge (op. cit.) remarks that '[t]he implacable stability of species in the face of [underlying] genetic ferment is a marvellous demonstration that large-scale systems exhibit behaviours that do not mirror exactly the events and processes taking place among their parts . . .' (p. 175).

All this holds, of course, for the psychological case *inter alia*; except that here it's one fillip worse since there is still another locus at which 'hidden constraints' may have their effects. This point isn't novel, but it's well worth insisting on.

Suppose that there is relatively unsystematic variation in whatever are the genetic determinants of neurophysiological structure, and suppose further that this relatively unsystematic genetic variation is faithfully expressed by a correspondingly unsystematic variation of neurophysiological phenotypes. Still, the selection of these neurophysiological phenotypes must itself be mediated by whatever processes govern their expression as psychological traits and capacities.[3] And nothing general—I mean *nothing* general—is known about these processes. The upshot is that *even if we knew for sure* that both genetic and neurophysiological variability are more or less incremental, random and unsystematic, it wouldn't begin to follow that the variability of psychological traits or capacities is too.

So, for example, as Chomsky and Gould have frequently remarked, it is perfectly possible that there is random variability of, and environmental selection for, overall brain size; and that this is as gradualistic and classically Darwinian as you like. That would still leave it open that, at some point(s) or other, and by processes currently utterly unknown, the adventitious consequence of such physiological change is radical *discontinuities* in behavioural repertoires. (Or, if you don't like brain size, try brain weight, or brain/body ratio, or the density of neural connections, or the amount of surface folding, or the proportion of cortex to old brain, or the proportion of neurons to glia or, for that matter, the colour of the damned thing. . . . Take your pick. All (except the last) have been in fashion, at one time or another, as candidates for conveying selectional advantage.) I repeat, *nothing* general is known about how physiological variation determines variation of psychological traits and capacities. Do not believe anything to the contrary that you may have read in the Tuesday *New York Times*. ·

To summarize this line of thought: it is, of course, an empirical issue how close to being true even a relaxed adaptationism will prove to be. For it to be very close, the effects of hidden constraints on the course of evolution have to be relatively small compared to the effects of unsystematic variation

---

[3] Likewise, the organism's phenotypic constellation of psychological traits and capacities is filtered and transformed by whatever processes 'read' it onto the creature's behaviour. Hence the need for a 'performance/competence' distinction *within* theories at the psychological level.

of genotypes and their phenotypic expressions. In the psychological case, however, the notion of 'phenotypic expression' is ambiguous; it may refer to a creature's neurophysiology or to its behavioural traits and capacities. So there are not just one but *two* places where hidden constraints may enter into determining the evolutionary process. Even if the *physiological expression of genotypic variation* is largely unconstrained, there is, as things now stand, no reason to suppose that the *behavioural expression of physiological variation* is too. In fact, insofar as there is any evidence at all, it seems to suggest that reading neural structures into behavioural repertoires must substantially amplify neurophysiological discontinuities. Our brains are, by any gross measure, physiologically quite similar to those of creatures whose psychological capacities are nonetheless, by any gross measure, unimaginably inferior.

Pace Dennett, even a relaxed adaptationism about our psychological traits and capacities isn't an article of scientific faith or dogma; we'll just have to wait and see how, and whether, our minds evolved. At the time of writing, *the data aren't in.*

2. Adaptation and Teleology

Suppose, however, that adaptationism is true; is it able to ground a notion of natural teleology? Or, to put it in terms a little closer to Dennett's, suppose there is a type of organism O that has some genotypic property G that was selected because G's (characteristic) phenotypic expression P increases O's relative fitness (on average, *ceteris paribus*, in ecologically Normal circumstances . . . blah, blah, blah.) Is it then reasonable to speak of P as a property that O was 'designed' to have? Or as a 'solution' to an 'engineering problem' that O's ecology posed?

It's important to see how this kind of question might bear on issues about intentionality. If adaptations are, in some not entirely figurative sense, design features of organisms, then maybe the root metaphysical problem about how to get meaning into a purely mechanical world order is on its way to being solved. Just how this might go will concern us presently. It will do, for now, to notice that DESIGNED FOR F-ING, and the like, belong to the same family of concepts as DESIGNED WITH F-ING IN MIND. Perhaps, then, the putative connection between natural selection and design is the thin end of the intentional iceberg. Perhaps there is a route from the indubitably mechanistic adaptationist story about how phenotypes evolve, to the intentional idiom which (eliminativists excepted) everybody thinks that a good theory of the mind requires. That, in a nutshell, is how Dennett is hoping that things will pan out.

