PHILOSOPHICAL PSYCHOLOGY, VOL. 15, NO. 4, 2002

A critical review of Jerry A. Fodor's *The mind doesn't work that way*

DANIEL A. WEISKOPF

ABSTRACT *The "New Synthesis" in cognitive science is committed to the computational theory of mind (CTM), massive modularity, nativism, and adaptationism. In* The mind doesn't work that way, *Jerry Fodor argues that CTM has problems explaining abductive or global inference, but that the New Synthesis offers no solution, since massive modularity is in fact incompatible with global cognitive processes. I argue that it is not clear how global human mentation is, so whether CTM is imperiled is an open question. Massive modularity also lacks some of the invidious commitments Fodor ascribes to it. Furthermore, Fodor's anti-adaptationist arguments are in tension with his nativism about the contents of modular systems. The New Synthesis thus has points worth preserving.*

1. Introduction

Jerry Fodor's latest book, *The mind doesn't work that way* (henceforth MDW), is a vigorous and sustained argument in support of the claim that there is something fundamentally mistaken about the approach to understanding the mind embodied in our current cognitive science. MDW revisits and elaborates themes from his earlier *The modularity of mind* (Fodor, 1983; henceforth MOM), particularly the controversial and pessimistic consequences adduced in that book's final chapters. Cognitive scientists and philosophers who were scandalized by MOM's wide-ranging negative claims about the limits of classical computational psychology will likely be equally outraged by MDW.

Fodor's discussion centers around three major themes in what he calls the "New Synthesis" in cognitive science. The New Synthesis is defined by four doctrines: (1) the computational theory of mind (CTM); (2) modularity; (3) nativism; and (4) adaptationism. Epitomizing these doctrines are evolutionary psychologists, particularly Steven Pinker, Henry Plotkin, Leda Cosmides and John Tooby. Roughly the first half of the book is devoted to arguing that, given the commitment to CTM, there can be no psychological theory of "central processes" such as belief fixation and the planning of action. The second half of the book is divided into discussions of the New Synthesis commitments to modularity and adaptationism. Many New Synthesists advocate the massive modularity hypoth-

Daniel A. Weiskopf, Department of Philosophy, Washington University, One Brookings Drive, St. Louis, MO 63130, USA, email: daweisko@artsci.wustl.edu

ISSN 0951-5089/print/ISSN 1465-394X/online/02/040551–12 © 2002 Taylor & Francis Ltd
DOI: 10.1080/0951508021000042067

esis—the claim that the mind is modular not just at the periphery, but through and through. Further, they believe that this massively modular architecture is a collection of Darwinian adaptations, and thus can be explained as the product of natural selection. Fodor concedes that the massive modularity hypothesis would, if true, block the pessimistic inference about the limits of CTM. Nevertheless, he argues that massive modularity is false; indeed, that it is in a sense incoherent. He does, however, advocate adaptationism about most modular systems, although on grounds different from those offered by New Synthesists.

The book thus is both ambitious and brief, which is not always a salutary combination. The writing, as usual, is quite entertaining, and although the main line of the argument is clear enough, on a number of crucial points one wishes that Fodor had paused to elaborate further. The remainder of my discussion will touch on some of these points. Following what I have described as the basic structure of the book, this review has three sections—the first dealing with the limits of CTM, the second dealing with the massive modularity hypothesis, and the third dealing with adaptationism about cognitive architecture. I will not cover all of the provocative arguments Fodor musters; here I can only convey the bare outlines of a case against the New Synthesis.

2. Frames and the limits of computation

In MDW, Fodor defines CTM by reference to "rationalist psychology," which says that (1) at least some mental states have logical form, and that (2) those states' causal roles in an organism's mental economy depend, at least in part, on their logical forms [1]. CTM is the conjunction of three theses:

(i) rationalist psychology;
(ii) for every mental state with logical form, there is a syntactically structured mental representation (MR) on which it supervenes [2]; and
(iii) the causal processes that determine transitions between mental states are sensitive *only* to the syntactic properties of the MRs that realize those states.

