On Fodor-Fixation, Flexibility, and Human Uniqueness: A Reply to Cowie, Machery, and Wilson

PETER CARRUTHERS

ABSTRACT: This paper argues that two of my critics (Cowie and Wilson) have become fixated on Fodor’s notion of modularity, both to their own detriment and to the detriment of their understanding of Carruthers (2006). The paper then focuses on the supposed inadequacies of the latter’s explanations of both content flexibility and human uniqueness, alleged by Machery and Cowie respectively.

1. Fodor-Fixation

Here is the story of modularity as told by Cowie, Wilson, and most other philosophers (I confess I was once in the grip of this account myself; see Carruthers, 2003): in the beginning there was Fodor (1983), who articulated a notion of modularity fit to do service in cognitive science. Along with that notion came an interesting and controversial thesis: peripheral systems of the mind are modular while central systems can’t be. The challenge to anyone wishing to defend a thesis of massive mental modularity is then to answer the Fodorian arguments against central-systems modularity while weakening the Fodorian notion of ‘module’ as little as possible (and in particular, while retaining the core idea that modules are encapsulated processing systems).

The error in this philosophical fable is that there exist a variety of notions of modularity that are independent of Fodor, which have a life of their own in common-sense thinking, in biology and evolutionary theory, and in artificial intelligence. It is therefore an open question

I am grateful to all those who participated in the discussion of Carruthers (2006) during an ‘Author meets critics’ session at the American Philosophical Association meeting in San Francisco in Spring 2007, and especially, of course, to my three commentators: Fiona Cowie, Edouard Machery, and Robert Wilson. Hereinafter, references to these three by name alone (without date) should be understood to cite their article in the present issue of the journal.

Address for correspondence: Department of Philosophy, University of Maryland, College Park, MD 20742, USA.
Email: pcarruth@umd.edu
which notion of ‘module’ should best be appropriated to figure in a thesis of massive mental modularity. The strategy that I adopt in my book is to examine the main arguments that have been offered in support of massively modular accounts of mind, lining up the notion of ‘module’ that figures in their conclusion with the content of those arguments themselves. This is plainly the right way to go about articulating the thesis of massive modularity, if one wants to end up with the doctrine that has the best empirical and theoretical warrant.¹

There are broadly three successful arguments for massive modularity. One is an argument from the organization and evolution of biological systems quite generally, originally articulated by Simon (1962). This requires, as minor premises, the claims that the mind is complex in its organization and that it is a direct product of natural selection. Each of these seems to me pretty much incontrovertible. The second is an argument from the organization of animal minds, one sub-component of which is a claim that learning mechanisms are multiple, and vary by domain. Of course this argument requires as an assumption that the massive modularity of animal minds wouldn’t have atrophied away in the five million or so years of evolution that have elapsed since our lineage parted company with the other great apes, as Wilson points out. But since humans retain all of the same animal capacities (together with some further additions), this assumption requires little defense. The third line of reasoning is an argument from computational tractability, which is due to Fodor (1983, 2000). But in disagreement with Fodor (and influenced by the ‘simple heuristics’ research program of Gigerenzer et al. (1999)), I claim that the argument only warrants the conclusion that cognition must be built out of multiple systems that are frugal in their use of information, not that they should be encapsulated.²

What emerges from these arguments is a notion of ‘modularity’ according to which modules are function-specific processing systems which exist and operate independently of most

¹ Note that the exercise, here, isn’t descriptive but normative. While I disagree with Wilson’s factual claim that massive modularists have historically intended something pretty close to Fodor-modularity (on this I agree with Barrett and Kurzban, 2006, and Machery and Barrett, 2006), this is actually beside the point. The goal, rather, is to articulate what massive modularists should mean by ‘module’ if they are to have a well-warranted theory.

² Although Fodor accepts his own argument that the mind must be built out of encapsulated systems, he doesn’t believe its conclusion. This is because he thinks that he has independent reasons for believing in the holistic character of central processes. Hence his pessimistic overall conclusion, that we should give up on doing cognitive science for the foreseeable future (Fodor, 2000). This is because to do cognitive science we have to assume encapsulation (Fodor thinks); but we also know that the central systems of the mind can’t be encapsulated.
others, and which have complex, but limited, input and output connections with others. Each of these systems will have a distinct neural realization, and will be frugal in its use of information, while having internal operations that are inaccessible to others. Moreover, the set of systems will be organized hierarchically so that all but the bottom layer of modules will be constructed out of other modules as parts. And then the claim that the mind is massively modular in organization means that it is composed out of many, many, such systems.

