

This work originally appeared in:

Dennett, Daniel. 1991. Granny's Campaign for Safe Science. In *Meaning and Mind: Fodor and his Critics*, ed. Barry Loewer and Georges Rey, p.p. 87-95. Oxford: Blackwell.

It is not available electronically from the publisher.

This is Daniel C. Dennett's final draft before publication. It has been modified to reflect the pagination of the published version of the work.

**MEANING IN MIND**  
Fodor and his Critics

Edited by  
Barry Loewer and Georges Rey

BLACKWELL

---

## Granny's Campaign for Safe Science

By DANIEL C. DENNETT

What is the thread tying together all of Jerry Fodor's vigorous and influential campaigns over the years? Consider the diversity of his *betes noires*. In "Operationalism and ordinary language," (1965) it was Wittgenstein and the "no private language" gang; in *Psychological Explanation* (1968d) and *The Language of Thought* (1975a), it was Ryle, Skinner and other behaviorists; in "Tom Swift and his procedural grandmother" (1978b, reprinted in 1981) it was AI in general and procedural semantics in particular; in "Three cheers for propositional attitudes" 1978c, reprinted with revisions in 1981) it was me and my "irrealist" way with stances; in "Methodological solipsism considered as a research strategy in cognitive science" (1980c, reprinted in 1981) it was the brand of "naturalism" that claimed that psychology had to traffic in meanings that were not inside the head; in *The Modularity of Mind* (1983) it was Bruner and the other New Look psychologists who infected perception with thought, but also, in the shocking punch line of the last chapter, AI again; in *Psychosemantics* (1987d) it was the meaning holists and those who would ground their naturalistic appeal to teleological formulations in what Fodor elsewhere has called "vulgar Darwinism" (these villains take another drubbing in his forthcoming "A Theory of Content" (in press-f); and in "Connectionism and cognitive architecture: a critical analysis", (Fodor and Pylyshyn, 1988a) it is the connectionists and their many friends.

What do these various heresies have in common? From Fodor's point of view, two things, obviously: (1) they are all wrong, wrong, wrong! and (2) they are endorsed by people who are otherwise quite decent company. That would be thread enough to tie Fodor's targets together if he were right, but as one who finds more than a morsel of truth in each of the derided doctrines, I must seek elsewhere for a uniting principle, and I think I have found it: they are all doctrines that make Fodor's Granny exclaim "Well I never!" and lurch alarmingly in her rocker. The cat is out of the bag. Jerry Fodor is a Granny's boy, a romantic conservative whose slogan might well be 'What is good enough for Granny is good enough for science.'

Now I don't know Jerry's real Granny from Marlene Dietrich, but his notional Granny, of whom we have all heard, believes in beliefs, in thoughts, in the genuine, intrinsic content of mental states.<sup>1</sup> In short, Granny believes in minds and mental events with all her heart and soul. And Jerry has eagerly sought to defend her: contrary to certain sophisticated opinions, Science does not suggest

that she is wrong; we can have Granny and Science too! Cognitive Science is - had better be - the vindication of Granny in spades.

Clever boy that he is, Fodor has seen the futility of other heroic defenses of Granny against the insidious march of Science: the crypto-dualisms of Elizabeth Anscombe, John Searle, Thomas Nagel, and Charles Taylor, for instance. He, like Noam Chomsky<sup>2</sup>, has had a subtler game plan: saving Granny's Cartesianism by outsciencing Science, by turning Granny's cherished views into redoubtable science itself: Cognitive Science. He is happy to call himself a mentalist, but he's the "thoroughly modern mentalist."

As he says in *The Language of Thought*:

Contemporary cognitive psychology is ... by and large conservative in its approach to the commonsense tradition .... Cognitive psychologists accept, that is, what the behaviorists were most determined to reject: the facticity of ascription of propositional attitudes to organisms and the consequent necessity of explaining how organisms come to have the attitudes to propositions they do. What is untraditional about the movement ... is the account of propositional attitudes that it proposes: ... having a propositional attitude is being in some computational relation to an internal representation.

(p. 198.)

As a bit of sociology of science, this is egregiously tendentious; and ever since, Fodor has been hard pressed to insist that you can't be a proper cognitive scientist unless you accept the "facticity" of propositional attitudes. This comes out most clearly, perhaps, in his recent broadside against the connectionists, who are, by his lights, enemies of cognitive science precisely because they don't accept the facticity of the "classical" mental types and processes.