The subtext is the thing to keep your eye on here. It is, no doubt, a question that's interesting in its own right, whether adaptationism licenses teleological notions like *selection for*. But what makes that question interesting in the

present metaphysical context is that *selection for* is presumably *intensional*.[4] Just as you can believe that P and not believe that Q even though P iff Q, so a creature might be selected for being F and not for being G even though all and only the Fs are Gs. Correspondingly, the issue we're most concerned with is whether a naturalistically grounded notion of *selection for* would be intensional in the ways, and for the reasons, that mental states are. If so, then maybe a naturalistic teleology is indeed a first step towards a naturalistic theory of mind.

But, promising though it may seem, I'm afraid this line is hopeless, and for familiar reasons. Design (as opposed to mere order) requires a designer. Not theologically or metaphysically (*pace* Paley, Bishop Wilberforce *et al.*) but just conceptually. You can't explain intentionality by appealing to the notion of design because the notion of design *presupposes* intentionality. I do think this is obvious and dreary and that Dennett should give up trying to swim upstream against it; (especially since, as we're about to see, there's a different route to what he wants that will probably suit his purpose just as well).

Anyhow, here's the (familiar) argument. Patently, not every effect that a process has is *ipso facto* an effect that it designs; short of theology, at least some effects of every process are merely adventitious. This must hold of the process of natural selection *inter alia*. So, in evolutionary theory as elsewhere, if you wish to deploy the idiom of posed problems and designed solutions, you must say something about what designing requires *over and above mere causing*. Lacking this distinction, everything a process causes is (vacuously) one of its designed effects, and every one of its effects is (vacuously) the solution to the problem of causing one of *those*.

To be sure, if solutions aren't distinguished from mere effects, it does come out—as Dennett would want it to—that the giraffe's long neck solved the problem of reaching to the top of things, and did so under precisely the ecological conditions that giraffes evolved in. But equally, and in exactly the same sense, I solve the problem of how to make a Jerry Fodor under the genetic and environmental conditions that obtained in making me; and the Rockies solve the problem of how to make mountains *just like the Rockies* out of just the materials that the Rockies are made of and under just the conditions of upthrust and erosion in which they were formed; and the Pacific Ocean solves the problem of how to make a hole of just that size and just that shape that is filled with just that much salt water; and the tree in

my garden solves the problem of how to cast a shadow just that long at just this time of the day.[5] This is, however, not a metaphysical breakthrough, it's just a rather pointless way of talking; neither I nor the Pacific get any kudos for being solutions in this attenuated sense. That's because problems are like headaches; they don't float free of people's having them. The Pacific and I didn't *really* solve anything because *nobody had* the problems that we would have been solutions to. ('Who would want *those*?' people always ask).

Serious talk about problems and solutions requires a serious account of the difference between designing and merely causing. Notice, moreover, that if your goal is a reductive theory of intentionality, then your account of this difference cannot itself invoke intentional idiom in any essential way. This really does make things hard for Dennett. In the usual case, we distinguish designing from mere causing by reference *to the effects that the designer did or didn't intend*. For example: the flowers Sam gave Mary made her wheeze and did not please her. They were, nonetheless, a failed solution to the please-Mary problem, not a successful solution to the wheeze-Mary problem. That's because Sam intended that receiving the flowers should please her and did not intend that they should excite her asthma. Suppose, by contrast, that Mary merely *came across* the flowers, and that they both pleased her and made her wheeze. Then the flowers didn't solve, or fail to solve, *anything*; they just had whatever effects they did. I think the intuitions here are about as clear as intuitions can be. It certainly looks as though the concept of design presupposes, and hence *cannot be invoked to explain*, the accessibility of intentional idiom.