Even without further elaboration, it is plain that this is an interesting and controversial thesis. *Contra* Wilson, the claim that the mind is composed of a great many physically distinct processing parts is *not* one that everyone will accept. And *contra* Cowie, this isn’t just the reductive functionalism of the nineteen-seventies and eighties. For the latter made no commitment to the physical distinctness of its functional ‘components’ (which were really just sub-functions). Indeed, functionalism was designed precisely so as to be agnostic about questions of physical realization. Nor could massive modularity as described above be something that Fodor could accept. For the latter famously believes that there are numerous central processes that are *holistic*, being sensitive to all of the information that is held in semantic memory, for example (Fodor, 2000).

Moreover, it is important to see clearly that there is a distinction between what ‘the mind is massively modular’ *means* (given the notion of ‘module’ sketched above) and what the thesis of massive mental modularity defended in my book actually claims. For while the arguments in support of massive modularity show that *all* modules will have the properties outlined above, they also show that *almost all* modules will be domain specific in their input conditions, that *most* modules will be innate, being unlearned while appearing reliably in the course of development that is normal for the genotype, and that *some* modules will be encapsulated. So the overall thesis of the book is actually a good deal stronger than can be gathered merely from reflection on what ‘massive modularity’ means (in the Carruthers sense).

Let me comment briefly on the innateness of (most) modules. Recall that one of the arguments for massive modularity is an argument from biological organization and evolvability generally. Taken by itself, this would warrant the claim that *all* modules are innate. But this is moderated by the reflection that in the case of the mind, many of the components will be learning systems of one sort or another, and that at least some of these are likely to have the function of building processing systems (such as the one that underlies the capacity to read) which one might
also want to recognize as modular. Indeed, there is evidence that behavioral skills in general are constructed out of modular (but learned) behavior-organizing cognitive systems (Kharraz-Tavakol et al., 2000; Manoel et al., 2002; Wolpert et al., 2003).

Finally, let me comment briefly on the separate affectability of most modules, since Cowie goes astray on this point, claiming that no modules will be dissociable. Of course it is true that if you destroy or otherwise do something that has an impact on the functioning of a given module, then this will have an effect on any down-stream module which normally relies upon that module for input. But it won’t have any impact on up-stream modules, nor on modules (normally the vast majority) with which the target module lacks any connection. And likewise, if you destroy a module you will also destroy all the sub-modules that are its parts. But this won’t have an impact on any modules considered with the same grain of analysis (except those that normally receive its output or the output of some of those parts), nor on the sub-components of those other modules. In consequence, dissociability of modules will be rife in a massively modular mind.

In conclusion of this section, then, I claim that the thesis of massive mental modularity defended in my book will only seem uninteresting to those philosophers who have become fixated on the particular construal of modularity-claims presented in Fodor (1983).

2. Content Flexibility

Machery does an excellent job of explaining the process of representation and rehearsal of natural language sentences, which is supposed to help with the problem of content flexibility according to Carruthers (2006). Before discussing his criticisms of the latter account, one thing should be emphasized. This is that the claim is not that the only way for contents deriving from different modules to get conjoined is via language. On the contrary, we should expect there to be lots of content integration going on within a massively modular mind—roughly, whenever two or more modules severally feed their outputs as input to the same down-stream module. But in the absence of language, what can get conjoined to what will be dependent upon, and constrained by, these contingent patterns of module connectivity.

With that said, let me now turn to Machery’s criticism. He claims that although the
language faculty might be able to combine sentences deriving from two distinct conceptual modules (hence combining concepts that might not otherwise get linked into a single representation), and can rehearse the resulting sentence in inner speech, such rehearsals can’t actually make any cognitive difference. This is because the only modules that will be capable of consuming and drawing inferences from the various combined conceptual components will be those that produced them in the first place. So Machery, in effect, concedes the role of language in underpinning content flexibility (for the contents of the rehearsed sentences do combine concepts that might not otherwise get conjoined), but denies that this account can be used to explain any other significant cognitive capacities.

I have multiple replies. The least radical is that language, and mental rehearsal of language, can serve to give a fact salience that it wouldn’t otherwise have had. Suppose, for example, that the knowledge that Alex is my brother is merely ‘theoretical’, not emotionally felt. (Perhaps he lives in a distant part of the country, and our very different ages mean that we have rarely interacted, even as children.) Then by rehearsing the sentence, ‘The cheater is my brother’ (to use Machery’s example), I might evoke filial emotions towards Alex that wouldn’t otherwise have been triggered, and which can serve to counter-balance my punitive ones. Or that rehearsal might serve to evoke a norm which wouldn’t otherwise have become activated, causing me to think, ‘Families should stick together’, or something of the sort.

