

[Published in *Mind* 110, April 2001]

Concepts: Where Cognitive Science Went Wrong, by Jerry Fodor. Oxford: Oxford University Press, 1998. Pp. xii + 174.

Concepts, the 1996 John Locke Lectures, synthesizes and develops Fodor's views on the eponymous topic. It's immensely stimulating. Anyone working in the area will need to study its trenchant critical discussion of key positions in philosophy, linguistics, and psychology. These readers will be rewarded as well by the book's many illuminating asides and its more constructive closing chapters. With its wealth of ideas and enjoyably Fodorian prose, *Concepts* auspiciously inaugurates the Oxford Cognitive Science Series. Oxford University Press is also to be commended for making *Concepts* immediately available in paperback, though it contains far too many typos.

Fodor's tale is told against the backdrop of the Representational Theory of Mind, according to which to have a thought is to stand in a certain functional relation to a token mental representation composed of concepts, and to think is to perform content-respecting computational operations over such representations. Concepts clearly are a crucial part of this view. So someone committed to it will want a compelling account of what concepts are, how they acquire their content, what relations they bear to one another, etc. According to Fodor, most cognitive scientists hold in some form or another that a concept's content is in part constituted by inferential relations defined over the concept's constituent structure—with the exception of a small set of *primitive* concepts required to avoid non-well-foundedness. But, he argues, the most prominent and promising such theories—that concepts are definitions, that they are prototypes, that they are nodes in a theory—are fatally flawed. Note that on this account the received view has two parts: (1) a concept's content is in part constituted by inferential relations involving other concepts (Inferential Role Semantics, or IRS), and (2) those relations are defined over concepts' constituent structure. Fodor also briefly considers—and rejects—a version of IRS *sans* (2), one that appeals to meaning postulates instead of positing constituent structure among concepts. (A concept is structured iff it contains a concept as a proper constituent or part. What it is for one mental representation to contain another as a constituent or part is not here explored. Presumably, the possibility of IRS *sans* (2) rules out an inferential explication. Given Fodor's token physicalism with respect to mental representations (p. 10), it's possible that "contain", "part", and the like are to be taken in their primary, or literal, spatial senses. Cf. Fodor's use of "literal" on p. 111.)

Fodor suggests that we reject both (1) and (2). There is indeed a potentially infinite number of structured concepts to which an inferential story applies: we must accommodate our potentially infinite conceptual capacities. But practically all *lexical* concepts (a) lack structure and (b) possess their content independently of their relations to other concepts—where a lexical concept is, roughly, what a morpheme expresses. (The "practically all" qualification covers *inter alia* at least some logical concepts towards which Fodor also favors an inferential approach. See his *The Elm and the Expert*, Cambridge, MA: MIT Press, 1994, pp. 71-75.) Fodor labels thesis (a) "conceptual atomism" and supplements (b) with "informational semantics," the claim that the content of these concepts is constituted by nomic mind-world relations. (Informational semantics is assumed here; Fodor defends his particular "asymmetric dependency" version in "A

Theory of Content” in his *A Theory of Content and other Essays*, Cambridge, MA: MIT Press, 1990, pp. 51-136.) Fodor calls the conjunction of these two claims “informational atomism.”

Here, in outline, are Fodor’s reasons for rejecting the two IRS alternatives he discusses at greatest length. The first claims that lexical concepts are definitions—that is, are composed of constituent concepts such that falling under their conjunction provides a necessary and sufficient condition for falling under the concept they compose. Fodor objects that (1) there are practically no examples of successful definition around; (2) the claim runs afoul of Quine’s critique of the analytic/synthetic distinction; and (3) definitions seem to play no role in explaining the acquisition or cognitive deployment of concepts—for example, one might expect that more complexly defined lexical concepts, assuming there are such, ought to be harder to learn and to apply, but no supporting empirical evidence has been found. Fodor also responds to some putative reasons empirical and philosophical *for* positing definitional structure. Lexical semanticists, for example, have argued that knowledge of *syntax* is acquired in part through “bootstrapping” from knowledge of semantical features of the lexicon: crudely put, by knowing that ‘eat’ denotes an action, one comes to know that it’s a verb. (Cf. Steven Pinker, *Language Learnability and Language Development*, Cambridge, MA: Harvard University Press, 1984.) This might provide empirical grounds for positing word definitions, which in turn would supply reason for thinking that the concepts words express have definitions. But Fodor rejoins *inter alia* that evidence of such semantical *features* falls far short of establishing lexical *definitions*. Fodor’s response to the philosopher’s thought that definitions are necessary to explain our intuitions of analyticity, entailment, and the like is to explain away the intuitions instead, appealing as appropriate to Quinean centrality (the epistemic dependence of many other beliefs on the one at issue) and Putnamian mono-criteriality (the epistemic dependence of some concept’s application on a single criterion). (See, respectively, W. V. O. Quine, “Two Dogmas of Empiricism,” *Philosophical Review* **60**, pp. 20-43; and Hilary Putnam, “The Analytic and the Synthetic” in H. Feigl and G. Maxwell (eds.), *Scientific Explanation: Space and Time*, Minnesota Studies in the Philosophy of Science **3**, Minneapolis, Minn.: University of Minnesota Press, 1962.)