If you found a watch on a desert island, you'd have a couple of options. You could argue that since it was clearly designed, there has to have been a designer; or you could argue that since there was certainly no designer, the watch can't have been designed. What is *not*, however, available is the course that Dennett appears to be embarked upon: there was no designer, but the watch was designed all the same. *That just makes no sense.*

There may be a way out of this somewhere in Dennett's present text, or elsewhere in his writings; but if there is I honestly can't find it. I think perhaps Dennett is led to underestimate the magnitude of this problem because he allows himself such free use of the metaphor of Mother Nature as the designer from whose planning the intentionality of selection flows. But surely none of that talk can be meant to be literal; reduction has to stop somewhere, but it can't conceivably stop there. 'Jerry Fodor may joke about the preposterous idea of our being Mother Nature's artifacts, but the laughter rings hollow; the only alternative views posit one skyhook or another' (p. 427). But Dennett doesn't seriously think that having a scientific world view requires believing *that*, does he? The story about Mother Nature is,

---

[4] See Sober (1984) whose generally admirable treatment of these issues is, however, mistaken at a crucial point. Sober claims that *selection for*—but not *selection*—is intensional, and also that what gets *selected* is objects, whereas what gets *selected for* is traits (properties). But, as we're about to see, *selection for* contexts are *extensional* (viz. transparent, like *selection* contexts) in respect of the substitutivity of coextensive trait terms when the coextension is (e.g. nomically) *necessary*. That, in fact, is one of the main reasons why *selection for* is so unlikely to ground a theory of the intensionality of the mental.

[5] Moreover, all these solutions meet Dennett's requirement of being 'algorithmic'. It is, however, unsurprising that they do so since, as Dennett rightly says, there aren't '. . . any limits at all on what may be considered an algorithmic process . . . (p. 59).' Step right up and play; every contestant is a guaranteed winner.

after all, a *fairy tale*. There *isn't* any Mother Nature, and no intentional agent literally planned me or designed the Rockies or saw to the tree in the garden. Could Dennett really have lost sight of that?[6]

You can't get natural teleology by postulating designerless designs. Still, both for the sake of the argument and because it may be true, I'm prepared to grant that a different approach to metaphysical reduction really might yield a certain kind of natural teleology. Consider a well-worn example (Hempel, 1965). Hearts pump blood and they make the familiar sort of noise, and this correlation is empirically reliable. Nonetheless, intuition strongly suggests that hearts were selected for being blood pumps, not for being noise makers. That is, in such cases intuition reads 'selected for . . .' as *intensional*.[7] The inference from 'xs were selected for F' and '(reliably) F iff G' to 'xs were selected for G' is *not* valid. What, if not appeals to the intentions of a designer, underwrites such intuitions?

Maybe appeals to counterfactuals do. Maybe, in the present case, the counterfactual that legitimizes the inference is that hearts *would* contribute to fitness (hence would be selected) in (nearby?) worlds where they pump blood but don't make noise; but they wouldn't contribute to fitness (hence wouldn't be selected) in worlds where they make noise but don't pump blood. At least this way of getting intensionality into the picture doesn't invoke the operation of intentional systems (minds);[8] to that extent it might be of use to a reductivist metaphysical programme.

This kind of approach to natural teleology has well-known problems, to be sure. For one thing, it's not entirely clear that the required counterfactuals are true. Teleological explanations generally offer sufficient but not necessary conditions for selection. Maybe if hearts didn't pump, they would be selected for something else (Hempel, op. cit.; Dennett recognizes this sort of point in a slightly different context, see pp. 259–60).

For another thing, maybe there just aren't any (nearby?) worlds of the kind that the counterfactual contemplates. It's arguably *a law*—viz. nomologically necessary—that hearts pump blood iff they make noise; and who knows what to make of counterfactuals whose antecedents are *necessarily* false. Not only: 'who knows what their truth conditions, if any, are?' but also 'who knows what, if any, roles they can play in empirical, e.g. biological, theories?'

Digression: Ruth Millikan has recently been pushing a line rather different from Dennett's (if I read her right; which, however, Dennett says I hardly ever do) according to which what a trait is selected for is those of its effects which (completely) explain why a creature has it. Correspondingly, the intensionality of 'selected for . . .' derives from the intensionality of '(completely) explains . . .' (Presumably from *xs being F (completely) explains y's being G* and *xs are F iff xs are H*, it does *not* follow that *xs being H (completely) explains ys being G*.)