Fodor is justly renowned as a philosophical champion of cognitive science, but if you look closely you will note that he is its champion only so long as it hews to the traditional line. Any breaking ranks on facticity earns an immediate reprimand from him. If he has to choose between Granny and science, he has made it clear that he'll choose Granny, and in several recent publications (see especially Fodor, 1985b), he has been candid in acknowledging the contortions he is willing to submit himself to in order to hold the faith.

Now that I have discovered the secret thread, I will use it to conduct a tour of recent history, following its winding path through labyrinths of controversy.

In one of the most often-quoted passages in contemporary philosophy, Quine declared a watershed:

One may accept the Brentano thesis either as showing the indispensability of intentional idioms and the importance of an autonomous science of intention, or as showing the baselessness of intentional idioms and the emptiness of a science of intention. My attitude, unlike Brentano's, is the second.

(1960:221)

There has been a growing consensus in philosophy of mind since then (for some details, see Dennett, 1987) that, one way or another, one must accept Quine's message and opt for one of the two choices: championing a non-physicalistic, non-naturalistic "autonomous science of intention" (something akin to Cartesian dualism), or siding with the natural sciences at the cost of taking propositional attitudes less than entirely seriously. Fodor has always thought

otherwise. He has tried to show that you can have what he calls intentional realism (taking Brentano's intentionality entirely seriously) and make it live with the rest of physical science: scientific mentalism, in short. This is a delicate balancing act indeed, and if cognitive science is to fulfill this role, it must be carefully protected from variations that slide off either into anti-scientific mentalism (dualism, and other bizarre doctrines such as Searle's) or into scientific anti-mentalism (eliminative materialism, behaviorism, instrumentalism, etc.).

Fodor has revealed in his iconoclasm: in Fodor (in press-f) he puts it this way: "it counts as conventional wisdom in philosophy that (i) the intentional/semantic predicates form a closed circle and (ii) intentional states are intrinsically holistic." The conjunction of these theses he sees as fatal to "naturalism," and he importunes us to "become cautious about taking intentional irrealism for granted; more cautious, at a minimum, than has been the philosophical fashion for the last fifty years or so." - a rough doubling of the actual time span, but perhaps that is the way it has seemed to Jerry, as he has cast about for a stable version of realism.

In Fodor (1968d) he resolutely turned his back on Quine, Ryle, and Wittgenstein and began sketching an account of mental events as inner processes, distinguished or identified by their functional properties. It looked, and was supposed to look, like quite a radical alternative. It appeared to be a way of undercutting what seemed at the time to be a stifling Behaviorist dogma: what goes on inside doesn't settle anything. One way of understanding the limits that Fodor imposes on his own theorizing (and that of anyone else who will listen) is that it must come out in the end that something about what is going on inside the head must trump the "behaviorists" criteria, their principles of charity, their procedural semantics, their skepticism about meanings. That is what it means for content to be Real.

Here is how Fodor put it at the time:

To qualify as a behaviorist in the broad sense of that term that I shall employ, one need only believe that the following proposition expresses a necessary truth: For each mental predicate that can be employed in a psychological explanation, there must be at least one description of behavior to which it bears a logical connection.

(Fodor, 1968d:51)

A mentalist is simply anyone who denies this (Fodor, 1968d:51), and ever since then Fodor has been seeing to it that he, and others, resist the siren song of this proposition. In his first book we get his discussions of function that make him one of the original functionalists - and it is easy to overlook the fact that nowhere does he tie the content of mental states to their functions, though some passages certainly seem to permit it. He says:

The functional character of a machine state is determined by its total role in the machine's computational processes as well as by its relation to machine behavior in the narrow sense of machine output.

(Fodor, 1968d:143)

Hence holism is true of functional attributions, but he doesn't explicitly relate the content of a state to its functional character. He does, however, go on to say:

It is a question of fact, and not of how we should decide to talk about the facts, whether, or to what extent, a given kind of organic behavior is the consequence of

psychological processes that are functionally equivalent to some machine processes ... In the present case, we demonstrate functional equivalence between machine and organic processes by determining that the psychological theory realized by the machine's program provides an adequate and simple account of the organism's behavior.