The second IRS approach holds that lexical concepts are prototypes: roughly, they are composed of concepts such that if something falls under (a sufficient number of) these constituents, it’s likely to fall under the concept they compose. According to Fodor, while there’s strong reason to think that many concepts are *associated* with prototypes, the view that lexical concepts *are* prototypes runs afoul of compositionality, the thesis, roughly, that “the syntax and the content of a complex concept is normally determined by the syntax and content of its constituents” (p. 94). (“Normally” accommodates idioms. “Roughly” marks not only the fact, noted by Fodor, that attempts to render such intuitive characterizations rigorous have yielded trivially satisfiable results, but also issues concerning the relevant strength of determination and the proper accommodation of context-sensitivity.) Compositionality explains productivity (our potentially infinite conceptual capacities) and systematicity (if you can think *Fa*-thoughts and *Gb*-thoughts, you typically can think *Fb*- and *Ga*-thoughts, etc.), as well as numerous specific features of our cognitive economy. But, argues Fodor, prototypes don’t compose. For example, there seems no obvious way to get from the prototypical *fish* and the prototypical *pet* to

the prototypical *pet fish*. (There are, however, various on-going attempts to reconcile such problem cases with compositionality. One recent attempt Fodor discusses is Hans Kamp and Barbara Partee, “Prototype Theory and Compositionality,” *Cognition* 57, 1995, pp. 129-91.) Further, many complex concepts, such as NOT A CAT, seem to lack prototypical features altogether.

Does it follow that such lexical concepts as FISH, PET, and CAT are not prototypes? Briefly sketching some lines of response in the negative will help bring out the importance of Fodor’s commitment to a controversial claim much stronger than compositionality. The conclusion doesn’t follow if *complex* concepts—i.e., the concepts composed of lexical concepts—aren’t prototypes. So, a first simple response is to note that if the content of complex concepts is given in some other way, it remains open that the content of the lexical concepts plus their syntax determines *that*. But suppose there’s no promising theory available in this direction (we’ll see that the stronger claim will preclude one, in any case), and we concede that lexical concepts are not prototypes. Still, the possibility remains open that, although lexical concepts aren’t *simply* prototypes, that’s *part* of what they are. Lexical concepts could have, say, two aspects: a prototype and something further. Perhaps, even, this something further could be precisely what Fodor takes practically all lexical concepts to be *in toto*—viz., unstructured mental particulars that expresses properties. Lexical concepts on this amalgamated view would then be structured, since they would contain a part—the prototype—that, moreover, has concepts as parts. And, if compositionality is satisfied on Fodor’s own account, it would *ipso facto* be satisfied here too. The position at which we’ve arrived may seem unattractive. According to it, lexical concepts are in part prototypes while complex concepts aren’t even prototypes in part, although some of them are *associated* with prototypes: why this bifurcation, why should it be that some concepts *are* in part prototypes and some are only *associated* with them? (A parallel objection would face versions of the simple response above.) Well, perhaps just as we’re considering the possibility that lexical concepts might be an amalgamation of a prototype and an unstructured mental particular, we might entertain the possibility that the same is true of those complex concepts that exhibit prototypicality. One would need to say something about what fixes the prototypical aspect of complex concepts that have them; but, since not all complex concepts exhibit such effects, it’s unclear whether a compositional account would be needed for this aspect of content.

