Fodor on Cognition, Modularity, and Adaptationism*

Samir Okasha†

This paper critically examines Jerry Fodor's latest attacks on evolutionary psychology. Contra Leda Cosmides and John Tooby, Fodor argues (i) there is no reason to think that human cognition is a Darwinian adaptation in the first place, and (ii) there is no valid inference from adaptationism about the mind to massive modularity. However, Fodor maintains (iii) that there *is* a valid inference in the converse direction, from modularity to adaptationism, but (iv) that the language module is an exception to the validity of this inference. I explore Fodor's arguments for each of these claims, and the interrelations between them. I argue that Fodor is incorrect on point (i), correct on point (ii), partially correct on point (iii), and incorrect on point (iv). Overall, his critique fails to show that adopting a broadly Darwinian approach to cognition is intellectually indefensible.

1. Introduction. Recent years have seen an explosion of interest in the idea that at least some aspects of human cognition are the product of Darwinian natural selection. This is the central idea around which evolutionary psychology is based, and is prima facie quite plausible. For many of the morphological, physiological, and anatomical traits of humans are indisputably adaptations, so it is reasonable to suggest that the same may be true of our cognitive traits too. Though solid empirical evidence on the issue is unfortunately in short supply, the idea that the cognitive mind has been fashioned by natural selection must surely merit serious consideration, if only because of the success of the Darwinian mode of explanation in general biology. Nonetheless, evolutionary psychology has encountered

*Received July 2002; revised November 2002.

†To contact the author, please write Samir Okasha, Department of Philosophy, University of York, UK; e-mail: so5@york.ac.uk. Or Instituto de Investigaciones Filosoficas, Circuito Mario de la Cueva, Ciudad de la Investigacion en Humanidades, UNAM, Delegacion Coyoacan, CP 04510, Mexico DF, Mexico; e-mail: samir@filosoficas.unam.mx.

Philosophy of Science, 70 (January 2003) pp. 68–88. 0031-8248/2003/7001-0007$10.00

considerable resistance from cognitive scientists, biologists, and philosophers alike. My aim in this paper is to examine some of Jerry Fodor's most recent arguments against the contentions of evolutionary psychologists, or "psychological Darwinists" as he calls them, from his book *The Mind Doesn't Work That Way* (2000).

Fodor's critique merits attention for three reasons. Firstly, Fodor has long been one of the most articulate and influential opponents of evolutionary approaches to cognition. Secondly, unlike many who are critical of evolutionary psychology, Fodor gives his opponents' arguments a fair and comprehensive treatment. Thirdly, Fodor develops some novel and intriguing arguments concerning the putative relations between innateness, modularity, and adaptationism. The burden of this paper will be that some of these arguments succeed while others do not.

2. Evolutionary Psychology, Modularity, and Adaptationism. Following Leda Cosmides and John Tooby (1992), two of the founders of the discipline, the majority of evolutionary psychologists are committed to the following claims:

1. There is good reason to think that the cognitive mind is the product of natural selection, i.e., an adaptation. There is a straightforward Darwinian story to be told about how human beings with sophisticated cognitive capacities evolved from ancestral protohumans who lacked such capacities.
2. There is good reason to think that natural selection would have favored a "modular" over a "nonmodular" cognitive architecture.
3. Therefore, it is a good bet that the mind consists of a number of specialized modules, each optimized for solving a problem that would have confronted our ancestors back in the Pleistocene.

Much research in evolutionary psychology consists in trying to find evidence for the existence of these "Darwinian modules," and in trying to figure out how they work. Typically, researchers start by identifying a possible "problem" that our hunter-gatherer ancestors might have faced; they then ask what a mental module capable of solving the problem might be like; finally, they look for evidence that the postulated module actually exists in contemporary human beings. This research methodology is obviously fallible, involving an inevitable element of speculation. Nonetheless, there have been some apparent successes. For example, recent theorists claim to have found evidence that we possess modules for detecting cheaters in social exchanges, for attributing mental states to others, for facial recognition, and more. In each case, the module is claimed to be an evolved response to a problem that confronted our ancestors. The overall picture of the mind that emerges is summed up by Tooby and Cosmides as follows: "our cognitive architecture resembles a confederation of hundreds or thousands of functionally dedicated computers (often called modules) designed to solve adaptive problems endemic to our hunter-gatherer ancestors" (1995, xiv). This is the core intellectual commitment of evolutionary psychology, and the main target of Fodor's attack.

The concept of modularity is central to much evolutionary psychological thinking, but has not always been understood identically by all authors.[1] In *The Mind Doesn't Work That Way,* Fodor spends some time carefully untangling the different possible senses of modularity. As Fodor uses the term, a mental module consists of a computational mechanism which is innately specified, special-purpose, and informationally encapsulated—it doesn't have access to all of the information which the cognitive subject possesses. Typically, a module will contain its own proprietary database—a store of innate information which every child is born with, and which will often only be accessible to the module in question. But it is important to note that for Fodor a module is a computational *mechanism,* rather than the innate store of information itself. In this respect Fodor's concept of a module is different from Chomsky's. In what follows, I use the word "module" exclusively in Fodor's sense, despite the fact that some evolutionary psychologists favor the Chomskian sense, and still others do not require modules to be informationally encapsulated. I also follow Fodor in using the contrast between "domain specific" and "domain general" cognitive mechanisms as a synonym for the contrast between modular and nonmodular mechanisms.

Fodor's position on the likely extent of mental modularity remains similar to the position articulated in *The Modularity of Mind* (1983). "Input systems" such as perception and language processing are probably modular, while "central processes" such as thinking and reasoning are probably not—they are performed by domain-general mechanisms. Fodor's view thus contrasts sharply with the view favored by most evolutionary psychologists, according to which the mind is largely or wholly composed of modules, many of which do perform central processes. Fodor dubs this view the "massive modularity hypothesis" and devotes much energy to debunking it. In particular he attacks the claim, defended by Cosmides and Tooby and others, that general adaptationist considerations support the massive modularity hypothesis; this is the main burden of Chapter Four of his book. There is no particular reason to think that natural selection would have led to a massively modular mind, Fodor insists.

