

# The Drink You Have When You're Not Having a Drink

ROBERT A. WILSON

---

**Abstract:** *The Architecture of the Mind* is itself built on foundations that deserve probing. In this brief commentary I focus on these foundations—Carruthers' conception of modularity, his arguments for thinking that the mind is massively modular in structure, and his view of human cognitive architecture.

Peter Carruthers' *The Architecture of the Mind* is a wide-ranging book that offers an articulation and defense of the idea that 'the mind'—in particular, the human mind—has a massively modular architecture. In this brief commentary I focus on Carruthers' starting points—his conception of modularity, his arguments for thinking that the mind is massively modular in structure, and his view of human cognitive architecture.

## 1. What is Modularity?

Carruthers acknowledges that there are many conceptions of modularity in the literature, and he avoids being dogmatic about just what modularity must be. Recall that Fodor (1983) listed nine features of modules: they are domain specific, mandatory, operant on inputs inaccessible to downstream processing, fast, informationally encapsulated, producers of outputs that are conceptually shallow, associated with fixed neural architecture, subject to patterned breakdowns, and ontogenetically sequenced/innate. Fodor also claimed (and still claims) that only the perceptual systems *plus language* are modular. Since Carruthers aims to articulate a notion of modularity that 'makes the best sense of the claims of massive modularists' (p. 8), where these are claims that the *whole or most* of our cognitive architecture is modular, he rejects this notion of a module, which he calls a *Fodor-module*. In particular, according to Carruthers, in order to give a charitable reading to the claims of massive modularists, speed, shallowness, innate channeling, and encapsulation need to be dropped from the list of properties that modules have,

---

I would like to thank Edouard Machery for organizing the publication of this book symposium, and my fellow symposiasts and audience members at a corresponding session at the APA Pacific meetings in 2007, where an earlier version was presented.

**Address for correspondence:** Department of Philosophy, Assiniboia Hall, University of Alberta, Edmonton, Alberta, T6J 2E5, Canada.

**Email:** [rob.wilson@ualberta.ca](mailto:rob.wilson@ualberta.ca)

... leav[ing] us with the idea that modules might be isolable function-specific processing systems, all or almost all of which are domain specific ..., whose operations aren't subject to the will, which are associated with specific neural structures (albeit sometimes spatially dispersed ones), and whose internal operations may be inaccessible to the remainder of cognition (p. 12).

So what we might call *Carruthers-modules* are domain specific, mandatory, associated with specific neural structures, and processing inaccessible. Roughly speaking, these are the first three of Fodor's conditions, plus his seventh, fixed neural structures. Three points about *Carruthers-modularity*.

First, while Carruthers clearly intends this as a loosening of the notion of Fodor-modularity, retaining some but not all of the properties that Fodor ascribes to modules, the method of argument used to arrive at Carruthers-modularity is problematic. The grounds for rejecting some of the properties that Fodor cites—such as conceptual shallowness and *perhaps* informational encapsulation—are that they are *incompatible* with the claim that all or most of cognition is modular. Fair enough. But the grounds for rejecting other properties—fast, ontologically sequencing/innate, and *perhaps* informational encapsulation—is that *it is possible that* processes like reasoning and remembering might not have, or are unlikely to have, these features. This is a kind of lowest common denominator reasoning that spells trouble. In fact, it constitutes a reason for rejecting *all* of the properties on Fodor's list, even domain-specificity, as Carruthers himself seems to acknowledge (pp. 8–9). This would leave us with a notion of modularity that might perhaps best be called—paying homage to an advertising campaign popular in Australia for non-alcoholic beverages in the 1980s that played on the word 'drink'—*Clayton's modularity*: the modularity you have when you're not having any modularity.

Carruthers will surely resist the charge that his softening of Fodor-modularity amounts to Clayton's modularity, but some of his early comments in the book stiffen the challenge here. He begins by introducing a very weak sense of 'module' according to which a module is 'a dissociable functional component' (p. 2), saying that the massive modularists 'are defending a thesis of considerable interest, even if "module" just means "component"' (p. 3). While Carruthers is no doubt correct to imply that sometimes evolutionary psychologists and others think of modules in this minimal way, it muddies the waters considerably to keep the notion of modularity *this* open in making sense of the claim that the mind is 'massively modular'. 'Component', after all, and even 'dissociable functional component', are not much different from 'faculty' or 'part', and any modularity thesis worth its salt has to say something more than that the mind is built up out of lots and lots of components, faculties, or parts.