Fodor, in any event, would reject the amalgamated account of lexical concepts (as well as the simple response). For he maintains not only that concepts are compositional, but further that there’s a sense in which constituent concepts must transmit to the concepts they compose *all* of their content. Not only is the content of a complex concept determined by that of its constituents and how they’re combined; further, “whatever content is, constituents must yield theirs to their hosts.” (p. 106) And, if we concede to Fodor that prototypes don’t compose, then at least one aspect of these proposed amalgamated concepts seems *not* to be transmitted to their hosts. This highly contentious transmission principle, however, receives scant defense here: Fodor writes that it “is required because whatever is true of cows as such or of brown things as such is *ipso facto* true of brown cows.” (p. 106) For two arguments in favor of a closely related principle, one must turn to Fodor’s “There are No Recognitional Concepts: Not Even RED” in his *In Critical Condition*, Cambridge, MA: MIT Press, 1998, chaps. 4 and 5.

Fodor suggests that the difficulties faced by these and other versions of IRS provide indirect support for his alternative conceptual atomism. Conjoined with informational semantics, however, his positive proposal faces some *prima facie* problems of its own. In the last chapters of his book, Fodor discusses two in particular. First, informational atomism can seem committed to the claim that all primitive concepts are innate; but nativism about practically all lexical concepts seems implausibly extreme. Second, informational atomism's account of what fixes conceptual content requires laws relating primitive concepts and what they are concepts of, and so its broad conception of the primitive thus requires that there be laws about practically all the kinds of things for which we possess lexical concepts. But many of these things—for example, doorknobs—don't seem to be the kinds of things about which there are laws. Let me comment on Fodor's response to the first problem—his response to the second will come up along the way.

In earlier writings (see especially “The Present Status of the Innateness Controversy” in his *Representations*, Cambridge, MA: MIT Press, 1981, pp. 257-316), Fodor defended radical nativism on the basis of what he calls the “Standard Argument”. It runs roughly as follows. Concepts are either learned or not. If they are not learned, they are innate. Learning a concept is an inductive process that consists of devising and testing hypotheses about the concept's constituent structure. But for this process to be possible at all there must be available for hypothesis testing primitive concepts that are not learned and are therefore innate. The primitive concepts can't be learned, since they don't have constituent structure—and any attempt to form a hypothesis about a primitive concept that uses the concept itself would obviously presuppose possession of the concept at issue. Primitive concepts are thus innate, and so, given conceptual atomism, practically all lexical concepts are innate.

Fodor now rejects this argument and so denies that he's so committed. He begins by noting that one might reject the inductivist model of concept learning. This model can seem plausible if one holds that to *possess* a concept is to know the relevant facts about its constituent structure. But this is precisely what intentional atomism denies. There is, however, an alternative route to the inductivist model. Everyone must accept “that what leads to acquiring a concept is *typically having the right kinds of experiences.*” (p. 127, Fodor's emphasis) So, everyone needs to explain why having *those* experiences results in the possession of *this* concept. Why, Fodor asks, “is it so often experiences of doorknobs, and so rarely experiences with whipped cream or giraffes, that leads one to lock to *doorknobhood?*” The inductivist model would provide an answer to what he calls the “doorknob/DOORKNOB (or, d/D) problem”. The lack of an alternative would thus lend this model support and therefore threaten to reinstate the Standard Argument for radical nativism.

But there *is* a non-inductivist answer to the d/D problem, Fodor argues, one that doesn't require concepts like DOORKNOB to be innate. The claim is that doorknobhood is a property “*constituted* by the kind of experience that leads to the acquiring of the concept DOORKNOB.” (p. 134, emphasis in original) That the property is thus “appearance”-dependent is said to *explain* why experience of prototypical instances leads to acquisition of just that concept; it provides a metaphysical explanation in place of the inductivist's attempted psychological explanation. The explanation, moreover, does not require that the *concept* be innate (it doesn't preclude this either), though it may well be

required that *something* be innate, if only some no doubt very complicated mechanism in virtue of which in the right circumstances typical things with that property strike us as they do. Finally, that doorknobhood is an appearance-dependent property provides an answer as well to the second *prima facie* problem for informational atomism listed above: for it follows that there *is* a “law about doorknobs (qua doorknobs); viz. that we lock to them in consequence of certain sorts of experiences.” (p. 146)