I am pretty comprehensively unmoved by this. Though it's true that mere extensional equivalence does not license the substitution of 'F' for 'G' in *F explains that . . .*, it looks like (nomologically) *necessary* equivalence does (see fn. 4). In particular, if 'F iff G' is a law, then for every explanation according to which trait T is selected for because it causes F, there will be a corresponding, equally complete, equally warranted, explanation according to which T is selected for because it causes G *in a world where, reliably, Gs are Fs*.

The standard philosophical illustration goes like this: is what explains the fitness of the frog's feeding reflex that it causes snaps at flies? Or is what explains its fitness that it causes snaps at ambient black dots in circumstances where it is nomologically reliable that the ambient black dots are flies? Unless there is something to choose between these explanations, there will be nothing to choose between the corresponding claims about what the behavioural repertoires of frogs are for. Notice that, though this example concerns the selection of an intentional trait, the problem it invokes is *entirely general*: Appeals to being F can explain nothing that isn't equally well explained by appeals to being G *in a world where it's necessary that Gs are F*.[9] In effect, explanation contexts are transparent to the substitution of (e.g. nomologically) necessary equivalents. (For a less sketchy discussion of this group of issues, see Fodor, 1990, and papers in Loewer and Rey, 1991).

As remarked above, in the present context the main issue is whether a naturalistic reduction can license *intentional* teleological idiom. Here, then, is what's to choose between Millikan's and Dennett's account of *selection for*. Dennett *can* distinguish between selection for nomologically equivalent F and G. That's because he helps himself to 'designed for'; and a thing can be designed for F-ing and not for G-ing even if it's a law that only Gs F; designing is *more* intensional than explaining. The down side, for Dennett, is that there is no design without a designer; so, willy nilly, he finds himself in bed with Mother Nature. Millikan tries rather harder not to cheat; *explains* is (arguably) a relation among *propositions*, so (maybe) it doesn't tacitly invoke

---

---

an agent.[10] The down side, for Millikan, is that explanation isn't intentional *enough*; unlike the intensionality of the mental (and, in particular, unlike the intensionality of design), explanation doesn't distinguish necessary equivalents. So, nor, of course, does a notion of *selection for* that 'explains' is used to explain.

Much of a not muchness, if you ask me.

Well, so much about several approaches to natural teleology, and the qualms that they inspire. I'm sort of inclined to doubt that there *is* any natural teleology; I sort of think that the only goals there are the ones that minds entertain. But never mind. For the sake of the argument, I hereby grant that there is such a thing as *selection for* after all, that it can be naturalistically grounded, and that selection for F doesn't imply selection for G even if 'F iff G' is reliable. The argument will now be that, even so, the intensionality of 'selection for' doesn't help with the metaphysics of the intentionality of minds.

4. Deconstruction

I'm proposing to deconstruct what I take to be Dennett's metaphysics of intentionality. Now, as everybody knows who's in this line of work, the way you set about a bit of deconstructing is: You winkle out of the text some aporia (that means, roughly, *glitch*, or *screw-up*) that reveals its deep, unconscious yen for self-refutation. (Texts are what yen these days, the author being dead.) It is, in this fashionably post-Modern spirit that I call your attention to what strikes me as distinctly a peculiarity in Dennett's views.

On the one hand, it's of the essence of Dennett's programme to argue that there is no *principled* difference between the intentionality of natural selection and the intentionality of mind. A *principled* difference would need a Skyhook to get around it; whereas Dennett's primary effort is to overcome the 'resistance philosophers have shown to evolutionary theories of meaning, theories that purport to discern that the meaning of words, and all the mental states that somehow lie behind them, is grounded ultimately in the rich earth of biological function)' (p. 402). Or, as he puts it in a summary passage: 'Real meaning, the sort of meaning our words and ideas have, is itself an emergent product of originally meaningless processes—the algorithmic processes that have created the entire biosphere' (p. 402).