CTM thus shows "how thinking could be both rational and mechanical" (p. 19). Mental processes are rational to the extent that, given true premises, they lead to true conclusions: "One true thought tends to lead to another in the course of cognition, and it is among the great mysteries about the mind how this could be so" (p. 18). The explanation of the mystery is that the MRs on which our thoughts supervene have syntactic structure, and the computational processes that operate on those MRs respect their semantic properties in virtue of directly responding to their syntactic properties [3].

Fodor's main argument for the incompleteness of CTM runs as follows:

(1) Computational processes are sensitive only to syntactic properties of MRs.
(2) Syntactic properties are local to the MRs that have them.
(3) Global inferences [4] are sensitive to nonlocal properties of a propositional

attitude (e.g. properties such as "being embedded in a certain kind of theory").

(4) The properties that global inferences are sensitive to are not reducible to syntactic properties of MRs.

(5) So, global inferences are not implemented by computational processes.

The general form of this argument is familiar from discussions of the frame problem in artificial intelligence. I will call this version of the problem the *computational* frame problem. Premises (1) and (3) are points of definition. One might question the argument's force on the grounds that there simply are no global processes in human mentation. Fodor does not give this possibility much time, although he at one point refers to the "apparent" effects of "apparently" global processes (p. 36), suggesting, at least, a less firm stance. In any case, the argument depends on the *reality* of global processes, which seems to be an empirical issue, and so it still remains to be determined whether humans perform genuinely global inferences. Supposing that we do, however, the most contentious premises are (2) and (4); I will discuss these in turn.

Characterizing the "locality" of syntactic properties in a way that makes clear their conflict with global inference is tricky. Fodor describes syntactic properties variously as being (i) intrinsic, (ii) essential, and (iii) context-independent. These are not clearly equivalent, however, nor is it completely obvious which reading his argument depends on [5]. However the relations among the three readings of locality work out, though, Fodor is committed to what he calls Principle E: "Only the essential features of a mental representation can determine its causal role in a mental life" (p. 24). Syntactic properties are essential to MRs being type-identified as the kind of MRs that they are, so they are partial determinants of causal role. The computational frame problem is whether they are sufficient to determine total causal role, or whether other essential features must play a role as well.

Take a case of the computational frame problem at work. A body is discovered in the study, and according to my theory, Boris is the murderer. The discovery of convincing evidence that Boris was in Minsk at the time of the killing would complicate this theory considerably, while it would probably not affect much the theory that Natasha is the killer. A thought may, when introduced into one theory, cause widespread revision of a large number of beliefs, perhaps even the rejection of the theory's central principles themselves. But the same thought, when introduced into another theory, may cause no change at all, or at most a very small change in some peripheral beliefs. So the property of a belief that it contributes a certain increment to the complexity [6] of a theory is a nonlocal property of that belief. The processes that mediate these two cascading processes of belief revision are computational, and therefore are responsive only to the syntactic profile of the MR realizing the thought. This profile is identical in both cases; so what explains the global nature of the revision in the first case, and the comparatively restricted revision in the second case?

A reasonable response is that the revision is determined by the syntactic features of the background theory itself, which is the only factor that differs in the two cases.

By conjunction we can form a single large MR out of all of the relevant context plus the thought itself [7]. There must, then, be some process that takes this large representation in its domain, and produces the variously revised and simplified versions of the theory as output. What would be wrong with this approach? Fodor follows three separate response strategies: (1) computational tractability, (2) locality of computational processes, and (3) the rationality of inference.

Consider the argument from computational tractability. Fodor notes that "the representations over which mental processes are actually defined are *much shorter* than whole theories" (p. 31). This suggests one interpretation of premise (4): the reason this attempt to reduce global properties to syntactic properties fails is because whole theories are just too big to be units of computation. In short, the problem is that "the units of thought are much bigger than in fact they could possibly be" (p. 33).