The Cosmides and Tooby argument that selection would have favored

1. See Segal 1996 and Samuels 2000 for useful discussions of the various notions of modularity at work in cognitive science.

a modular cognitive architecture, and thus that the mind is probably modular, clearly rests on a presupposition: that the cognitive mind is the product of natural selection in the first place. In Chapter Four Fodor does not contest this presupposition, but purely for the sake of argument. For in Chapter Five he attacks it roundly, in the course of a more general argument against the view that evolutionary considerations *must* be relevant to psychology. Fodor insists that this view is incorrect—there is no a priori reason to think that our cognitive architecture is an adaptation at all. However in the second half of Chapter Five, Fodor introduces a surprising twist: there *is* a link between adaptationism and modularity after all, he thinks. In so far as the mind *is* modular, there is good reason to think that adaptationism is true of it, for modules have certain features that only natural selection can explain. In short, the evolutionary psychologists argue that adaptationism about the mind implies modularity; Fodor rejects both the validity of this inference and the truth of its premise, but defends the validity of the *converse* inference from modularity to adaptationism; however, he thinks that the mind is largely nonmodular. And in a final, intriguing twist, Fodor claims that the inference from modularity to adaptationism, though generally a good one, does *not* apply to the language-learning module. This module is an exception to the general rule that modularity is indicative of the operation of natural selection. I examine all of these arguments below.

3. Is Human Cognition the Product of Natural Selection? It is obviously true that the cognitive capacity of *Homo sapiens* far exceeds that of any other species, living or extinct. The exact phylogeny of human cognition is not known with any certainty, but the broad picture is clear. Hominids with brain sizes similar to those of modern humans first appear in the fossil record 250,000 years ago. However, archaeological evidence suggests that it was not until 60,000 years ago, or perhaps even less, that the distinctive features of modern minds first arose (Mithen 1996, 2000). So in a relatively short period of time, protohumans with cognitive capacities much inferior to ours gave rise to humans with cognitive capacities comparable to ours.

What explains this remarkable transition? One possible answer, favored by all evolutionary psychologists and many others too, is that the cognitive mind evolved under the pressure of natural selection. There was a selective advantage associated with being able to perform more and more sophisticated cognitive tasks, and so the genes that coded for these cognitive capacities (or for the underlying brain structures on which they depended) gradually spread through the species to fixation. Now as everybody admits, there is no *direct* evidence that a Darwinian story of this sort explains the origin of the mind—but this is always the case when we put forward adaptationist hypotheses in biology. Faced with a complex organ or trait which clearly adapts an organism to its environment, biologists routinely assume that the organ or trait is an adaptation, i.e., that it evolved by natural selection. A priori adaptationist theorizing of this sort raises important methodological questions, extensively debated by biologists and philosophers of biology in the last thirty years. But even biologists opposed to extreme adaptationism accept the principle that where we have a complex, highly adapted phenotypic trait, natural selection is by far the most plausible explanation of its existence.

Evolutionary psychologists apply this general principle to the human mind. The mind is an organ of great complexity, and clearly bestows a fitness advantage on its possessors—evidenced by the impoverished survival and reproduction prospects of the mentally retarded. So surely it is reasonable to assume that the mind is the product of natural selection? As Steven Pinker puts the point: "natural selection is the only explanation we have of how complex life can evolve . . . [so] natural selection is indispensable to understanding the human mind" (1997, 55) This quotation from Pinker is cited by Fodor, along with similar quotations from Henry Plotkin (1997), Cosmides and Tooby (1992), and Richard Dawkins (1996). Fodor remarks disapprovingly, "books about psychological Darwinism simply can't get from their prefaces to their conclusions without saying this sort of thing" (2000, 87). He then argues at length that Pinker, Plotkin, Cosmides and Tooby, and Dawkins are all wrong: "the complexity of our minds, or of our behaviour, is simply irrelevant to the question of whether our cognitive architecture evolved under selection pressure" (2000, 87).

Fodor's argument for this claim is at first blush quite remarkable: he claims that the complexity of an organ or trait has actually got nothing to do with the likelihood that it is an adaptation, conventional wisdom in evolutionary biology notwithstanding. I quote the relevant passage in full:

> what does matter to the plausibility that a new phenotypic property is an adaptation has nothing to do with its complexity. What counts is only how much genotypic alteration of the nearest ancestor that lacked the trait would have been required in order to produce descendents that have it. If it would have needed a lot, then it's very likely that the alteration is an adaptation; if not, then not. In the present case, what matters to the plausibility that the architecture of our minds is an adaptation is how much genotypic alteration would have been required for it to evolve from the mind of the nearest ancestral ape whose cognitive architecture was different from our own." (2000, 87–88)

Fodor goes on to argue that it is entirely plausible that a small genetic alteration could have produced the large difference in cognitive capacity between us and those of our most recent ancestors who lacked modern minds. This is plausible, according to Fodor, because cognitive capacities supervene on cognitive structure, which in turn supervenes on brain structure (presumably), and no-one knows what the rules of supervenience are; hence small neurological changes could lead to radically different cognitive structures and/or cognitive capacities. In defense of this claim Fodor points out that our brains are very similar to those of apes by any gross measure, and yet our cognitive capacities very different. Since small neurological changes could very easily be the product of small changes to the genetic material (and since we are also genetically very similar to apes), Fodor concludes that there is "no reason at all to believe that our cognition was shaped by the gradual action of Darwinian selection." (2000, 88) The complexity and fitness-enhancing nature of human cognition do not therefore constitute evidence for its being an adaptation.