It is because he cares so much about the continuity of the intentionality of 'words and ideas' with the intentionality of selection that Dennett is so hard pressed to deny that the distinction between 'derived' and 'original' intentionality is principled. Since '. . . your selfish genes can be seen to be

---

[10]   I'm being very nice and assuming, concessively, that explanation isn't itself a pragmatic notion. If what explains what depends on who wants to know what, then Millikan's line is just as question begging as Dennett's.

the original *source* of your intentionality—and hence of every meaning you can ever contemplate or conjure up . . .', the soi-disant 'original' intentionality of your thoughts is every bit as derived as the patently parasitic intentionality of thermostats and street signs. Thermostats and street signs derive their intentionality from us; we derive our intentionality from Mother N (specifically, from the reproductive ambitions of our selfish genes). Only the evolutionary process is originally intentional *sans phrase*.

That is one part of Dennett's story. But there is another part, equally strongly insisted on, that might well strike one as distinctly in tension with the sorts of passages just cited. For, it turns out that '. . . Mother Nature (natural selection) can be viewed as having intentions, [only] *in the limited sense of having retrospectively endorsed features for one reason or another . . .*' (p. 462). 'The Local Rule is fundamental to Darwinism; it is equivalent to the requirement that there cannot be any intelligent (or far-seeing) foresight in the design process, but only ultimately stupid opportunistic exploitation of whatever luck . . . happens your way' (p. 191). In this respect, it appears, Mother Nature is *not*, after all, much like you and me. 'If [for example] you were playing chess under hidden constraints, you would adjust your strategy accordingly. Knowing that you had secretly promised not to move your queen diagonally, you would probably forgo any campaign that put your queen at risk of capture . . . but you have knowledge of the hidden constraints and foresight. Mother Nature does not. Mother Nature has no reason to avoid high-risk gambits; she takes them all, and shrugs when most of them lose' (p. 259). It is, indeed, precisely because of her lack of foresight that Mother N is forever getting trapped in local maxima: 'If only those redwoods could get together and agree on some sensible zoning restrictions and stop competing with each other for sunlight, they could avoid the trouble of building those ridiculous and expensive trunks' (p. 255; Dennett is quoting Dennett, 1990, p. 132).

So, on the one hand, 'There may be somewhat nonarbitrary dividing lines to be drawn between biological, psychological, and cultural manifestations of this structure [decision making by 'satisficing'], but not only are the structures—and their powers and vulnerabilities—basically the same, the particular contents of "deliberation" are probably not locked into any one level in the overall process but can migrate' (p. 504). But, on the other hand, Mother Nature is a *blind* watchmaker; and her blindness consists (*inter alia*, see below) in this: her reasoning is never *prudential*; she maximizes her utilities *only in retrospect*; she *never* adopts (or rejects) a policy because of outcomes that she has *foreseen*. Whereas, you and I do that sort of thing all the time. *Why doesn't this difference between her and us count as principled?* I call that an aporia (or do I mean an aporium? Probably not.)

None of this, please note, takes back so much as a word of what I've agreed to concede. In particular, I am still reading 'select for' and the like as intensional. But though Mother N's *selecting for* . . . and my *intending to* . . . have their intensionality in common, I'm now wanting to emphasize a residual difference between them. Mother N never rejects a trait because she

can imagine a more desirable alternative; or ever selects for one because she can't. We do.

Is this difference principled? I'll argue that it is, anyhow, qualitative; it shows that the intentionality of minds is a different kind of thing from the (putative) intensionality of selection. First, however, it's important to get clear on the scope of the problem. What we've got so far is only the entering wedge.

It's true but a bit misleading to say that Mother N can't foresee the consequences of selection. For, if she can't see forward, she can't see sideways or backwards either; all she can see is where she is. So, just as Mother N never prefers F to G because it will (likely) bring it about that H, so she also never prefers F to G because if *would* bring it about that H if, counterfactually, it *were* the case that K; and she never prefers F to G because Fs but not Gs used to be Js (though they're not anymore). In short, it's not the (likely) future as such that Mother Nature is blind to; it's the *nonactual as such* that she can't see.

In fact, even that underestimates the magnitude of her myopia. Just as the process of selection is blind to whatever is unactual, so it is unaffected by any actual thing that is, *de facto*, causally isolated from the process of selecting. Hence, for example, one aspect of the impact of geography upon evolution: let it be that Fs are predators on Gs; and let it be that some genetic variation produces a strain of G*s that are invisible to Fs. Still being F-invisible won't provide these G* variants with any selectional advantages over the old fashioned sort of Gs if, as it happens, *all the Fs live in some other part of the forest*. Selectional advantage is the product of, and only of, *real causal interactions*. Merely possible competitions don't enter into Mother Nature's calculations; not even if they are possible competitions among *actual* creatures. None of this, surely, is in the least contentious; in fact, it's close to tautology: *Nothing is ever the effect of merely possible causes. Nothing* is. Not *ever*.