However, a large data structure is intractable only if cognitive processes can only access it piecemeal. Perhaps there are processes that can simultaneously "see" and operate over larger representational chunks. If some of these processes are among the atomic operations of the architecture, then tractability can be restored. To this, Fodor says that "in classical models, the *architectural* processes are all local, just like the computations" (p. 45). More specifically, classical processes "are (or reduce to) operations defined over symbols that belong to the *primitive* vocabulary of the language that the machine computes in" (p. 45). It isn't clear what motivates this restriction. Often one wants a primitive operation to simultaneously perform operations on two data structures at once (in programming, this is sometimes done to prevent so-called "race conditions"). These atomic operations are not architecturally reducible to two separate processes. In addition, one could imagine an architecture that is classical in Fodor's sense, but which consists of many parallel processes, each of which operates simultaneously over part of a large data structure. Here the operations are all local, but carrying them out in parallel may offset the computational cost. The link between classical computation and local processing needs to be argued for at greater length.

Finally, Fodor suggests that the problem with operating on very large MRs may not in fact have anything to do with the computational/architectural limitations described above. Instead, he suggests that "[w]e don't understand how such tractable processes could be rational (say, in the sense of reliably truth preserving) and not reducible to local operations" (p. 45). That is, there is something epistemologically suspect about such solutions to the computational frame problem. "It isn't simply that whole theories are generally too big to get one's head around—too big to think about all at once. It's also that assessments of confirmation can be, should be, and generally are called for in respect of objects much less elaborate than the totality of one's cognitive commitments" (pp. 31–32). Call this the *epistemological* frame problem: in order to be rational, the processes that underwrite global inferences must involve surveying less than the totality of a theory or belief system.

Now, it seems possible that we could discover that global inferences are not rational in this sense. If our most adequate psychological theory involved processes that operated over very large numbers of beliefs, would this be a reason to conclude that the theory was mistaken, that we are not in fact rational, or that the condition on rationality needs to be reassessed? Fodor treats failure to solve the epistemological frame problem as an objection to a theory of how the mind works; thus his pessimism about the generality of CTM. But one could with as much propriety treat it as a discovery that rationality might be different from what we supposed. No one yet has any idea what nonlocal atomic computational processes there might be, of course. The point is that the theory of rationality does not *a priori* rule out the possibility that we might have some.

3. Modularity solutions to the frame problem

The main point of MOM is to to characterize the concept of a module, and to make it plausible that peripheral systems are such modules. Hence, it consists largely of evidence drawn from linguistics and psychology, which is intended to show that input systems typically display the properties characteristic of modules. The examples range over phoneme perception, parsing, the structure of the lexicon, object identification, and capacity of visual recognition memory. No doubt this careful attention to the methods and results of empirical inquiry is part of what accounts for the popularity of MOM among cognitive scientists.

By contrast, MDW displays almost none of this concern with the details of the empirical literature. In fact, the only purported module that receives anything like an extended discussion is the so-called "cheater detection" module postulated by Cosmides and Tooby (1992), and even that is intended to show that evidence for such a module is just a materials effect. In part, this shift of focus away from the details of experimental results can be attributed to the striking success of MOM. It is no longer necessary to show the concept coherent, nor to make the case that at least some of the mind is modular.

According to MOM, modules are domain-specific, mandatory in their operation, fast, and informationally encapsulated. They have relatively shallow or impoverished outputs, and extramodular processes, especially those involved in belief formation, have little or no access to the intermediate representations they use in computing those outputs. Their ontogeny exhibits a characteristic pacing (e.g. there are critical periods for the development of language and normal vision), and they have regular patterns of breakdown or malfunction. Finally, they are associated with a relatively stable neural architecture, which explains this maturational pacing and pattern of deficits, since different parts of the brain may have different rates of growth and development, and different propensities to recover from injury. The neural architecture associated with a module also explains its informational encapsulation, since the absence of an anatomical pathway between two parts of the brain prevents either one from accessing the other's informational contents [8].