In effect, Fodor is proposing that human cognition arose as a "saltation"—the product of a sudden and relatively small neurological reorganization in an ancestral ape, probably in a single generation, rather than the product of cumulative directional selection over many generations, as evolutionary psychologists typically believe. In defense of the empirical plausibility of this suggestion, Fodor emphasizes that a prerequisite of successful Darwinian explanations is what he calls "a roughly linear relation between alteration of some physiological parameter and the consequent alteration of a creature's fitness" (2000, 89). This corresponds closely to what Richard Lewontin (1985a) has called the requirement of "continuity" on Darwinian explanations. The Lewontin/Fodor idea is quite simple. For directional selection to produce complex phenotypic adaptations, it must be the case that every slight modification of the phenotypic trait in question, in a given direction, causes incremental gains in fitness. For example, if selection is to cause a butterfly species to evolve a new wing camouflage pattern, it must be the case that even small changes to the existing wing pattern, in the direction of the new pattern, bestow an increase in fitness. All the intermediate steps between the original pattern and the final pattern must be associated with a fitness increase, if natural selection is to gradually drive the species from the former to the latter.

In many cases there is good reason to think that the continuity requirement is satisfied. For example, if a deer population is subject to heavy predation, then a deer which runs only slightly faster than average will presumably have a fitness advantage over other deer, so natural selection can gradually drive up the running speed in the population. But in the case of cognition, says Fodor, there is no reason to believe that continuity is satisfied, since small neurological changes might very easily cause large changes in cognitive structure/cognitive capacity. As Fodor says: "make an ancestral ape's brain just a little bigger (or denser, or more folded—or, who knows, grayer) and it's anybody's guess what happens to the creature's cognitive-cum-behavioural repertoire. Maybe the ape turns into us." (2000, 89–90) So in the case of cognition, there is no evidence for the existence of a neurological parameter which is linearly related to biological fitness in the required way. So there is no reason to believe that cognition is the product of gradual, Darwinian selection—it could just as easily be the product of a saltation.

I believe that Fodor's argument is flawed in two respects. The first concerns adaptationist methodology. It is quite true that in the case of cognition, we have no direct evidence that the continuity requirement is satisfied. As Fodor says, no-one knows the laws which determine how cognitive structure supervenes on neurological structure, so no-one knows whether there is a neurological parameter slight changes in which would occasion slight improvements in cognitive structure, hence fitness. But this situation is quite common in evolutionary biology. Biologists frequently offer adaptationist explanations in the absence of direct evidence that the continuity requirement is satisfied. Indeed direct evidence for continuity is rarely available, though in some cases, such as the running speed example above, it seems a priori quite plausible. In general biologists simply assume that continuity *must* be satisfied (or have been satisfied in the evolutionary past) because they are *antecedently* certain that the organ or trait in question must have evolved by natural selection. Take bat echolocation for example. This trait is so complex, and adapts bats so wonderfully to their environment, that biologists automatically assume that it arose by natural selection. So they assume that gradual modification of some rudimentary precursor of the echolocation device must have led to incremental increases in fitness; they then advance hypotheses about what the precursor device was like, how it got gradually modified, and why the gradual modifications were associated with incremental fitness gains. Similar stories are told about the vertebrate eye.[2] The crucial point is that

2. Fodor appears to misunderstand this point, citing Dawkins' (1996) discussion of the evolution of the eye in *support* of his claim that prior knowledge of continuity is needed before it is legitimate to assume a trait is an adaptation. With reference to Dawkins, he writes: "many of the great successes of Darwinian theory have consisted precisely in showing that there actually is such a parameter in a case where, prima facie, it mightn't seem that there could be one" (2000, 89). But in his discussion of the eye, Dawkins does not *show* that there is a parameter which satisfies the continuity requirement; rather he assumes that there must be, for otherwise the eye could not have evolved, and offers some highly plausible empirical conjectures about what the parameter(s) is, and how its gradual modification could have produced better and better eyes. It is odd that Fodor misses this, for Dawkins is quite explicit on the point. Discussing the lens, he says that his story is meant to show "how lenses *might* have evolved in the first place" (160; my emphasis); discussing focusing, he says that it is "not at all difficult

biologists do not wait for *direct* evidence of continuity before advancing adaptationist hypotheses; where a trait seems obviously to be an adaptation, they take that to show that continuity must in fact have been satisfied.

Why does this methodology not apply in the case of cognition? For Fodor has not shown that continuity is *not* satisfied; his point is just that, given how little we know about the neurological basis of cognition, we do not know that it *is* satisfied. But as in the case of bat echolocation, why cannot we argue that since the cognitive mind is so complex, and so obviously fitness-enhancing, it must be the product of natural selection, therefore continuity must have been satisfied? It is unclear how Fodor would respond to this argument. He appears to concede—though perhaps only for the sake of argument—that the cognitive mind is indeed complex and conducive to fitness (2000, 88). And as we shall see in Section 5, Fodor *does* allow that in some cases it is legitimate to assume a priori that a trait must be an adaptation. He is thus not totally averse to a priori adaptationist theorizing. So I think, therefore, that Fodor's view must be that human cognition does not exhibit the features which license us to make *a priori* adaptationist assumptions. Whether this is because he holds that cognition is *not* actually complex and fitness-enhancing (and only grants the contrary assumption for the sake of argument), or because he holds that cognition *is* complex and fitness-enhancing but that complexity and fitness-enhancingness are not the features that license us to make a priori adaptationist assumptions, I do not know; I return to this issue in Section 5.

The second problem with Fodor's argument is this. Even if it is true that human cognition arose as a saltation—a minor neurological change in an ancestral ape that led to dramatic cognitive reorganization—this does not mean that cognition is not an adaptation. For natural selection could still have been responsible for causing cognition to *spread* through the ancestral population, and for it to be *maintained* in subsequent generations. Even if a novel phenotypic trait arises as a single mutation, if the trait then spreads through the population, and is maintained in the population over many generations, these latter facts clearly need explanation. And the most obvious explanation is that the trait conferred a selective advantage on its bearers. Alternative explanations are of course possible:

to *imagine* the gradual evolution of a mechanism for changing focus" (167; my emphasis); discussing the pupil, he says that it "isn't difficult to see how this advanced mechanism *might* have got its start" (168; my emphasis); and discussing the transition from compound eyes to camera eyes his hypothesis is meant to offer "a glimpse of what the evolutionary progression *might* have been like" (184; my emphasis). Dawkins is adopting the standard methodology of assuming that a complex trait must have evolved by gradual Darwinian selection, and offering plausible, empirically-based conjectures about what the intermediate stages in its evolution were like.

genetic drift could have caused the mutant gene to spread, and the absence of any back mutations might explain its retention in subsequent generations. But neither of these is really very plausible, particularly since the trait in question—having cognitive abilities far superior to those of one's fellow ancestral apes—presumably would have been fitness-enhancing. So even if Fodor's saltationist theory of the *origin* of cognition is correct, there is still every reason to believe that natural selection accounts for its spread and subsequent maintenance—in which case cognition still would qualify as an adaptation, on the usual understanding of this concept. Traits that are maintained by stabilizing selection are still adaptations, whether they originated in a single mutation or not.