Another reply is that by combining together the items of information necessary to solve some problem, rehearsed sentences can enable information to get used that wouldn’t actually be looked for otherwise. I take it that this is the moral of Spelke’s reorientation experiments with infants and adults (Hermer-Vazquez et al., 1999; Shusterman and Spelke, 2005). By rehearsing the sentence, ‘The toy is to the left of the red wall’, children and adults can reorient themselves successfully and find the target object following disorientation, when otherwise they would have defaulted to the use of geometric information alone—hence failing in the task on 50 percent of occasions in a rectangular space.³

³ The main data are that children who lack productive use of ‘left’ and ‘right’ use only geometric information when reorienting, ignoring the relevance of the red wall—as do rats—and that adults who are engaged in shadowing speech—and who are therefore unable to rehearse sentences—do likewise. And it should be noted that the animal literature is rife with examples of this general sort, where a specific type of goal only issues in searches for some types of information while ignoring others. (See Chapter 2 of my book.) Mental rehearsal of sentences enables us
More radically, in Chapter 5 I show how natural language sentences can be generated and rehearsed creatively, hence doing far more than combining together two existing module-produced judgments. For example, metaphors might be engendered creatively by ‘boosting’ and formulating into language weakly activated concepts that are drawn from different domains and/or modules, utilizing semantic associations. Thus a child confronted with a banana might entertain the sentence, ‘The banana is a telephone’, prompted by the similarity in shape between the banana and a telephone handset. When this sentence is rehearsed, inferences appropriate to telephones can be drawn (one can talk to grandma on it, for example) which would never have been entertained otherwise, and the child can be launched into an episode of pretend play.  

Even more importantly, however, what Machery overlooks is the role that language plays in so-called ‘System 2’ thinking and reasoning. Let me say a little about the latter by way of background, and about dual-systems theories of cognition generally, before developing the point.

Dual systems accounts of human reasoning are now quite widely accepted, at least in outline (Evans and Over, 1996; Stanovich, 1999; Kahneman, 2002). Most researchers agree that System 1 is really a collection of different systems that are fast and unconscious, operating in parallel with one another. The principles according to which these systems function are, to a significant extent, universal to the human species, and they aren’t easily altered (e.g. by verbal instruction). Moreover those principles are, for the most part, heuristic in nature (‘quick and dirty’), rather than deductively or inductively valid. It is also generally thought that most, if not all, of the mechanisms constituting System 1 are evolutionarily ancient and shared with other species of animal. System 2, in contrast, is slow, serial, and conscious. The principles according to which it operates are variable (both across cultures and between individuals within a culture), and can involve the application of valid norms of reasoning (although System 2, too, can involve

---

4 I argue in Chapter 6 that creatively generated and rehearsed natural language sentences also play a crucial role in scientific thinking in adulthood, and in suppositional thinking more generally (also being employed by hunter-gatherers when tracking prey, for example; see Liebenberg, 1990).

5 I should emphasize that in my view it is only the general principles of operation of System 1 systems that are hard to alter, rather than their contents. For many of these systems have been designed for learning, enabling us to extract new information from our environment in quick and reliable-enough ways. I also want to distance myself from a claim that some dual-systems theorists endorse, that System 1 systems are associationist in character (Sloman, 1996, 2002).
the use of heuristics). These System 2 principles are malleable and can be influenced by verbal instruction, and they often involve normative beliefs (that is, beliefs about how one should reason). Moreover, System 2 is generally thought to be uniquely human.

In my book (chapters 4, 6, and 7) I propose that System 1 should be identified with the set of central modules for belief-formation, goal-formation, and decision making, while System 2 is realized in cycles of operation of System 1, utilizing mental rehearsals of action (resulting in either visual imagery or inner speech). It is because System 2 involves cycles of mental rehearsal that it is slow; it is because (roughly speaking) only one action can be rehearsed at a time that it is serial; and it is because the resulting images are ‘globally broadcast’ (in the sense of Baars, 1988, 1997) and made available to the full range of conceptual modules that some aspects of its operation are conscious. Since System 2 is action-based, moreover, it can be influenced directly by verbal instruction and guided by normative beliefs (as can actions generally).

When a sentence that combines concepts drawn from two or more central modules is formulated and mentally rehearsed, therefore, it becomes available to enter into far more processes than any of its component concepts could have done alone (without rehearsal). The thinker might have normative beliefs about what one should infer from a sentence of that type, for example. Or the sentence might fit as a component into some abstractly-described set of action-sequences which have been copied from others and have now become habitual. And of course the sentence is now also an object of conscious reflection—the subject might wonder whether or not it should be believed, for example, or acted upon.