Second, consider the post-Fodorian deployment of the concept of modularity, most notably in the hands of developmental psychologists and evolutionary

psychologists, which has pursued modularity in recesses into which, Fodor himself claimed, no self-respecting cognitive scientist should venture. Like Fodor, these enthusiasts of modularity are often less than fully forthcoming about precisely what properties modules have. And like Fodor, they are right not to get too hung up on this, making hay while the sun shines. But in self-consciously flouting Fodor's First Law for the Non-Existence of Cognitive Science—the more 'central' (code words: Quinean and isotropic) a process is, the less anyone understands it—they are working with something like *Fodor-modules*. When Dan Sperber (1994) argues that there is no barrier to extending the notion of modularity to the *conceptual* realm, it was *Fodor modules* whose extension he was increasing, not some significantly more anemic notion of a module (see Hirschfeld and Gelman, 1994 more generally). Sperber argues that Fodor is mistaken in thinking that modules *in something close to Fodor's own sense of modularity* can't be conceptual (and can't in turn shape cultural representations). When Leda Cosmides (1989) argues for the existence of a *social exchange module* that underlies performance on the various permutations on the Wason reasoning tasks, again these are modules *in something close to Fodor's own sense of modularity*. In particular, both *domain-specificity* and *informational encapsulation* are not only features of these early contestations of Fodor's pessimism about the reach of modularity and the corresponding limitations of cognitive science; they are important in steering clear of Clayton's modularity.

My final point here picks up on the charge that Carruthers sails dangerously close to Clayton's modularity. In anticipation of something like this objection, Carruthers makes two pre-emptive moves. The first is to suggest that it doesn't really matter whether we accept Carruthers-modularity as a form of modularity: what matters are the substantive proposals about animal and human cognitive architecture, whatever we call them. It is true that the specific proposals concerning human cognitive architecture made in the second half of the book are worth considering, whether or not their components are called 'modules'. But since Carruthers himself invokes a very general notion of a module (= component) and a more specific notion (= loosening from Fodor-modules), casting the book in terms of modularity while divesting oneself of commitment regarding precisely which notion of modularity is in play is trying to have your cake and eat it too. The book is, after all, a defense of the *massive modularity thesis*, with specific proposals about what the resulting modules look like.

The second move is to point to appeals to modularity beyond the domain of cognition—those in play in discussions of hierarchically-ordered, complex systems à la Herb Simon, and those in play in discussions of modularity in evolutionary developmental biology à la Gunther Wagner. These are powerful appeals, and they provide motivation for the massive modularity thesis independent of the idea of extending Fodor-modularity to 'central systems'. The primary arguments that Carruthers offers for the massive modularity thesis appeal to these 'beyond-cognition' considerations in favor of modularity. The resulting arguments are our next topics.

## 2. The Argument from Design

The first of these arguments is neatly summarized (p. 25) as follows:

- (1) Biological systems are designed systems, constructed incrementally.
- (2) Such systems, when complex, need to have massively modular organization.
- (3) The human mind is a biological system, and is complex.
- (4) So the human mind will be massively modular in its organization.

The design in (1) is due to evolution by natural selection. The Argument from Design is intended to be quite general in its scope, however, and there are sources of design other than natural selection, as Carruthers makes clear. Thus, a version of this argument can be run with ‘computational system(s)’ substituted for ‘biological system(s)’ throughout. Carruthers takes Herb Simon (1962) to have made this argument in both forms.

As general as the Argument from Design is, formally similar arguments cast in terms of *physical* systems or *social* systems lack even the *prima facie* plausibility that the Argument from Design has. Their failures are instructive. The corresponding *Argument from Physical Design* is:

- (1′) *Physical* systems are designed systems, constructed incrementally.
- (2′) Such systems, when complex, need to have massively modular organization.
- (3′) The human mind is a physical system, and is complex.
- (4) So the human mind will be massively modular in its organization.