I suspect that Fodor misses this obvious point because he is overly impressed with a feature of Darwinian explanations that has been much emphasized by Richard Dawkins, among others. Dawkins stresses the *gradualness* of natural selection—the fact that many rounds of cumulative selection are required to produce the exquisite adaptations that we find in nature. Dawkins is quite right that the power of natural selection to produce a close adaptive fit between organism and environment depends heavily on its cumulative nature, and Fodor is quite right that cumulative selection requires continuity. But Fodor wrongly concludes that where a novel phenotypic trait arises as a single-generation saltation, rather than being gradually fashioned by cumulative selection, the trait's prevalence in the population can be explained in wholly non-Darwinian terms. But since natural selection can explain the spread and maintenance of a trait, not just its origin, Fodor's view that "saltational" explanations are diametrically opposed to Darwinian explanations is simply incorrect.

To be fair to Fodor, there is a tradition in evolutionary biology of opposing saltational to Darwinian explanations, in much the way that he does. The idea of saltational change is usually associated with Richard Goldschmidt, who coined the expression "hopeful monster" to describe the sudden appearance of evolutionary novelties in a single generation. Orthodox neo-Darwinians, including Dawkins, have heaped much scorn on the idea of "hopeful monsters." But as John Maynard Smith has pointed out, two quite different concepts have been conflated until the label "hopeful monster." Firstly there is the idea of "a complete repatterning of the chromosomal material, giving rise in a single step to a new species or higher taxonomic group" (1989, 135). This idea clearly does conflict with neo-Darwinian ideas about speciation, and is anyway unsupported by genetic evidence; as Maynard Smith says, it was rightly rejected. Secondly, there is the idea of a genetic mutation which has a very large effect on the organism's phenotype. Most such mutations would presumably be harmful, but it is possible that the occasional one would produce a viable organism, or even a highly fit organism. Maynard Smith stresses that this notion of a hopeful monster is entirely compatible with neo-Darwinism, for if such a hopeful monster arose, selection would still determine whether it survived and spread its genes in the population or not. He writes

> I do not see . . . any contradiction between neo-Darwinism and the idea of hopeful monsters, at least in the sense of a mutant of large phenotypic effect. The essential point is that the fate of hopeful monsters, like that of other mutants, depends on the operation of natural selection in populations. (1989, 136)

Maynard Smith's point is directly relevant to Fodor's suggestion that human cognition arose by a sudden neurological change in an ancestral ape which caused dramatic cognitive re-organization. For this is clearly a hopeful monster hypothesis of the second sort, not the first. Fodor is not positing a sudden massive chromosomal rearrangement leading to a new species overnight, but a genetic mutation of large phenotypic effect. And as Maynard Smith says, there is nothing un-Darwinian about such a hypothesis, for selection is still what determines the fate of the hypothesized mutation. And if Fodor's hypothesis is true, the fate of the mutation in question was to spread rapidly through the population and to be retained in subsequent generations. The obvious explanation is that the mutation, and the neurological and cognitive changes for which it coded, conferred a large selective advantage on organisms which carried it.

To conclude the section, Fodor's reasons for denying that human cognitive capacity, and the cognitive structure on which that capacity depends, is an adaptation are not convincing.

4. Would Natural Selection Have Favored a Massively Modular Cognitive Architecture? Evolutionary psychologists believe that the mind's architecture is massively modular. Whether this is so is obviously an empirical question. Evidence for massive modularity has been alleged from studies in developmental psychology, cognitive psychology, neuropsychology, neuropathology, and more. But most people agree that this evidence is far from decisive. However, many evolutionary psychologists hold that massive modularity is still a reasonable working hypothesis, because there are general adaptationist reasons for thinking that evolution would have favored a modular over a nonmodular cognitive architecture.

Cosmides and Tooby offer three such reasons. Their first is that since "what counts as fit behaviour varies markedly from domain to domain," it is impossible for an animal equipped only with a domain-general learning mechanism to learn to behave in a minimally adaptive way (1994, 91). (Like Fodor I have difficulty understanding this argument; see Samuels 2000 for a plausible gloss on what Cosmides and Tooby might be driving at.) Their second is that if an animal contains just domain-general mechanisms, the information available to it will be limited to what it can glean from perception, which is insufficient to enable it to survive and reproduce. Whereas if the animal's mind is composed of modules, each with a rich store of innate information about its environment, it will not suffer this handicap. Their third argument is that a domain-general cognitive architecture will be paralyzed by "combinatorial explosion." Since a domain-general mechanism "lacks any content, either in the form of domain-specific knowledge or domain-specific procedures that can guide it towards the solution of an adaptive problem," it will have to "evaluate all alternatives it can define" (91). However as the complexity of the problem increases, "alternatives increases exponentially," paralyzing the mechanism. Specialized modules suffer no such handicap.