(Fodor, 1968d:149)

If "demonstrating functional equivalence" is to be the touchstone of a psychological theory, as this passage suggests, one would suppose that "content equivalence" would go along with it, but Fodor refrained from explicitly making this step, unlike such functionalists as Sellars and me, and only recently has he explicitly rejected this step (in Fodor, 1987d). There are two brands of functionalism, he says, and he endorses - has always endorsed, apparently - only the weaker version:

... all you need is the claim that being a belief is a matter of having the right connections to inputs, outputs, and other mental states. What you don't need - and what the philosophical motivations for Psychofunctionalism therefore do not underwrite - is the much stronger claim that being the belief that *p*, being a belief that has a certain content is a matter of having the right connections to inputs, outputs, and other mental states.

(Fodor, 1987d:69)

So Fodor's functionalism is eviscerated; he disagrees with those of us, from Sellars to the present, who in one way or another take "functional role semantics" to be the chief beauty of functionalism. I had argued (in Dennett, 1969) that content was a function of function, and this, it turns out, is the chief point Fodor and I have covertly disagreed about ever since. For Fodor, content is not a function of function after all. Or at any rate, it is not much of a function of function - for he does grant that functional role "is a marginal - a not very important - determinant of meaning" (Fodor, 1987d:72).

What, then, is content a function of? This has been the nagging problem for Fodor. At one point, during the early days of the Language of Thought, his friend and collaborator Zenon Pylyshyn could jokingly remind us of the lady who said that the earth rested on the back of an elephant who stood on a turtle, and when asked what the turtle stood on, replied "It's turtles - all the way down!" For Jerry, said Zenon, "it's syntax - all the way down!" That, we can presume, was not to be taken entirely seriously, but then what did Jerry himself say content was a function of? He didn't say. He did, however, have a lot to say about what wouldn't work: procedural semantics.

Fodor's first broadside against AI, in Fodor (1978b) set the tone and method for his later attacks: a priori "refutations" of what might charitably be called rational reconstructions or idealizations of the actual claims of his targets. Unlike Dreyfus's parallel campaign, Fodor's has never been particularly concerned with the details of the models under attack (though he has been happy to endorse the particular criticisms developed by others), but has rather invoked and defended "principles" that, he claimed, were ineluctably violated by any and all versions of the heresy. For instance, he has argued that AI can be seen, when you look at it just right, to be just warmed-over "associationism" or "phenomenalism" or some other ism that philosophers decisively refuted ages ago. This has not set well with

his targets, and it is interesting to compare their complaints, early and late. Johnson-Laird (1978) and Smolensky (1988c) have both charged that Fodor simply invented and demolished strawmen, and have amply defended their charges.

This "principled" turning of his back on AI has always baffled me. For someone who has claimed that to believe is to be in a computational relation to a representation, he has been remarkably uninterested in actual computational relations. Why, if one was sure that the road to the scientific salvation of mentalism was via a language of thought, would one be so unwilling to consider in detail the actual attempts to construct such languages of mental representation, by people working in what Haugeland (1985) calls GOFAI (Good Old Fashioned AI)? If their models were faulty, why not try to devise a better one? Because, Fodor apparently believed, the whole enterprise of GOFAI was ill-founded. What to put in its place? One might have supposed that Fodor would look with interest, if not outright enthusiasm, at the radical alternative now brewing: connectionism. The other arch-critic of GOFAI, Dreyfus, has done just that. But Fodor has declared that enterprise to be, if anything, even more severely benighted, in all possible versions.

That would not seem to leave much room to maneuver for the serious scientist of mental representations, and that hunch is confirmed by Fodor, in his most sweeping attack on AI, in Fodor (1983). Here we find an a priori argument from first principles that supposedly spells doom:

- 1 "The condition for successful science (in physics, by the way, as well as psychology), is that nature should have joints to carve it at: relatively simple sub-systems which can be artificially isolated and which behave, in isolation, in something like the way they behave in situ. Modules satisfy this condition ..." (Fodor, 1983:128).
- 2 But "belief-fixation," the basically analogical and holistic process that occurs in the "central system" inboard of the modules, lacks this condition for successful science (p. 128).
- 3 Therefore, there is no possibility of discovering psychological laws governing the processes of belief-fixation, no possibility of a science of the central mind at all.