Mother Nature never prefers any Fs to any Gs *except on grounds of (direct or indirect) causal interactions that the Fs and G actually have with the selection process*. This looks to me like a whopping difference between Mother N and us. Indeed, if you think of intentionality in the Good Old Brentano way, it is exactly what makes mental states inten*t*ional and not just inten*s*ional; viz. their capacity for having nonactual objects.[11]

Here's a way to summarize the point: mental states are *both* intensional *and* intentional. Or, if you prefer, mental state predicates resist *both* substitution of coextensives *and* quantifying in. So, in the first case, Lois may fail to realize that Clark is Kent even though (ho hum!) she knows quite well, thank you, that Clark is Clark. And, in the second case, Lois may believe that Santa Claus comes down the chimney even though (ho hum!) there isn't any Santa Claus. Now, I don't really think that the intensionality of selection explains the intensionality-with-an-s of mental states. That's because, as I

---

[11]   This last is by way of cashing footnote 7.

remarked above, however it's intensionality is grounded, it's likely to turn out that 'selected for' is transparent to the substitution of *necessarily* coextensive terms; whereas 'realizes', 'thinks about', 'believes that' and the like certainly are not. But even if selection did explain why the mental is intensional-with-an-s, it *still* wouldn't explain why the mental is intentional-with-a-t. That's because a state that is intentional-with-a-t can have 'Ideal' or 'intentionally inexistent' objects, and its causal role can (typically does) depend on which Ideal object it has. Whereas, to repeat, only what is *actual* is visible to Mother Nature. *Only* what is *actual* can affect the course of a creature's selectional history.

This sort of point is occasionally remarked upon in the literature on natural teleology; see, for example, Allen and Bekoff (1995): 'For a thing to possess a biological function, at least some (earlier) members of the class must have successfully performed the function. *In cases of psychological design, the corresponding claim is not true*' (p. 61; my emphasis). That is: unlike the traits that Mother N selects (for), the goals to which intentional activity is directed can be entirely Ideal. Philosophers whose project is the reduction of intentionality to natural teleology seem not to have understood how much difference this difference makes.

But, you might reasonably wish to argue, this can't be a problem for *Dennett*'s kind of reductionism unless it's a problem for *everybody's* kind of reductionism. For, you might reasonably wish to continue to argue, if *nothing* can be the effect of a merely Ideal cause, then *a fortiori*, thoughts, decisions, and actions can't be the effects of merely Ideal causes. And similarly for other intentional goings on. Remember that Brentano thought that the intentionality of the attitudes shows that Naturalism can't really be true; and Quine thinks that Naturalism shows that the attitudes can't really be intentional; and the Churchlands think (at least from time to time), that the intentionality of the attitudes shows that there can't really be any attitudes. If, in short, there's an argument that Mother N is blind to merely possible outcomes, then there must be the same argument that you and I are blind to merely possible outcomes. So either Dennett wins or everybody loses.

Now, I do think that's a puzzle; but it's not one that I just made up. In fact, it's just the old puzzle about intentionality-with-a-t (as distinct from the old puzzle about intensionality-with-an-s). How can what is only Ideal affect cognitive and behavioural processes? How can I think about, long for, try to find, blah, blah, the gold mountain *EVEN THOUGH THERE ISN'T ANY GOLD MOUNTAIN*? This is a question which, though it arises for mental processes, *has no counterpart for processes of natural selection* since, to repeat, although 'selected for . . .' and the like are maybe sometimes opaque to the substitution of equivalents (viz. when the equivalence is not—nomologically or otherwise—necessary), they are patently always transparent to existential quantification.

Preliminary moral: you cannot solve the problem of meaning by reducing the intentionality of mind to the intensionality of selection, because the intensionality of selection fails to exhibit the very property that makes meaning

problematic. Mother Nature has no foresight, so she can't ever select for a trait that isn't there. But *you can mean something that isn't there*; you can do that any time you feel like.