Fodor criticizes these three arguments on a number of scores. His most telling criticism is that Cosmides and Tooby conflate the issue of modularity with the issue of innateness. They assume that if the mind is nonmodular, i.e., contains only domain-general mechanisms, then it must be a tabula rasa, i.e., contain no innate knowledge. But this is simply false: a creature whose cognitive mechanisms are domain-general can perfectly well possess innate knowledge. (Samuels (2000) makes a related point.) Cosmides and Tooby's arguments above exploit the fact that an animal needs a rich store of innate information about its environment if it is behave adaptively. But this would only show that selection would favor modularity over nonmodularity *if* it were true that only a modular mind could contain innate information. In Fodor's words: "*pace* Cosmides and Tooby, there is no warranted inference from a creature's possessing a domain-general cognitive architecture to its lacking an innate cognitive endowment. Arguments that it must have such an innate endowment are therefore neutral as to whether its cognition is modular" (2000, 70).

Fodor argues that the issues of innateness and modularity are actually independent in both directions: there can be nonmodular mechanisms with lots of innate knowledge and modular mechanisms with little or no innate knowledge. He writes "you can . . . have perfectly general learning mechanisms that are born knowing a lot, and you can have fully encapsulated mechanisms (e.g., reflexes) that are literally present a birth, but that don't know about anything except what proximal stimulus to respond to, and what proximal response to make" (2000, 69). Cosmides and Tooby's conflation of innateness with modularity is thus wholly mistaken, he concludes. Even if general adaptationist considerations show that the mind is likely to contain lots of innate information, that is equally compatible with a modular or a nonmodular cognitive architecture.

Does Fodor's point refute all three of Cosmides and Tooby's adaptationist arguments for massive modularity? With respect to arguments one and two, the answer is surely "yes." Grant to Cosmides and Tooby that to behave adaptively, a creature must know that "what counts as fit behaviour varies from domain to domain" (argument 1), and must possess more information "than could be gleaned from perception alone" (argument 2). It does *not* follow that the creature's cognitive architecture must be massively modular. Innate content could do the trick instead. The creature could have innate knowledge of the various different standards for fit behavior, and could innately possess whatever information is necessary for adaptive action but cannot be gleaned from perception. If so, it could presumably behave adaptively even if its cognitive mechanisms were entirely domain-general. Arguments 1 and 2 thus fail to establish the adaptive superiority of modularity.

With Cosmides and Tooby's third argument—from "combinatorial explosion"—matters are somewhat different.[3] In this case, it is unclear that innate content instead of modularity can do the trick. The basic worry is that a domain-general mechanism, because it is not informationally encapsulated, will be unable to compute the correct solution to an adaptive problem in a feasible amount of time. Since it has access to all the information in the mind, the number of possible solutions it will have to assess is unfeasibly large. It is hard to see how additional innate content can alleviate this problem of computational tractability.[4] Only if the mechanism is modular, i.e., encapsulated with respect to much of the information in the mind, will the number of alternatives that need to be assessed be reduced. Innate content does not ensure computational feasibility.

Does this mean that Cosmides and Tooby's third adaptationist argument succeeds? Can we legitimately infer that the mind is probably massively modular, because evolution would not have produced mechanisms unable to feasibly compute solutions to adaptive problems? Fodor argues that we cannot. He agrees with Cosmides and Tooby that feasible computation is only possible for modular systems; indeed this is the burden of the first half of his book. However, he argues that we are unentitled to simply assume that all mental processes are classical computations. So the most Cosmides and Tooby's argument shows is that "*either* we have the kind of cognitive architecture in which massive modularity avoids an ex-

3. This paragraph was prompted by an anonymous referee's comments, for which I am grateful.

4. Samuels (2000) asserts that innate content without encapsulated mechanisms can solve the problem but does not explain how. Cosmides and Tooby's own position on the issue is not clear. In their presentation of the combinational explosion problem quoted on page 11 above, they say that the problem arises because a domain-general mechanism "lacks any content, *either* in the form of domain-specific knowledge *or* domain-specific procedures" (my emphasis). This suggests that they think innate content alone would alleviate the problem. But as argued in the text, it is hard to see how.

plosion of Classical computation, or that (at least some) of our mental processes aren't Classical computations" (2000, 71). Given his independent belief that central cognitive processes *couldn't* be classical computations, Fodor of course favors the second disjunct.

Where does this leave us vis-à-vis the claim that evolution would probably have favored a massively modular architecture? In my view, Fodor has certainly refuted Cosmides and Tooby's first two arguments for this claim. The status of their third argument is less clear cut, for it depends on broader issues concerning the legitimacy of helping one's self to the computational theory of mind. But since Cosmides and Tooby do not have an argument for the computational theory's being true of all human cognition, and Fodor does have an argument for the contrary, it is fair to say that Fodor has the better of the dispute. Of course, future empirical work may yet vindicate the massive modularity hypothesis. But Fodor is surely right that general adaptationist considerations do not lend it much support.

5. Is There a Good Inference from Modularity to Adaptationism? The foregoing arguments notwithstanding, Fodor *does* think there is a link between modularity and adaptationism. However the link is almost exactly the reverse of the one that Cosmides and Tooby argue for, he maintains. There is no good inference from adaptationism about the mind to massive modularity, for reasons that we've seen, but there is a good inference in the other direction. That is, if the mind were massively modular, there would be good reason to think that adaptationism must be true of it. "There's a plausible line of argument that leads from massive modularity to psychological Darwinism," Fodor writes (2000, 91).

The line of argument Fodor suggests is this. According to the massive modularity hypothesis, the mind is composed of a large number of innately specified modules, which contain innate beliefs in their databases. These innate beliefs are substantive, and they are contingent—typically they are about features of the organism's environment. Furthermore, many of these beliefs are true. Fodor gives the following examples of innate, contingent true beliefs which infants are born with: that unsupported objects fall, that the auditory location of a sound source predicts its visual location, that objects continue to exist even when visually occluded. If the massive modularity thesis is true, and there are modules for practically everything we do, then there are many, many more such beliefs. The question naturally arises: "how did the infant come to have these contingent, true beliefs?". Obviously not through learning, for by hypothesis the beliefs are innate. The only plausible answer, Fodor claims, is that they were put there by natural selection. So to the extent that the mind is modular, there is good reason to think that it has been shaped by natural selection.

Modules contain detailed, contingent information about the environment that only natural selection can explain. So if the mind is massively modular, adaptationism is very likely to be true of it.