So it is no wonder Fodor has no use for either connectionism or GOFAI; he has an argument to prove not only that AI, in all its forms, is impossible, but that no conceivable successor science could do any better! "I am suggesting that there is a good reason why nothing is known about [the neuropsychology of thought] - namely, that there is nothing to know about it" (Fodor, 1983:128).

Safe science is "classical" science, Fodor has always insisted, but with Granny urging him on, he admits that the safest science is no science at all. One can see signs of this underlying scientific nihilism in other passages in Fodor's work. Most striking, perhaps, is the gag with which he ends "Methodological Solipsism":

My point, then, is of course not that solipsism is true; it's just that truth, reference and the rest of the semantic notions aren't psychological categories. What they are is: they're modes of Dasein. I don't know what Dasein is, but I'm sure that there's lots of it around, and I'm sure that you and I and Cincinnati have all got it."

(Fodor, 1980c:71)

And in Fodor (1985b) he recognized what he calls the "idealization problem" - the principle of charity in disguise - as an unsolved problem and admitted he saw "no reason to assume that the problem can be solved" within his Realistic boundaries. That paper ends with the following observation. "But of the semanticity of mental representations, we have, as things now stand, no adequate account."

It was into this breach that Fodor launched *Psychosemantics* (1987d), in which he attempted to hang on to his realism by adopting a denotative semantics: "concepts are individuated by reference to the properties they express, thoughts by the states of affairs they correspond to, and so forth." Independently of any functional role that a concept or mental representation might play within an interanimated set of its fellows, it can, as a matter of brute metaphysical fact, denote (or express) a property, thanks to its causal (but not functional) link to that property.

So if Fodor's hope for a Realistic but still Naturalistic theory of content is to be fulfilled, it must be fulfilled by an account of content that meets some strict requirements: a theory of content has to be (1) "factive" - not a matter of interpretation; (2) "naturalistic" - but only in the sense that it is physicalistic and atomistic. It avoids "meaning holism," and the reason it must avoid meaning holism is that meaning holism would prevent the existence of psychological laws, and without laws, there is no proper Science. But this means that at least the content part of the theory cannot be functionalistic, because holism reigns for functional attribution. That is the message from Darwin.

This means that Fodor cannot avail himself of any teleological formulations, and this threatens to put him in mysterious waters indeed - for it is going to be passing strange if biological creatures have come to rely on Dasein for their psychology.

In Dennett (1987, p. 308) I offered a diagnosis of Fodor's puzzling antipathy to Darwin: he "sees that the most one can ever get from any such [Darwinian] story, however well buttressed by scrupulously gathered facts from the fossil record, etc., is a story with ... the potential for indeterminacy." He has subsequently confirmed this diagnosis, and gone on the attack, in Fodor (in press-f). What is particularly interesting about his attack is that it brings the conservatism of his Intentional Realism into even sharper focus than in Fodor (1987d). Here is what he says:

contrary to advertisements you may have seen, the teleological story about intentionality does not solve the disjunction problem. The reason it doesn't is that teleological notions, insofar as they are themselves naturalistic, always have a problem about indeterminacy just where intentionality has its problem about disjunction.

(Fodor, in press-f)

Let us consider, as so often before, the frog with its bug-detecting eyes and responsive tongue. Fodor, criticizing efforts by Israel (unpublished) and Millikan (1984) to explain the content of the frog's perceptual states, points out that while one evolutionary/optimal story explains that the frog's ON state is about flies, "there is nothing to stop you from telling the story quite a different way. On the alternative account, what the neural mechanism is designed to respond to is little

ambient black things (or, *mutatis mutandis*, characteristic patterns of ocular irradiation as of little ambient black things)." And, Fodor claims, "Darwin doesn't care which of these ways you tell the teleological story" (in the press-f, p. 17).

He is almost right. A Darwinian story is quite capable of distinguishing, as advertised, between large and interesting classes of errors and the "proper" functioning of those detectors. One does it by considerations of cost-effectiveness, as I outline in Dennett (1987:290-305). Such considerations do indeed depend on a certain "interest-relative" tactic of interpretation - roughly, the decision to ignore nitpickers. Hence, such considerations will not serve to rule out heroically Pickwickian interpretations, e.g., the sort of phenomenalist interpretation that insists that the frog lives its whole life caring only about "little ambient black things" or its ocular irradiation patterns and other states of its receptors. We human beings are ever so much more particular about the objects of our intentional states, but in the limit the same conclusion holds for us: Darwin doesn't care whether we grew up on Earth or Twin Earth.