The great virtue of modules, from Fodor's point of view, is that they are immune to frame problems. Specifically, they are immune to what I have called the computational frame problem "if the database [to which the module has proprietary access] is small enough to permit approximations to exhaustive searches" (p. 63). Notice that this leaves the status of modules with regard to the epistemological frame problem wide open. In particular, it leaves open the possibility that modular cognition is not *rational* cognition, in the sense in which that problem defines it. How seriously one takes this objection depends on how seriously one takes the claim that what goes on in modules is importantly like our paradigms of rational inference (namely, those that involve beliefs and other components of central cognition). Suppose that the MRs filling a module's database are appropriately called the module's beliefs. Then modular cognition is not rational; and if a process's being rational is necessary to our having an adequate understanding of it, as Fodor at one point suggests, then we have no such understanding of modules either. Again, one response to this would be to abandon the requirement of the epistemological frame problem. But if we abandon it for modular cognition, why not for central cognition as well?

Setting these problems aside, the New Synthesis claims that salvation from the computational frame problem is found in the massive modularity hypothesis. While never formulated precisely, massive modularity is approximately the view that "there is a more or less encapsulated processor for each kind of problem that [the mind] can solve" (p. 64). Instead of an architecture in which a general-purpose central cognitive system is surrounded by modules devoted to sensorimotor processing, New Synthesists posit an architecture in which there is nothing (or almost nothing) besides modules and their interconnections [9].

A massively modular architecture faces a problem about what determines the inputs to each module. Imagine that there are two modules, M1 and M2, each of which deals with a different domain. Domain-specificify entails that each module is triggered only by MRs bearing certain features. What kind of process could assign features to MRs in a way that ensures that they get routed to the appropriate modules? Fodor offers a dilemma: either there is a single process, BOX_1, that assigns the relevant features, or there are two processes, BOX_2 and BOX_3, each of which is responsible for routing MRs to M_1 or M_2, respectively. On the first horn, there is at least one process that is less domain-specific than M_1 or M_2, and massive modularity is false. On the other horn, the question again arises: what determines whether MRs are routed to either BOX_2 or BOX_3? There can be no infinite regress here, so we will at some point be forced back to the first horn, which again defeats massive modularity.

The problem with this argument arises with its first horn. It isn't clear why the existence of a mechanism that is less domain-specific than some downstream modules is incompatible with massive modularity as defined above. Fodor says that such a mechanism is incompatible with the claim that the mind "consists of nothing but systems that are, more or less, all equally domain specific" (p. 73). The notion that the mind might be nothing but modules doesn't say anything about the relative size of their domains, though. All that seems to matter is that there might be no processes capable of making all-things-considered decisions that draw on all of the organism's beliefs and utilities.

Fodor wants to block the notion that there might be a mechanism in charge of routing MRs to modules that takes all representations in its domain. He suggests that this mechanism would be, essentially, the empiricist's sensorium: "your sensorium is assumed to be less modular (less domain specific) than *anything else in your head*" (p. 74). This supposition leads to the absurd conclusion that "every cognitive distinction corresponds to a sensory distinction" (p. 74). This *reductio* is not convincing, however. Imagine a mind structured as follows. At the input layer is a domain-general sensorium, which feeds into an initial layer of modules, which feeds into a further layer, and so on. The features that route MRs to the initial layer of modules must be definable in sensory terms, since no other MRs are accessible to the system at that stage. However, the outputs of the layer one modules needn't be definable in terms of those sensory representations. Modules have access to a proprietary database, and may express their outputs in MRs that are not meaningful to the systems that send them input. So although initial routing decisions may be based on sensory MRs, routing at deeper layers is likely to be in terms of the properties tagged by the output of the deeper modules. There is nothing at all in the supposition of a domain-general sensorium that commits a cognitive scientist to the invidious empiricist principle about the equivalence of conceptual and sensory content.