But then, what sort of story *should* one tell about intentionality-with-a-t? There's a commonsense approach which, in my view, probably points in the right direction but which, no doubt, Dennett would think invokes a skyhook: merely intentional objects can affect the outcomes of cognitive processes *qua, but only qua, represented.* What *appear* to be the effects of Ideal objects on causal processes are invariably mediated by *representations* of these objects; and representat*ions*, unlike their represent*ees*, are actual by stipulation (they're physical tokens of representation types). The fact that elephants fly can't make anything happen because there isn't any such fact. But tokens of the symbol type 'elephants fly' mean that elephants fly, and they can make things happen because there are as many actual ones of those as you like.

That doesn't, of course, *solve* the problem of intentionality; it merely replaces it with the *unsolved* problem of representation (i.e. of 'meaning that'). Now, I don't know whether representation is a skyhook, *and neither does Dennett.* Either there is a naturalistic theory of representation—in which case, it *is* the solution to the problem of intentionality—or there is no naturalistic theory of representation, in which case I, for one, will give it all up, become an eliminativist about the mental, and opt for early retirement (well, early*ish*). But, either way, the present point stands: You can't reduce intentionality to 'selection for' because *selection for doesn't involve representation.* (*A fortiori*, selection for Fness doesn't involve representing anything as F). That, in a nutshell, is the difference between what Mother N does when she *selects for* tall petunias, and what Granny does when she *breeds* for them.[12]

You can, in short, suppose that the whole (neo-)Darwinian story is true; and you can suppose that 'selection for' is intensional; you will not thereby have succeeded in supposing any representation into the world. And, according to commonsense (and according to me) it's representation that you need to explain intentionality.

Three caveats and we're finished: first, I haven't claimed that only minds can represent. On the contrary, we informational semanticists are inclined to think that the representational is a (possibly large) superset of the intentional. So, for example, it's OK with us if genes, or tree-rings, or the smoke that means fire, have (underived) representational content. It's just that, according to the previous argument, they don't have it *qua selected.*

Second, I hope you don't think that I think that the line of argument I've

---

12 Sober (op. cit., p. 202) remarks that: 'Artificial selection is a variety of natural selection; the relation is one of set inclusion, not disjointness.' Well, yes and no. It's quite true that, if you are typing selectors by the *effects* of their sorting, Granny and Mother N are both just filters on phenotypic variance. But if you are typing selectors by *how they achieve their effects*, the difference could hardly be more striking. In the whole Natural Order, as far as anybody knows, only Granny and her kind filter with, literally, *premeditation.*

been pursuing shows that natural selection couldn't have *resulted in* intentional processes. Of course it could; or, anyhow, of course it could for all that I know. The issue about the 'source' of intentionality is not the *historical* question 'what made intentional things', but the *metaphysical* question 'what makes intentional things intentional'. If what produced intentionality turns out to have been the action of electricity on the primal soup, it wouldn't follow that intentionality is either a kind of soup or a kind of electricity.

Third, a point that should be obvious: it's OK for somebody who is a reductionist/naturalist about intentionality to not believe that intentionality reduces to adaptation or natural teleology. Dennett, having appropriated 'naturalistic' for his own brand of reduction, seems unable to contemplate a view that says 'yes, intentionality is reducible, but no it's not reducible to Darwin'. Hence such really bizarre misreadings as '... Searle[] and Fodor[] ... concede the possibility of [a clever robot] but ... dispute its "metaphysical status"; however adroitly it managed its affairs, they say, its intentionality would not be the real thing' (p. 426).

But I don't say that, and I never did (and, as I read him, Searle doesn't and didn't either. I may have mentioned that at one point, Dennett chastises me for being unsympathetic in my reading of him and Millikan. Talk about your pots and your kettles!). What I say is that, however clever a robot (or a creature) is at managing its affairs, its intentionality, its *having affairs to manage*, does not *consist in* the cleverness with which it manages them. If it did, intentionality would supervene on behaviour and behaviourism would be true. Which it's certainly not.