How does this argument comport with Fodor's earlier argument that human cognition is just as likely to have arisen by sudden saltation as by Darwinian selection? As we have seen, Fodor insists that there is no good inference from the complexity of the cognitive mind to its having arisen by selection, so why does he think there *is* a good inference from the cognitive mind containing innate true contingent beliefs to those beliefs having been put there by selection? Why is a priori adaptationism acceptable in the second case but not in the first? The answer, according to Fodor, is that a saltation hypothesis is implausible in the case of innate true contingent beliefs. Unlike in the case of cognitive structure/capacity, it is "surely not conceivable that relatively small, fortuitous changes in brain structure should produce massive increments in a creature's stockpile of true, contingent beliefs" (2000, 93). In other words, the innate beliefs found in the databases of the mental modules postulated by evolutionary psychologists display too close a "fit" with the environment for chance to explain their presence. Some sort of "instruction" of the mind by the world is the only plausible explanation, and since the beliefs are innate (ruling out learning), natural selection is the only instructional mechanism left. So according to Fodor, although the cognitive mind could easily have arisen as a saltation, it is very unlikely to have arisen ready stocked with innate contingent information; if the mind contains such information, it must have been put there by natural selection.

There is something right about Fodor's line of thought here, but it is misleading in one important respect. It is unclear why Fodor thinks he has provided an argument from *modularity* to adaptationism. In reality, he has provided an argument from the existence of innate contingent knowledge to adaptationism. But Fodor himself emphasized, in his attack on Cosmides and Tooby examined in the previous section, that the issues of modularity and of innateness are independent! As he rightly pointed out, a mind which contains only domain-general cognitive mechanisms, and is thus entirely nonmodular, can still have a large endowment of innate knowledge. But this makes it very hard to see why Fodor thinks there is a valid inference from massive modularity to adaptationism. For his fundamental point—that innate true contingent beliefs could only have arisen by natural selection—surely applies whether the architecture of the mind is massively modular or not.

This point bears some emphasis. Suppose for the sake of argument that the mind consists wholly or mainly of domain-general cognitive mechanisms, which have access to a rich store of innate, true contingent beliefs. As we have seen, Fodor allows that this is a logical possibility. (Samuels

(2000) argues that the mind may very well be like this.) Clearly, the question then arises: how did these beliefs get into the mind? Again, natural selection would seem to be the most plausible answer. For the beliefs contain true, contingent information about the environment, and as before, it seems most unlikely that they could have arisen by chance mutations; so detailed a "fit" with the environment is a clear indicator that natural selection has been at work. It seems obvious that Fodor should accept this argument, for it is identical in all relevant respects to his own argument. Thus the correct inference is not from *modularity* to adaptationism, but from innate, true contingent beliefs to adaptationism, whether those beliefs are processed by modular or nonmodular cognitive mechanisms. Modularity has simply got nothing to do with it. It is very strange that Fodor misses this point, for in doing so he has committed precisely the confusion of modularity with innateness of which he earlier convicted Cosmides and Tooby.[5]

As a consequence of this, the extent to which Fodor's argument represents a concession to evolutionary psychology is largely obscured. Since Fodor holds that the massive modularity thesis is false, and says so repeatedly, when he argues that there is a good inference from massive modularity to adaptationism about the mind, it is natural to assume that he is not conceding any ground to the latter. But since what his argument really shows is that there is a good inference from innate contingent content to adaptationism, matters are a bit different. For Fodor himself is a *fan* of innate contingent content—he has long maintained that "poverty of the stimulus" arguments show the mind to contain lots of innate knowledge. So he is actually making a quite substantive concession to adaptationism about cognition, though he does not say so. This helps explain what would otherwise be a very puzzling feature of Fodor's argument: the examples of infants' innate true contingent beliefs that he gives (cited above) are ones that he himself accepts! So in fact, Fodor is granting that the cognitive mind does indeed bear the hallmarks of natural selection, to no small extent. (The precise extent depends on exactly how much innate true contingent content Fodor thinks there is.) Appearances to the contrary notwithstanding, Fodor has actually provided a significant argument in *favor* of taking an evolutionary approach to cognition, at least for anyone sympathetic to the nativist view that the mind is not a tabula rasa.

In my view Fodor is basically right that innate true contingent beliefs can only have been put into the mind by natural selection. But I think he overestimates the extent to which this argument depends on the beliefs

5. Fodor (personal communication) admits that his argument is misleadingly expressed, and confirms that the inference he really wishes to defend is from innate contingent knowledge to adaptationism, rather than from modularity to adaptationism.

being *contingent.* Fodor repeatedly stresses that having any particular innate contingent belief increases one's biological fitness only if one is born into a world in which that belief is true. And in most worlds, the belief will be false. So if one is born with innate contingent true beliefs, one's mind seems to be particularly well-adapted to the environment—it contains information which is fitness-enhancing in one's actual environment but wouldn't be fitness-enhancing in most other environments. Obviously, this isn't so if one's innate beliefs are necessarily true. In that case, they will be true whatever one's environment is like, so the fact that they are true in one's actual environment will not seem so striking. Presumably because of this difference, Fodor appears to hold that a Darwinian explanation is called for only where we are dealing with innate beliefs that are contingently true.

But is this correct? Suppose for the sake of example that infants were born with innate knowledge of some complex geometrical truths. This would obviously cry out for explanation. Given all the possible geometrical falsehoods the child could have been born believing, why should it have been born knowing truths? The fact that geometrical truths are true in all possible worlds does not answer this question. If it could further be shown, or at least plausibly argued, that knowledge of the truths in question is fitness-enhancing, or would have been fitness-enhancing in an ancestral environment, it would be reasonable to suggest that natural selection accounts for their presence. Perhaps Fodor would reply that in the case of innate true beliefs that are contingent, our intuitive sense of a surprising "fit" between world and mind is greater than in the case of necessarily true beliefs. This may be so, but the fact remains, if we have substantial innate knowledge of necessary truths this presumably requires *some* explanation, and natural selection seems like the obvious candidate.