Having said this, though, it is true that massive modularity does pose problems. Take Fodor's example: how does the mind determine what is in the input domain of the cheater detection module? Social exchanges are not, obviously, marked by any kind of apparent sensory properties. Figuring out whether something is a social exchange or not takes reasoning that is likely to be global; but this is exactly the kind of reasoning that CTM allegedly lacks the resources to model. Problems that are further from the highly encapsulated activity of the periphery will need to draw on more and more of the context in even setting up their triggering conditions. So perhaps massive modularity is doomed in any case, although to really show this would require showing that the processes governing these inferences are global enough to be problematic.

4. Adaptationist arguments about cognitive architecture

Fodor's final chapter examines and rejects three arguments the New Synthesis gives for thinking that the innate structure of the mind is a Darwinian adaptation. The first of these arguments rests on general considerations about the consistency of psychological theory with evolutionary biology, and the second rests on considerations about the notion of function that is best employed in psychology. Since I think that Fodor's arguments here are essentially correct, I will comment only on the third argument, which adverts to the complexity of structure that the mind displays.

According to the complexity argument, the only way that a complex, adaptive system such as the mind could have arisen is by the process of natural selection. So the mind's complexity provides a reason to think that it is an adaptation. To this, Fodor replies that (i) complexity of mind/behavior is irrelevant to whether something is an adaptation; and (ii) all that does matter to whether a trait is an adaptation is "how much genotypic alteration of the nearest ancestor that lacked the trait would have been required in order to produce descendents that have it" (p. 88). Traits that would have required a lot of genotypic alteration are likely to be adaptations; traits requiring a little are less likely to be so.

The first, negative, point about adaptations seems quite mistaken. It is not *definitive* of something's being an adaptation that it is complexly structured, but this does not make it irrelevant. Complexity is extremely compelling evidence that a structure or behavioral type is an adaptation. It is the powerful appearance of design that typically leads evolutionary biologists to suppose that the trait in question was shaped by selection. If the structure of our own minds is sufficiently complex and has the right sort of "appearance of fit" with our environment, then, applying this standard reasoning, we will suspect it of being an adaptation.

This is not conclusive proof that a trait is an adaptation, however. According to the usual understanding, a trait is an adaptation only if it has been modified during its evolutionary history in ways that increase its task performance, and these modifications persist because of the higher fitness that results from its performing its task better (West-Eberhard, 1992). The kind of modification of the trait from its ancestral forms that Fodor adverts to is not the whole story of whether something is an adaptation. Adaptations have been been shaped by natural selection for certain of their effects. This does not, in itself, say anything about the magnitude of the genotypic differences between organisms having the trait and those lacking it.

The genotypic difference criterion is important to Fodor's argument because he suggests that it is compatible with all that we know about cognition that the difference between our minds and the minds of our evolutionary ancestors might be the result of some small, incremental change in a physical parameter of the brain (he suggests that perhaps the brain's simply getting bigger might do it). If the relationship between architecture and neural structure is "nonlinear" in this way, then maybe our kind of cognitive architecture is, in effect, a saltation and not an adaptation. Whether one is persuaded by this depends on whether one really thinks that the difference between our brains and those of our ancestors is just "more of the same." The question is one that we can make some progress in investigating, and the present verdict seems to be that the differences are much more pervasive than mere size or any other simple increase in a parameter already present in other primate brains (Preuss, 2000).

In many respects, the most important change from MOM to MDW is Fodor's present agnosticism about the neural basis of more or less *everything* at the cognitive level. In MOM, Fodor suggested that it appeared that the peripheral systems (sensory plus language) typically had the kind of stable neural architeture distinctive of modules, and furthermore that the "association" areas of the neocortex were typically plastic in their connectivity. He was willing, tentatively, to conclude that this neural evidence further supported his conclusions about cognitive architecture, namely, that modules are informationally encapsulated and central cognition is not. So MOM suggested that cognitive psychology might be constrained, or at least *informed*, by neurobiology.