The conclusion of this section is that Machery has failed to see the wood for the trees. He is quite right that a rehearsed sentence that results from combining two module-produced judgments will get broken up again and worked on by those very modules that had produced its components.6 But there are numerous other sorts of effect that mental rehearsal can have, ranging from the recall of information that wouldn’t otherwise get looked for, to serving as the basis for a

---

6 Even this might have significant effects, however. Since cognitive systems are to some degree noisy in their operations, the mere recycling of information can sometimes cause large overall changes. Consider what happens in video feedback (Crutchfield, 1984). Directing a camera at a blank television screen, in circumstances where the camera’s output will display on that very screen, causes all sorts of interesting things to happen. Rich patterns of color and shape tend to result from the cycling of the feedback loop alone, without the injection of any initial content, and without design. Something similar might happen in the human mind, once it begins cycles of mental rehearsal.
whole new system of flexible thinking and reasoning (System 2).

3. Human Uniqueness

Chapter 3 of my book proposes that there are a number of different innately channeled modules and behavioral tendencies that underlie human uniqueness. These include, at a minimum, a folk-physics module, an innately structured natural kinds learning mechanism (folk-biology), a mind-reading system (which often operates in conjunction with other systems for supposition and for simulation), a language acquisition module, various innate biases in cultural learning (such as a tendency to imitate those who are prestigious), and a system for norm acquisition, normative reasoning, and normative motivation. In addition, Chapter 5 suggests that an innate disposition for creative generation and rehearsal of action schemata issues in pretend play in childhood, and sets up the developmental sequence that eventually culminates in System 2 reasoning, together with a capacity for creative hypothesis generation and inference to the best explanation.

Cowie will have none of this. She thinks that it requires too much genetic change to have taken place in the 5–7 million years that have elapsed since the hominid line diverged from the other great apes. And she thinks that a bigger, highly plastic, brain combined with cumulative environmental change and niche construction can do the trick. I shall reply to these points in turn. But first I want to comment on the list of distinctively human traits that I provided in

---

7 Note that I don’t think that the folk-biology system comes with a lot of innate content, as Wilson claims in interpretation of me. Rather, it contains a set conceptual templates which issue in hierarchical, mutually exclusive, classifications of natural kinds as a result of learning from the local environment, together with a disposition to make appropriate inductive projections across nodes in the hierarchy (Atran, 2002). This is one of the places where Wilson is led astray by assumptions about what modules must look like.

8 Again Wilson misreads me on this point. My view is not that there is a module underlying the entire set of processes involved in mind-reading. Rather, just as the language faculty must interact with other (semantic and pragmatic) systems when engaged in the process of language comprehension, so a core mind-reading system (whose contents are partly innate and partly learned) must utilize systems of mental rehearsal for suppositional thinking, as well as a whole suite of inference systems that can reason from suppositions, in the course of doing its work. This picture is entirely consistent, I believe, with the views of Nichols and Stich (2003).
Chapter 3 of my book, which is reproduced (in part) by both Cowie and Wilson in the course of their criticisms. These include all of the systems listed above, together with many other dispositions and capacities (22 in all), including gossip, music, a sense of humor, and so on and so forth. Wilson remarks that some of the items on this list remind him of ‘the most intellectually crass forms of pop sociobiology’.

In contrast with the claims about me made by both Cowie and Wilson, I want to emphasize that I do not think that each item in the list corresponds to either a distinct module or a distinct adaptation. On the contrary, I wrote, ‘Whether each of them involves a distinct mental module, or whether some can be fully explained in terms of others, is a matter for discussion, both here and in later chapters’ (2006, p.154). (I also pointed out that I wouldn’t have the space to discuss them all.) And both critics appear to have overlooked the fact that some of the listed capacities (such as the capacity for scientific reasoning) are fully explained in terms of others in later chapters.

The point of providing a list of uniquely human cognitive capacities and dispositions (which is highly incomplete, I should stress—there are many that I did not list; see Brown, 1991, for a more complete accounting) is to present a challenge to any theorist who proposes that there was just one big change underlying human uniqueness (whether it be mind-reading, imitation, language, or—with Cowie—a bigger brain). The challenge is to explain how the one ‘great leap forward’ in human evolution can explain all the rest. But this challenge is never actually taken up, in any detail—and with good reason, since one has only to begin working at it to see that the task is a hopeless one. Or so I argue in my book. Cowie’s mere assertion to the contrary should carry no weight.