I also say that, however clever a robot (or a creature) may be, and however much its cleverness may, in point of historical fact, have been the cause of its having been selected, its intentionality does not *consist in* its having been selected for its cleverness. Dennett really must see that saying these sorts of things is quite compatible with being as naturalist and reductionist as you like about minds. 'Jerry Fodor may joke about the preposterous idea of our being Mother Nature's artifacts, but the laughter rings hollow; *the only alternative views posit one skyhood or another.*' Nonsense. There are lots of alternatives to adaptationist accounts of intentionality. Some are eliminative and some are reductive; some are naturalistic and some aren't; some are emergentist and there are even one or two that are panpsychist. What they have in common is only their being reconciled to a thought that Dennett apparently can't bring himself to face; a truly Darwinian idea, and one that is, if not precisely dangerous, at least pretty disconcerting: *there isn't any Mother Nature*, so it can't be that we are her children or her artifacts, or that our intentionality derives from hers.

Dennett thinks that Darwin killed God. In fact, God was dead a century or so before Darwin turned up. What really happened was that the Romantics tried to console themselves for God's being dead by anthropomorphizing the natural order, but Darwin made it crystal clear that the natural order couldn't care less. *It wasn't God that Darwin killed, it was Mother Nature.*

Department of Philosophy
Rutgers University

References

Allen, C. and Bekoff, M. 1995: Biological Function, Adaptation and Natural Design. *Philosophy of Science*, 62, 4, 609–22.
Dennett, D.C. 1990: Teaching an Old Dog New Tricks. *Behavioral and Brain Sciences*, 13, 76–7.
Dennett, D.C. 1995: *Darwin's Dangerous Idea*. New York: Simon and Schuster.
Eldredge, N. 1995: *Reinventing Darwin*. New York: John Wiley and Sons.
Fodor, J. 1990: A Theory of Content, Part I. *A Theory of Content and Other Essays*. Cambridge, MA.: MIT Press.
Hempel, C.G. 1965: The Logic of Functional Analysis. In C.G. Hempel, *Aspects of Scientific Explanation*. New York: Macmillan.
Loewer, B. and Rey, G. (eds.) 1991: *Meaning in Mind*. Oxford: Blackwell.
Orr, H.A. 1996: Dennett's Dangerous Idea. *Evolution*, 50, 1, 467–72.
Sober, E. 1984: *The Nature of Selection*. Cambridge, MA.: MIT Press.

Granny versus Mother Nature—No Contest

DANIEL C. DENNETT

Abstract: Fodor's doubts about neo-Darwinism are driven by something other than familiarity with evolutionary biology, so they should be set aside. His claim that a theory of intentionality cannot be constructed on an evolutionary foundation because there is no representation in the process of natural selection reveals that he has been blind to the chief beauty of Darwin's vision: its capacity to explain not just how the living can come, gradually, from the non-living, but also how meaning can come, by incremental steps, out of the meaningless.

I've been looking forward to seeing Jerry Fodor's reaction to my book, since his candidly avowed antipathy toward evolutionary arguments was one of the spurs for writing it. For instance, it was his abrupt comment to me in 1985 to the effect that Searle was right about robots lacking original intentionality that set me to writing 'Evolution, Error and Intentionality' (1987), and that contributed in turn to some of his recent outbursts against evolutionary approaches to these issues. Nothing clears the air quite so briskly as one of Jerry's jaunty tantrums. He is the master of blithe self-exposure, and on this occasion he is true to form: he reveals that his views are much more radical than I had realised; indeed, he reveals that they are much more radical than he himself has realised. Such huge disagreements cannot be resolved or even duly criticised in a few pages; the best course here is just to use Fodor's declarations to sharpen the contrasts. Figuring out 'who wins' can be saved for another occasion.

Fodor 'deconstructs' one of the main strands of argument in my book into three steps: (1) secure adaptationism against its critics; (2) show how it permits us to speak of biological functions (show how it grounds 'natural teleology', in his terms); and (3) use that concept of function as the basis for a functionalist theory of meaning or intentionality. So far, so good; that's just what I take myself to have done. He has doubts about all three steps, but is willing to suspend judgment on the first two and dig in his heels on the third. This is a retreat worth noting. In the old days, Fodor used to say that any bio-functional theory of meaning such as mine or Millikan's was hope-

Address for correspondence: Center for Cognitive Studies, Miner Hall, Tufts University, Medford, MA 02155, USA.
Email: ddennett@diamond.tufts.edu.