The attitude of MDW, by contrast, can be summed up by the following passage:

> Since psychological structure (presumably) supervenes on neurological structure, genotypic variation affects the architecture of the mind only via its effect on the organization of the brain. And, since nothing at all is known about *how* the architecture of our cognition supervenes on our brains' structure, it's entirely possible that quite small neurological reorganizations could have effected wild psychological discontinuities between our minds and the ancestral ape's. (p. 88)

Later he repeats the point: "Nothing at all is known about the laws according to which cognition supervenes on brain structures, or even about which brain structures it is that cognition supervenes on" (p. 89). This is as extreme a form of neural agnosticism as one could imagine, and it seems entirely unmotivated. Ignorant we may be of many things concerning the mind/brain relation, but many cognitive functions, including even some "central" processes such as affect, working and long-term memory, and categorization, have been reliably associated with particular anatomical regions of the brain [10]. One would like to know how much evidence, and of what kind, would be required to really *teach* us something about how the mind relates to the brain.

This neural agnosticism comes back to bite Fodor in his discussion of why the modular, innately specified parts of the mind are likely to be adaptations after all. Modules contain databases of contingent but true information about the creature's environment. It is wildly unlikely that this body of true representations was formed as a saltation (i.e. by a one-step change in neural structure; the epistemic fit between the representations and the world is too tight). But if saltation is not the explanation, then natural selection must be. So, despite the failure of the New Synthesis arguments from consistency, teleology, and complexity, there is a reason to think that the modular mind is an adaptation after all.

Selectional mechanisms depend at each stage on some small, incremental change in phenotype mediated by a correspondingly small genotypic change. The relevant phenotypic trait here is presumably some aspect of the brain. How cognition supervenes on the brain, though, is allegedly totally unknown to us. In fact, it is "compatible" with what we know that just such small changes in brain structure might produce very large changes in cognitive structure. Even if we then suppose that an organism starts out with its brain wired in such a way as to give it some innate, true beliefs, what reason is there to think that any change away from this state will make only an "incremental" contribution to its stock of beliefs? For all we know, a small change to the neural structure of an organism that believes that unsupported objects will fall could produce an organism that believes that fire quenches thirst. Why suppose, that is, that small changes to neural structure will even preserve the gross subject matter of internal representations unless the relationship between neurology and cognition is "linear" [11]?

The strong appearance of fit between mind and world that Fodor points to here sounds suspiciously like the kind of complexity that New Synthesists typically think underwrites adaptationism. If Fodor's arguments against complexity are sound, they are also sound against his own adaptationist argument. This means that despite his attempts to remain agnostic, Fodor is rather strongly committed as to what the mind/brain relationship must look like: it must be such that small changes in neurology do *not* produce wild discontinuities in cognition; otherwise, there is no explanation for the existence of highly articulated, contingent, true content. The Classical account of how that relationship might actually look, though, remains to be told.

5. Conclusion

What is the moral of Fodor's book? In a nutshell, "what our cognitive science has found out about the mind is mostly that we don't know how it works" (p. 100). It is important to notice that this claim depends on a possibly contentious reading of what would be required for us to understand how the mind works. The central argument concerning the frame problem, if sound, shows that there are many processes in mentation that cannot be implemented by classical, syntax-driven computations. The implication is that to *really* understand the mind, we would have to know what algorithm it implements. What we want, in short, is a mechanistically satisfying account of thought, and the current foundational ideology of cognitive science does not provide a complete one.

How seriously one takes this depends on how seriously one takes the demand for a mechanistic implementation theory of mentation. I can imagine that many cognitive scientists will simply be unmoved by the frame problem. Certainly those psychologists who are not in the business of providing formal or computational models of the phenomena they study may not see the threat. The practice of cognitive science is frequently much less reductionistic than its ostensible commitment to CTM might otherwise suggest.