That is to say, one can tell a purely phenomenalist story (for instance) about the frog or about us; "there is nothing to stop you," as Fodor says, but also, there is nothing to recommend it. Ever since Descartes we have had to live with the startling but no doubt negligible fact that a human mind well-equipped to live in the real world is, by definition, equally well-equipped to live in the fantasy world conjured up by the evil demon. It would be a wonder if Darwin's vision gave us a scientific talisman with which to ward off the evil demon.

So in a way Fodor is right: the resources of the theory of natural selection cannot provide a foundational, interpretation-free criterion of real content - of what the frog's states are really about. Recall Fodor's early definitions of behaviorism and mentalism quoted above: a behaviorist believes the following is a necessary truth, and the mentalist denies it: for each mental predicate that can be employed in a psychological explanation, there must be at least one description of behavior to which it bears a logical connection. The theory of natural selection is, in Fodor's terms, behavioristic - beyond a certain point, as Fodor says, "Darwin doesn't care" which way we tell the story. So Fodor's attack on the teleological/functional story convicts it of only what I have been insisting upon all along: there is no determinacy of function to be obtained from evolutionary considerations.

But so what? Only if you believe, with Granny, that there is an ultimate fact of the matter should this be an objection. That is, Granny believes that the following situation is (always?) logically possible:

according to our canvassing of the frog and its lifeworld, all considerations of efficiency and cost-effectiveness (all teleological considerations, that is) weigh in favor of attributing to it the belief that p, but in fact the frog believes something else.

"Behavioristic" theories such as mine (or such as any properly naturalistic theory that bases content ascription on function rather than mysterious denotational power) simply deny this possibility.

I fear this will shock Granny right off her rocker, but I would try to break it to her gently by pointing out all the interesting scientific problems still around to solve once we give up the defense of Intentional Realism as a lost cause. Besides, I would add, there is something ludicrous about the spectacle of her grandson trying to dissuade - one might as well say prohibit - scientists from attempting

empirical explorations of models on the grounds that he has an a priori proof that all such models are hopeless. This tactic has been tried before by philosophers, and it has never washed well. In the past, it has been philosophers' categories that have typically been shown not to carve (models of) nature at the joints, and one way or another, the banished models have been seen to leak back through the defenses of the censors. So on tactical grounds alone (if not because he has undergone a theoretical conversion), Granny should advise her hero to drop it.

Let me summarize the results of this survey of Fodor's campaigns over the years. He has, in fact, had a quite single and steady vision, which he has defended against all comers: the "mentalism" he defined in 1968 in opposition to "behaviorism." But as various inroads have been made in cognitive science, threatening to establish one "logical connection" or another between "behavior" (including internal, neural behavior) and mental predicates, Fodor's mentalism has been driven back, back, back, to the extremely conservative and well-nigh mystical position that he currently defends:

- (i) no science of the "central system" is possible;
- (ii) content is real, and determinate, but strictly independent of both behavior and biological function (to the extent that that is determinable).

Granny is no doubt quite pleased with this development, but while she and Jerry wait around for the day that scientists give up trying to understand the mind (when they get to say "I told you so!"), the rest of us have our own retort to fling: if we're so wrong and you're so right, how come we're advancing and you're retreating?

## NOTES

- 1** Many of the ideas herein were provoked/inspired/unearthed by Kathleen Akin (1988), and in discussions with me during the last three years.  
Jerry's Granny is not to be confused with his Aunty, who "speaks with the voice of the Establishment" (1987d:135), and turns out to be a "New Connectionist Groupie" (1987d: 139).
- 2** Chomsky has often expressed the view that if scientific psychology can't be "like physics," we would do better to look to novels for our knowledge of human thinking. A relatively early expression of this view was in 1978, recounted by me in Dennett (1988a: 285). He has recently reiterated the view: "Thus it is quite possible - overwhelmingly probable, one might guess - that we will always learn more about human life and human personality from novels than from scientific psychology" (Chomsky, 1988: 159).
- 3** See Dennett (1988b:384-9) for further discussion of the difficulties with Fodor's denotative semantics.



Jerry Fodor and his Granny