Let me turn now to the issue of the time that has elapsed since hominids first diverged from the last common ancestor of ourselves and our nearest relatives, the chimpanzees (5–7 million years), as well as to the question of the extent of the genetic differences that exist between the two groups (often said to be a mere 1.5%). In actual fact, 5 million years is quite a long time, especially if selection pressures are intense. A mere 10,000 years of evolution separates polar bears from black bears, for example, despite the many differences between them. Moreover, 5 million years was, manifestly, plenty of time in which to evolve the many bodily differences between humans and chimpanzees, including (at a minimum), body size, arm length, head size, upright posture, dropped larynx, opposable thumbs, less flexible feet, reductions in gut
size, sweat glands, numerous changes in facial morphology, hairlessness, and concealed ovulation and menopause in females. It is far from clear why the evolution of cognitive adaptations should require any more time than bodily ones.

The claim of a mere 1.5 percent difference (98.5 percent in common) is one that is often bandied about. But this probably vastly underestimates the genetic disparity between the two species. One reason is that even when genes are indistinguishable, they can be spliced differently in different species during the process of transcription. Indeed, recent research suggests that a significant source of diversity amongst mammals lies in species-specific alternative splicings of genetic sequences (Pan et al., 2005). Another reason for doubt is that the figure of 1.5 percent only includes base-pair substitutions. When insertions and deletions are also included, the differences between humans and chimpanzees are much more significant, yielding a figure closer to 13 percent (Anzai et al., 2003). And most importantly, the quoted figure only concerns protein-coding genes. But we now know that much of the DNA previously regarded as ‘junk’ is actually involved, inter alia, in regulating the gene-expression of particular genes or sets of genes (ENCODE consortium, 2007). And when one looks specifically at sequences of DNA known to be involved in gene regulation, what emerges is that the differences between chimpanzees and humans are of the order of 15 percent (Ebersberger et al., 2002). So there is no question but that there are plenty of materials, here, with which to explain the existence of multiple species-unique mental modules in humans.

Of course I don’t deny the importance of niche construction, either in general, or in human evolution in particular. But I would add that it is common for a constructed niche to create pressure for the evolution of adaptations to that niche. One clear example is the evolution of the genes that confer lactose tolerance after normal weaning, which are an adaptation to the cultural practice of dairy farming in northern latitudes (Durham, 1991). And something similar is almost certainly true of our mind-reading capacities, which are an adaptation to the increasingly sophisticated social arrangements and interactions that our ancestors built for themselves.

Nor do I deny the importance of cultural accumulation, and the ‘ratchet effect’ that this provides for human capacities. And there is nothing in Cowie’s potted history of the last 10,000 years with which I would want to take issue.\footnote{I would, however, want to emphasize that although an explicit scientific method may only have appeared in the 1600s, in fact that kind thinking is continuous with forms of reasoning that are present amongst all other cultures,}

Indeed, one of the goals of my book is to lay out
the modular systems and innate dispositions that make cultural accumulation possible. One is a mind-reading system (involved in sophisticated forms of imitation and the interpretation of speech), another is a deepened folk-physics system (which is involved in our technical and tool making abilities), another is language (which enables the transmission of information not available through observation), and yet another is a capacity for creative supposition, which is crucially involved in the genesis at least some cultural advances. The chances that all of these might be mere side effects of bigger brains and enhanced general learning capacities strike me as vanishingly small. Cowie has nothing to show on her side, here, beyond mere hand waving.

4. Conclusion

A symposium format means that an author gets only about as much space for his replies as each of the three critics gets for their criticisms. So I have had to be selective, focusing on those topics that are closest to the main themes of my book. I have said nothing about the extended mind thesis, for example, which Wilson discusses at length; although I am pleased to learn from him that it may be consistent with a thesis of massive mental modularity. (I suppose it is a good thing that one’s views should be consistent even with likely falsehoods, provided that the latter have at least some chance of being true.)

The take-home message is that philosophers of mind need to liberate themselves from the baleful shadow cast by Fodor’s (1983), recognizing that there are many more ways of conceiving of, and defending, a massively modular mind than was ever thought of in Fodor’s philosophy. And once so liberated, there are many more possibilities for explaining the distinctive properties of human and animal minds, also.

Department of Philosophy

and which are practiced \textit{inter alia} by hunters when tracking their prey (Liebenberg, 1990). Chapters 5 and 6 of my book try to explain in some detail what makes creative theory generation and inference to the best explanation possible, in terms that are consistent with the massive modularity thesis. To assert, in contrast (as Cowie does), that our brains got bigger hence we got capacities that resulted in the capacity for science, is to provide no explanation at all.
References