This point is slightly academic, since most of the innate beliefs posited by contemporary cognitive psychologists, including Fodor and his evolutionary psychologist opponents, are contingent. Nonetheless, it does highlight an important methodological problem. In general, the plausibility of engaging in a priori adaptationist theorizing depends on one's estimate of how likely it is that the trait or organ in question could have arisen by nonselective processes, and such estimates are partly subjective—they depend on "scientific common sense." As we saw in the case of cognitive structure/cognitive capacity, Fodor denies that we should assume a priori that we are dealing with an adaptation. In the case of innate contingently true beliefs, he thinks we should assume this a priori—their "fit" with the environment is too close to be explained in any other way. In the case of innate necessarily true beliefs, he apparently thinks we should not assume adaptationism a priori; I have suggested that this may be incorrect. It would be unfair to accuse Fodor of lacking a fully principled method of determining when and when not to assume adaptationism a priori. For the lack of such a method is a quite general problem in evolutionary biology; it is not specific to theorizing about the evolutionary function of cognition. But it does mean that knockdown arguments of the sort Fodor tries to run are unlikely to be successful.

At this point it is worth returning briefly to an issue discussed in Section 3. Fodor argued, to recall, that since small neurological changes could have caused abrupt changes in cognitive structure, we do not know that the continuity requirement on Darwinian explanations was satisfied, hence should not assume that the complexity of the human mind shows it to be an adaptation. Against this argument I noted that biologists rarely if ever know that continuity was satisfied, but rather assume that it *must* have been satisfied, if a trait shows clear evidence of having been fashioned by cumulative selection. Now Fodor *himself* adopts this methodology, implicitly, when he defends the inference from innate contingent knowledge to adaptationism. For he adduces no evidence, either direct or indirect, for the existence of a parameter whose gradual modification could have led from ancestral minds, without an endowment of innate contingent knowledge, to modern minds, with such an endowment. Nor does Fodor even offer any speculations on what the intermediate stages might have been like. He simply assumes that the continuity requirement must have been satisfied, for otherwise natural selection could not have done the work which we can see that it has done. In my view this assumption is perfectly acceptable. Fodor's official view—that a priori adaptationism is only legitimate where the continuity requirement is antecedently known to be satisfied—is belied by his own practice, and rightly so.

6. Is the Language Module an Exception? I turn to the final twist in Fodor's story. Having argued that modules which contain innate contingent information must have been fashioned by natural selection, there being no other way the information could have got there, Fodor immediately claims that there is an exception to this general principle: the language-learning module. In this special case, the inference to Darwinism doesn't work. It is "an irony of the history of cognitive science," he writes, that "knowledge of natural language, which was the first and is still perhaps the best candidate for being a module, happens to be thoroughly atypical of the usual relation between innate content and natural selection" (2000, 94).

What is so different about the language module? According to Fodor the difference is this. Like other modules, the language module contains a store of innate contingent information: on the standard Chomskian view which Fodor accepts, it contains innate knowledge of the principles of universal grammar. As before the question arises: how did the innate knowledge get there? But in the language case, the answer is different.

There is no need to invoke an instructional mechanism by which the world can shape the content of our innate beliefs. The reason, says Fodor, is that "the facts that make a speaker/hearer's innate beliefs about the universals of language true (or false) *aren't* facts about the world; they're facts about the minds of the creature's conspecifics" (2000, 95; emphasis in original.) One's innate beliefs about linguistic universals will be true so long as one's conspecifics have the very same beliefs, and that is guaranteed by the fact that the genes which determine those innate beliefs are common to all conspecifics. In Fodor's words: "in special cases like language,[6] what makes one's contingent beliefs true is that they are about the minds of creatures whose innate cognitive capacities are determined by the same genetic endowment that determines one's own" (p.95). Therefore "there is no particular need for what the language organ believes to have been shaped by natural selection" (2000, 96). Fodor thus endorses the view Chomsky has long defended: the language faculty is both innate and modular, but not an adaptation.

This is an intriguing argument, but I do not think it is correct. Fodor is right, of course, that the facts which make a creature's innate beliefs about linguistic universals true are facts about the minds of its conspecifics. For example, if I innately believe that all rules for sentence formation are structure-dependent, my belief will be true so long as all humans have the same innate belief. Whereas if I innately believe that the auditory location of a sound source predicts its visual location, my belief will not be true simply if all my conspecifics share the same belief—its truth depends on the environment being a certain way. But Fodor is wrong to see this difference as relevant to the issue of adaptationism. For from an evolutionary point of view, an organism's conspecifics are *part* of its environment, indeed a crucial part. Facts about the cognitive and behavioral traits of conspecifics are no less facts about an organism's environment than are facts about the physical features of its habitat. (See Lewontin 1985b for a famous elaboration of this point).

This objection to Fodor's argument may sound merely terminological, but it is not. For once we accept that the cognitive traits of conspecifics form part of an organism's environment, it becomes clear that innate knowledge of language is an instance of a very well-known evolutionary phenomenon: frequency-dependent fitness. This occurs when the fitness of a given phenotype depends on how many other organisms in the population have the same phenotype. Many interesting biological traits exhibit frequency-dependent fitnesses, and there is a well-established body of the-

6. Fodor suggests that the theory-of-mind module may also be a "special case," i.e., an exception to the principle that innate contingent knowledge must have been put in the mind by natural selection, for the same reason.

ory—evolutionary game theory—devoted to analyzing the consequences. The classic example is the hawk-dove game (Maynard Smith 1982). In one-on-one contests with conspecifics for a scare resource, animals have two possible strategies: fighting (playing hawk), or backing down (playing dove). If the population contains mostly doves, a hawk will be at a great advantage—it will usually be pitched against doves, so will usually win the contest without having to fight. But if the population contains mostly other hawks, then a hawk will usually find itself pitched against other hawks, so will often end up in fights; in such a situation, doves may be at an advantage. The fitness of the strategy "playing hawk" thus depends on how many other animals adopt that same strategy, and likewise for "playing dove." Both of these behavioral phenotypes exhibit (negative) frequency-dependent fitness.