Despite the truth of (3′), since physical systems in general are not designed, (1′) is false and so (2′) moot. Consider now the corresponding *Argument from Social Design*:

- (1★) *Social* systems are designed systems, constructed incrementally.
- (2★) Such systems, when complex, need to have massively modular organization.
- (3★) The human mind is a social system, and is complex.
- (4) So the human mind will be massively modular in its organization.

Put the apparent falsity of (3★) to one side for now, and concentrate on (1★) and (2★). Although *some* social systems are designed and constructed incrementally—particularly those that are *institution-based* or *institution-focused*—this is not true *in general* of social systems, and so (1★) is not true in general. Moreover, it is unclear just what (2★) is claiming about the organization of those social systems that are modular. What does it mean for a *social system* to be modular?

Return to the original Argument from Design. The Argument from *Physical Design* tells us that (3′) is true but irrelevant, and it suggests that the appeal to

complexity in (3) in the original argument does not carry much weight: it is the appeal to *design* in (1) that matters. The Argument from *Social Design* tells us that, the contestability of (3\*) to one side, even supposing that the mind *were* a social system of some kind leaves open the question as to whether it is a *designed social system* and just what modularity amounts to for such systems. It suggests that either (1) in the Argument from Design needs to be read as making a *universal* claim about biological systems, or that (3) in that argument needs to say that the human mind is a *designed biological system*. Both claims are problematic.

Suppose that, as I have argued elsewhere (e.g. Wilson, 2004, 2005) some biological and cognitive systems are *extended* in the sense articulated by Andy Clark and Dave Chalmers in defending the *extended mind thesis* and specified in an earlier defense of *wide computationalism* (Clark and Chalmers, 1998; Wilson, 1994). Both (1) and (3) remain true when we consider such extended cognitive systems. But what of (2)? Do (complex) extended biological and cognitive systems need to have massively modular organization? Although Carruthers himself is not very sympathetic to the extended mind thesis, this is a question about his views—what they allow and what they entail—not his sympathies. Consider two ways to conceptualize the massively modular organization of extended biological and cognitive systems.

The first is to conceptualize the system as having some of its modules existing completely beyond the boundary of the organism, such that even though many modules are organism-bound, this is not true of the overall biological/cognitive system of the organism. Certainly some biological systems are like this, such as what Scott Turner (2000) calls the *extended physiology* of coral reefs and termite mounds. The second is to conceptualize the modules themselves, and not just the system that they constitute, as extending beyond the boundary of the organism. I have argued that this is true of visual modules, such as those posited in the multiple spatial channels theory of form perception (Wilson, 1994, 1995, ch. 3), and that there are general reasons to expect vision to engage in what I call *exploitative representation*, a type of representation that incorporates worldly elements rather than constructing costly internal codes or images (Wilson, 2005, ch. 8; see also Shapiro, 1997).

Here is one place where precisely what a module and massive modular organization are make a difference to the plausibility of the views that one is committed to. For if a module is simply a component, and so to have massively modular organization is simply to have lots of components, then it is difficult to see why 'going extended' *in either of these two ways* would make a difference to (2) in the Argument from Design. The same holds true for *functionally specified component*, or even, I think, for *Carruthers-module*. Strengthening the requirements for modularity puts more pressure on (2), however, insofar as extended modularity *in that sense* (e.g. Fodor-modularity) becomes harder to make sense of, much as (I suggested in passing) making sense of the massive modularity of *socially* designed systems requires additional clarificatory work. My own view is that one can not only make sense of extended Fodor-modules, but that we have lots and lots of these; but I suspect that this is an option that Carruthers himself would be, at best, uncomfortable with endorsing.

### 3. The Argument from Animals

Unlike the Argument from Design, which Carruthers clearly articulates, the other chief argument for the massive modularity thesis, his Argument from Animals, is never explicitly stated in the book, despite being the subject of discussion for about 100 pages. Here's one stab at specifying the argument:

- (i) Animal minds are massively modular
- (ii) Human minds are incremental extensions of animal minds
- (iii) So human minds are massively modular.

All but one paragraph of Chapter 2, 'The Architecture of Animal Minds', is devoted to defending Premise (i). That remaining paragraph, the final one in the long chapter, starts:

Given that animal minds are organized along massively modular lines, then normal biological reasoning should lead us to expect that massively modular architectures will be preserved in the minds of members of *Homo sapiens*, too. And now we have some specific suggestions about what that core architecture should look like ... together with a range of likely belief-generating and desire-generating modules contained in it ... (p. 149).