One might wonder why it doesn't suffice to answer the d/D problem in the case of DOORKNOB—and thus to defuse the attempted reinstatement of the Standard Argument—simply to point out that, as a matter of *nomological* necessity, minds that are materially constituted as ours actually are acquire the concept DOORKNOB as a result of experiencing what are actually typical doorknobs. This proposed explanation isn't a mere restatement of what's to be explained, since it adverts to *how we are materially constituted*. Clearly, as an explanation it doesn't go very deep: one wants to know what the relevant facts concerning constitution are. But that would be for the relevant sciences to fill in as they can. This nomological claim is *entailed* by the conjunction of Fodor's metaphysical claim and his naturalism, and so perhaps the *metaphysical* claim can be said to *explain* in part the *nomological* claim. But explanation isn't generally transitive; why prefer the stronger claim as the *explanans*? I fear, however, that I may have an insufficient grasp of the d/D why-question's relevant contrast class. Fodor's discussion of appearance-properties (and how our concepts of them then enable us to acquire concepts of natural kinds and then concepts of natural kinds *as such*) is rich and suggestive, but his addressing in tandem—in two rather compressed chapters—these metaphysical matters along with questions of nativism (and much more besides) does pose a challenge to the reader attempting to disentangle the various themes.

A final remark. Early in the book, Fodor defends the “logical” priority of being a concept to having one and thus his methodological preference for inquiring into what concepts are independently of questions concerning what constitutes concept possession. He holds further that answers to these latter questions fall out trivially from an account of the former. “It's a general truth that if you know *what an X is*, then you also know *what it is to have an X*. . . . This applies to concepts in particular . . . if you commit yourself on one, you are *thereby* committed, willy nilly, on the other.” (p. 2) I'm not so sure either about the general truth or its particular application to concepts. I know (kind of) what quarks are, but I'm not sure what it is to have one—similarly for cities, numbers, and many other things. (Nor am I sure *whether* there is such a thing as having them, in particular whether having isn't here categorially precluded. The man behind the deli counter *asks* me if I have a number, for instance. Is this just a *facon de parler*, or is being assigned a number a way of having it?) In some cases, I have a *variety* of ideas about what having some kind of thing might be. “Who has a baseball?” “Well, I have one in my hand, but it's *Jimmy's* baseball.” Such cases perhaps show that “have” is ambiguous (in particular, polysemous) or context-sensitive or sense-general (that is, denotes a determinable: there are various more specific ways one can have something)—or some combination thereof. In any event, there would seem often to be more one needs to know, beyond what an X is, to know what relation ‘have’ expresses on some particular occasion or, perhaps, what specific sort of having is at issue. (Henceforth, for ease of exposition, I'll just talk of various senses “have” might have.) If this is often true, might it be true for concepts in particular? Fodor holds that, since for practically all lexical concepts the

concept's being what it is—and, in particular, having the *content* it has—doesn't logically depend on anything about any other concept, neither does *having* the concept. (See, for example, p. 158. Fodor does allow, however, that, as a matter of *nomological* necessity, we might be so constituted as to require the possession of some concept in order to possess some other concept.) But, though it's a possible position that possession of a concept in principle requires possessing no other—or at least doesn't require possessing any other in particular, this doesn't follow from the independence of concepts' *content* from that of others. Even if one agrees with Fodor about concept individuation, one still might hold that *having* some particular lexical concept requires also having some other concepts—and perhaps having some other concepts in particular. Indeed, Fodor need sacrifice nothing central to his view by allowing that this claim may well be true in *one* demanding sense of “have”. Thus, he could placate someone who, while granting conceptual atomism, held that it's impossible to *have* the concept DOORKNOB unless one had some ensemble of attitudes about them and thus had the concepts constituting those attitudes. Fodor could yet deny the claim in some less demanding sense. What might we take this sense to be? In what sense could it be metaphysically possible that we have a concept without having any other? Here's one suggestion: this weaker having just requires that the concept—the mental particular that's token-identical to something physical, and that stands in the sort of nomic mind-world relation that informational semantics requires, and so on for whatever else must hold of concepts—be a part of the person's body. Despite his deflationary remarks about concept possession, there's reason to think Fodor might welcome, or at least allow, this distinction between two senses of “have”. First, his talk of concept *acquisition*—of *coming* to have a concept—implicitly acknowledges two grades, the stronger requiring availability to one's cognitive economy. (See, e.g., p. 129.) Second, he at one point uses *being able to think a thought containing the concept* as a criterion for *having* it (p. 135, fn. 10)—which, though it doesn't require having any other concept in particular, does require having other concepts. Here, then, Fodor would appear to acknowledge the more demanding sense.

(I thank Jerry Fodor and Richard Samuels for commenting on a draft of this review.)

Department of Philosophy
University of Pennsylvania
433 Logan Hall
Philadelphia, PA 19104
USA

Steven Gross