Now consider the language module. As we have seen, a creature's innate beliefs about linguistic universals are only true if other creatures have the same beliefs. Now in general, innate beliefs will only confer a fitness advantage if they are true, as Fodor stresses (2000, 95). It follows that innate beliefs about linguistic universals exhibit (positive) frequency-dependent fitness: they benefit an organism greatly if they are common in the population, but little at all if they are rare. This immediately calls into doubt Fodor's argument that natural selection is not required to explain the innate beliefs in the language module. For if this were correct, then by parity of argument it would follow that selection is not required to explain the very phenomena that evolutionary game theory was explicitly designed to explain! If Fodor were right, then any trait which has gone to fixation in a population, and whose fitness exhibits positive frequency dependence, would not be in need of a Darwinian explanation. But this conclusion is most implausible, given that a well-articulated body of evolutionary theory is devoted to providing just such explanations. Hence Fodor must be wrong.

That Fodor is wrong can be shown in a more direct way, by examining his attempted nonselectionist explanation of how our minds came to have innate true beliefs about linguistic universals. His explanation, to recall, is simply that qua conspecifics we all share the genes that code for these innate beliefs, and the beliefs are such that if everybody has them, they will be true. But this does not address the question of how those genes came to be fixed in the species in the first place. What needs explaining is the fact that, from an ancestral condition in which humans or proto-humans didn't have the innate beliefs in question, there evolved humans who did have those beliefs. Obviously, it is grossly implausible to suggest that a huge number of mutations of identical effect simultaneously occurred in the population. Much more likely, one or a few mutations arose which coded for the beliefs, and they spread to fixation in the population—

either because of natural selection, genetic drift, or some other cause. Whatever accounts for the spread, one point is clear. We cannot obtain a plausible account of the phylogeny of the language module if we simply take as our starting point the ubiquity of the relevant genes in the species. For how those genes came to be ubiquitous, from an ancestral state in which they were not ubiquitous, is precisely what we would like to know. Fodor's nonselectionist explanation of the innate knowledge in the language module is thus no explanation at all.

I conclude that Fodor's argument for the language module being an exception to his general principle about the relation between innate contingent content and natural selection is not compelling. Adaptationism about the language module is neither more nor less plausible that adaptationism about other modules which contain innate contingent content.

7. Conclusion. My analysis shows that some of Fodor's arguments succeed while others do not. His view that the cognitive mind is as likely to have arisen by sudden saltation as by gradual directional selection is not convincing, for it depends on the untenable principle that gradualist Darwinian explanations are only acceptable where we have antecedent knowledge that the continuity requirement is satisfied. In any case, even if cognition did arise by sudden saltation, natural selection very likely explains its spread and maintenance. By contrast, Fodor's attack on the Cosmides and Tooby claim that selection would have favored a modular over a nonmodular cognitive architecture is largely successful. Fodor's claim to defend the inference from massive modularity to adaptationism is a mischaracterization: what he really argues is that there is a good inference from innate contingent knowledge to adaptationism. This argument is convincing, though its restriction to contingent knowledge is debatable. Finally, Fodor's view that language is a special case is not compelling: the difference he highlights between innate knowledge of language and other innate contingent knowledge does not have the relevance he claims for it. Overall, despite his searching critique, Fodor has not shown that it is intellectually unrespectable to adopt a broadly Darwinian approach to cognition and language.

REFERENCES

Cosmides, Leda, and John Tooby (1992), *The Adapted Mind: Evolutionary Psychology and the Generation of Culture.* Oxford: Oxford University Press.

Cosmides, Leda, and John Tooby (1994), "Origins of Domain Specificity: The Evolution of Functional Organization", in Lawrence A. Hirschfeld and Susan A. Gelman (eds.), *Mapping the Mind: Domain Specificity in Cognition and Culture.* Cambridge: Cambridge University Press, 85–116.

Dawkins, Richard (1996), *Climbing Mount Improbable.* Oxford: Oxford University Press.

Fodor, Jerry (1983), *The Modularity of Mind.* Cambridge, MA: MIT Press.

——— (2000), *The Mind Doesn't Work That Way.* Cambridge, MA: MIT Press.

Lewontin, Richard (1985a), "Adaptation", in Richard Levins and Richard Lewontin (eds.), *The Dialectical Biologist.* Cambridge, MA: Harvard University Press, 65–84.

——— (1985b), "The Organism as the Subject and Object of Evolution", in Richard Levins and Richard Lewontin (eds.), *The Dialectical Biologist.* Cambridge, MA: Harvard University Press, 85–106.

Maynard Smith, John (1982), *Evolution and the Theory of Games.* Cambridge: Cambridge University Press.

——— (1989), "Current Controversies in Evolutionary Biology", in *Did Darwin Get it Right?.* New York: Chapman and Hall, 131–147.

Mithen, Steven (1996), *The Prehistory of the Mind.* London: Thames and Hudson.

——— (2000), "Mind, Brain and Material Culture: An Archaeological Perspective", in Peter Carruthers and Andrew Chamberlain (eds.), *Evolution and the Human Mind: Modularity, Language and Meta-Cognition.* Cambridge: Cambridge University Press, 207–217.

Pinker, Steven (1997), *How the Mind Works.* New York: W. W. Norton.

Plotkin, Henry (1997), *Evolution in Mind.* London: Alan Lane.

Samuels, Richard (2000), "Massively Modular Minds: Evolutionary Psychology and Cognitive Architecture", in Peter Carruthers and Andrew Chamberlain (eds.), *Evolution and the Human Mind: Modularity, Language and Meta-Cognition.* Cambridge: Cambridge University Press, 13–46.

Segal, Gabriel (1996), "The Modularity of Theory of Mind", in Peter Carruthers and Peter Smith (eds.), *Theories of Theories of Mind.* Cambridge: Cambridge University Press, 141–157.

Tooby, John, and Leda Cosmides (1995), Foreword to *Mindblindness* by Simon Baron-Cohen. Cambridge, MA: MIT Press, xi-xviii.