I take this to provide the reasoning behind (ii), and Chapter 3 to flesh out the kind of incremental extensions of animal minds that Carruthers thinks we possess in our own minds.

The formal invalidity of the Argument from Animals, at least as I have reconstructed it above, does not detract from its interest. In chapter 2 Carruthers introduces the following claims as part of his defense of (i):

- perception/belief/desire/planning cognitive architectures can be found in creatures as phylogenetically separated from *Homo sapiens* as insects are (section 1);
- the ventral/dorsal visual systems—commonly referred to as the *what/where systems*—exemplify massively modular design (section 2);
- there are a variety of distinct belief-generating systems in animals, including those for spatial navigation, for the representation of time, number, and space, for causation, and for social cognition (section 3);
- emotional and social motivational modules structure desire (section 5);
- the multiple memory systems that are widely recognized exemplify modular organization (section 6).

The first point about these claims is that they provide instances of massively modular organization only given a relatively weak understanding of what modules are, an understanding weaker than that of Carruthers—modules, for example. There

is a lot of organization to the mind, but the massive modularity hypothesis says something specific about the nature of that organization: minimally (I think), that it is relatively *domain-specific* and relatively *informationally encapsulated*. It is striking that almost none of the work that Carruthers draws on here (Gallistel's on animal navigation is the exception) has been presented in support of even moderately strong modularity theses by those who have developed the work. The discussion of the what/where system, for example, is focused on articulating the boundaries for these two systems, debating how best to think of the distinction (what versus where, what versus how), and strengthening the evidence for a double-dissociation between the two systems. *Very little of this tells us about many of the properties that characterize Fodor-modules or even Carruthers-modules*. The same is true of the putative belief-generating modules, and in general of the multiple memory systems (working versus long-term, episodic versus semantic, explicit versus implicit), even if there are some interesting domain-specific phenomena associated with each.

The second point concerns the subjects of this research. Putting aside claims about invertebrate belief-desire-planning architectures, most of this work has been reported in a relatively narrow phylogenetic range of the animal kingdom—roughly, from rats to primates—and some of that which Carruthers relies on has been found *primarily* in human populations (note that such data can't be used directly to support (i)). The first of these points suggests that perhaps (i) should read:

- (i+) Advanced animal minds are massively modular,

or perhaps,

- (i#) Animal minds that are phylogenetically close to our own are massively modular,

and I wonder whether Carruthers would accept either of these as friendly amendments to the Argument from Animal Minds as I have presented it.

Suppose that we accept that animal minds are massively modular. What about (ii), the claim that human minds are incremental extensions of animal minds, and the inference from (i) and (ii) to the massive modularity thesis about human minds? The inference assumes that massive modularity is a property conserved across incremental extension. There are two questions to ask about this: Is massive modularity conserved across a single incremental extension? Is the conservation of massive modularity transitive?

The answer to the first question is 'yes': once you have massive modularity you can't lose it in a single incremental extension. But the answer to the second question is 'no', since otherwise this would imply that massive modularity *could never* be lost through a series of incremental extensions, and that is obviously false: *any* phenotypic feature can be so lost. (Carruthers seems to overlook this point, as do proponents of intelligent design who appeal to properties like 'irreducible complexity'.) In thinking about the relationship between animal and human minds it is this second question

that is most relevant, since current animal and human minds are separated by at least 5 million years of evolutionary and other changes. So there are likely *many* incremental changes separating us from our animal ancestors. Whether any one of these makes the resulting mind more or less modular is an open question. These facts suggest that (i) and (ii) do not provide very strong support for (iii).

#### 4. The Modularity of the Human Mind

On pp. 154–157, Carruthers provides us with a list of 22 mental capacities that are candidates for what makes human cognition distinct. To give the flavor of the list, a representative sampling:

1. A folk-physics capacity, enabling deeper causal reasoning about physical phenomenon than may be available to chimpanzees ...
2. A folk-biology capacity, involving a weak form of biological essentialism and also enabling ‘deeper’ causal reasoning and inductive generalizations concerning living organisms.
3. A mindreading capacity, enabling us to attribute mental states to other people, and to predict their likely actions in the light of their beliefs and goals ...
  
13. A sense of humor, and a disposition to tease, to tell jokes, and to play tricks.
14. An interest in stories, and a disposition to invent and transmit such stories.
15. A capacity for music; a sense of rhythm; and a disposition to dance and to sing.
  
20. A capacity to make inferences to the best explanation in general, and to reason scientifically, in particular ...
21. An indefinitely flexible capacity for practical reasoning ...
22. A capacity, not just to reason theoretically and practically in intelligent and flexible ways, but also to reflect upon, and to modify, one’s own reasoning practices ... .

These triads of capacities are importantly different from one another vis-à-vis the overall project of the book.

The first three capacities have been prominent in cognitive developmental psychology over the past 20 years, and it is here that the contestable language of modularity is most at home, even if there remain ongoing debates about *the modularity* of folk physics, biology, and psychology. I shall have more to say about these in concluding.

The next three capacities have not been seriously treated *as modules* in the cognitive science literature. I suspect that this is why Carruthers provides neither



any references for work on these nor takes them up in the remainder of the book. They have the feel of being more at home in an early 19<sup>th</sup> century tract on phrenology, an early 20<sup>th</sup> century pamphlet on eugenics, or the most intellectually crass forms of pop sociobiology, than in a serious work on the cognitive architecture of the human mind. There is no reason to credibly posit such widely shared psychological tendencies *as parts of the human cognitive architecture*, let alone as *modular* parts of a massively modular architecture.

The final three capacities identify putative aspects of human reasoning and inference, and each is the subject of more sustained discussion in the second half of the book. Carruthers' discussions here lead him to consider ways in which inferential processing can be modified and restructured in light of an organism's experience, as well as quite radical ways in which social contexts, informational tools and techniques, and instruction contribute to the augmentation of cognitive processing in the domains of inference. Where I would see these as restructuring the architecture of cognition itself, Carruthers sees phenomena such as the rise of explicit scientific methodologies and rule systems governing both deductive and inductive reasoning as providing epistemic clues about the *internal* massively modular architecture that underpins science and reasoning.

I close by returning to the putative 'folk modules' of physics, biology, and psychology. When Carruthers discusses these in more detail (pp. 166-186), he tends to view debate over them as internal to those who are committed to some version of the massive modularity thesis, a tendency underwritten by his earlier weakening of the concept of modularity. By contrast, I think that much of the debate here, especially in the last 10 years, is over the appropriateness of viewing the corresponding cognitive architecture in those terms. Consider two examples.

Frank Keil's early work on the development of biological knowledge (e.g. Keil, 1989) is often viewed as providing evidence for a distinct folk biological module—that is, one distinct from *and not derivative from* a folk psychological module, *contra* Susan Carey's (1985) landmark work on conceptual development. However, Keil's more recent work is much more skeptical about the 'folk biology module' construal of his findings, appealing as it does to what he calls *modes of construal* and focusing on ways in which views of individuals as bearing rich, internal biological knowledge are misleading (e.g. Keil, 1994; Wilson and Keil, 1998; Rozenblit and Keil, 2002; Lutz and Keil, 2000). I read Keil as now *rejecting* the rush to think of cognitive development in terms of the triggering or unfolding of a relatively small number of distinct modules of domain-specific knowledge, at least if folk biology is meant to be one such module.

The literature on folk psychology, to which Carruthers devotes more discussion, especially to the model of folk psychology defended by Shaun Nichols and Steve Stich (2003), should be seen in much the same way. Carruthers views Nichols and Stich as proposing a cognitive architecture for mindreading that exemplifies part of our massively modular architecture, quickly dismissing their own distancing of their view from the language of modularity by attributing this to their own focus on Fodor-modules (p. 174). This allows Carruthers to skate over the difference

between providing a *model* of mindreading and postulating a *module* for mindreading. But this is not only to miss the nuances of Nichols and Stich's own discussion (see especially their pp. 131–49, and pp. 200–13), but to ignore the larger, ongoing debate over how to conceptualize mindreading and 'folk psychology'. (For two recent examples, see Melser, 2004 and Ratcliffe, 2006.) Perhaps Carruthers thinks that those, such as Melser and Ratcliffe, who dissent from the modularity-broadly-construed view of mindreading are not worth taking seriously. But surely at least Nichols and Stich are not in this category, and their overall account challenges the claim that mindreading is a *Carruthers-module*. As they say in summarizing their view, on their account:

... mindreading depends on a motley array of mechanisms. Some aspects of mindreading are best explained by appeal to an information-rich theory, while other mindreading abilities are best explained by simulation, and still others are explained by processes like default attribution, which are not comfortably classified as either theory driven or simulation based (p. 212).

Carruthers may well take Nichols and Stich to be positing lots of *components* that underwrite our mindreading capacities, but it is only confusing, in my view, to lump this together with the claim that their work supports the massive modularity hypothesis.

## 5. Conclusion

To sum up, in brief: I have suggested that while the early, non-dogmatic discussion of the concept of modularity is a welcome, refreshing feature of Carruthers' approach to the massive modularity thesis, the outcome of that discussion has left less welcome stains on his overall view of cognitive architecture. Whether this, in turn, suggests that the book has a less modular design than Carruthers himself implies in his preface (p. xi), I leave as a question for further discussion.

*Department of Philosophy  
University of Alberta*

## References

- Carey, S. 1985: *Conceptual Change in Childhood*. Cambridge, MA: MIT Press.
- Carruthers, P. 2006: *The Architecture of the Mind*. New York: Oxford University Press.
- Clark, A. and Chalmers, D.C. 1998: The extended mind. *Analysis*, 58, 10–23.
- Cosmides, L. 1989: The logic of social exchange: Has natural selection shaped how humans reason? *Cognition*, 31, 187–276.
- Fodor, J. 1983: *The Modularity of Mind*. Cambridge, MA: MIT Press.
- Hirschfeld, L.A. and Gelman, S.A. (eds) 1994: *Mapping the Mind: Domain Specificity in Cognition and Culture*. New York: Cambridge University Press.

- Keil, F.C. 1989: *Concepts, Kinds, and Conceptual Development*. Cambridge, MA: MIT Press.
- Keil, F.C. 1994: The birth and nurturance of concepts by domains: The origins of concepts of living things. In Hirschfeld and Gelman 1994.
- Lutz, D.R. and Keil, F.C. 2000: Early understanding of the division of cognitive labor. *Child Development*, 73, 1073–1084.
- Melser, D. 2004: *The Act of Thinking*. Cambridge, MA: MIT Press.
- Nichols, S. and Stich, S.P. 2003: *Mindreading: An Integrated Account of Pretence, Self-Awareness, and Understanding Other Minds*. New York: Oxford University Press.
- Ratcliffe, M. 2006: *Rethinking Commonsense Psychology: A Critique of Folk Psychology, Theory of Mind and Simulation*. Basingstoke: Palgrave Macmillan.
- Rozenblit, L. and Keil, F.C. 2002: The misunderstood limits of folk science: An illusion of explanatory depth. *Cognitive Science*, 26, 521–62.
- Shapiro, L. 1997: A clearer vision. *Philosophy of Science*, 64, 131–153.
- Simon, H.A. 1962: The architecture of complexity. *Proceedings of the American Philosophical Society*, 106, 467–482.
- Sperber, D. 1994: The modularity of thought and the epidemiology of representations. In Hirschfeld and Gelman 1994.
- Turner, J.S. 2000: *The Extended Organism: The Physiology of Animal-Built Structures*. Cambridge, MA: Harvard University Press.
- Wilson, R.A. 1994: Wide Computationalism. *Mind*, 103, 351–372.
- Wilson, R.A. 1995: *Cartesian Psychology and Physical Minds: Individualism and the Sciences of the Mind*. New York: Cambridge University Press.
- Wilson, R.A. 2004: *Boundaries of the Mind: The Individual in the Fragile Sciences: Cognition*. New York: Cambridge University Press.
- Wilson, R.A. 2005: *Genes and the Agents of Life: The Individual in the Fragile Sciences: Biology*. New York: Cambridge University Press.
- Wilson, R.A. and Keil, F.C. 1998: The shadows and shallows of explanation. *Minds and Machines*, 8, 137–159. Modified version reprinted in F.C. Keil and R.A. Wilson (eds) (2000): *Explanation and Cognition*. Cambridge, MA: MIT Press.