What seems clear, though, is that the frame problem depends on whether or not mentation is global. Seventeen years on from the publication of MOM, there is still discouragingly little empirical evidence that bears on the question. Can we really, at any time, potentially draw on any element of our belief systems? Analogies between individual cognition and reasoning in science are *prima facie* persuasive, but philosophers and psychologists have been led astray in the past by just such persuasive considerations. What is needed to make clear the scope of the threat to cognitive science would be an investigation into the kind of data that subjects are actually able to draw on in making everyday inferences.

Fodor's book, then, may be premature in its pessimistic conclusions. But these conclusions are nevertheless forcefully argued. In many respects, Fodor's challenge to cognitive science is the same as Descartes': how could a machine do what *we* do? The strength of this challenge remains no matter what the right story about what we do turns out to be.

Notes

[1] This contrasts with empiricist psychology, according to which insofar as mental states have internal structure at all, their structure is a matter of associative integration only.

[2] To put (ii) another way: the logical form of mental states is *implemented* by (or realized in) the syntactic structure of the subjacent mental representations. The claim of supervenience, rather than identity, allows for the possibility that the MR that realizes a mental state might have a more richly articulated structure than the state that supervene on it. Perhaps some of the structure that it contributes is inaccessible to consciousness, for example.

[3] The presupposition that thinking is largely rational plays a strong role in motivating CTM, but it is never argued for in any detail, which seems like a serious omission given the large literature on human judgment that indicates many persistent failures of rationality. In any case, the fewer inferences that are rational, the less support there will be for CTM. However, there is another line of evidence for CTM that Fodor does not discuss in MDW: this is the well-known argument that runs from the systematicity of thought to the conclusion that thoughts have logical/syntactic structure (Fodor & Pylyshyn, 1988). Although discussion of systematicity is outside the bounds of this review, analogous remarks might be made about it, namely, to the extent that one doubts that thought is systematic, one will doubt that thoughts have syntactic structure.

[4] Fodor also calls these inferences "abductive," "holistic," and "inferences to the best explanation." I will continue to use "global" as the cover term for all of these.

[5] Here is a first approximation of the differences among the readings: a property is intrinsic to an object iff the object's having the property is independent of any other objects. A property is essential to an object's being *F* iff the object could not lose the property and still remain *F*. A property is context-independent iff an object has it in any context (what the context is may depend on how an object is described, and not every change in other objects may count as a change in context). Sorting out the metaphysical relations among these notions is a complex task, and seeing how they apply to the particular case of the syntax of MRs is equally difficult.

[6] Fodor often says that simplicity and centrality of theories (sets of propositional attitudes) are the properties to which global processes are sensitive. One might object that these are objective properties of theories, to which we have no guarantee of epistemic access; but I think that Fodor's point works just as well if one replaces them with *judged* simplicity and centrality. The question remains: how is this property detected in such a way that it can affect the course of thought?

[7] Strictly, this is not quite an acceptable way to put it, since the whole question of framing concerns what exactly is "relevant"' in any particular case. A rather radical move would be to conjoin all of the organism's beliefs in this way.

[8] It is an unfortunate feature of debates over modularity that theorists working in different fields tend to use it in very different ways. "Module" for linguists does not mean quite what it does for neuropsychologists, and it means something different again for developmental psychologists; and so on. For an attempt to separate out several notions of modularity that are often conflated in the literature, see Segal (1996).

[9] For exposition and criticism of the massive modularity hypothesis, see Samuels (1998).

[10] For examples, see the articles in Gazzaniga (2000).

[11] Maybe Fodor would respond that the relationship is linear between neurology and innate representations, but possibly nonlinear between neurology and innate architectural mechanisms. His analogy with Turing machines on p. 90 suggests this. However, this reply seems *ad hoc*. Why should there be such a metaphysical difference between content and process, even granted the Turing story about cognition?

References

COSMIDES, L. & TOOBY, J. (1992). Cognitive adaptations for social exchange. In J. BARKOW, L. COSMIDES & J. TOOBY (Eds) *The adapted mind* (pp. 163–228). Oxford: Oxford University Press.

FODOR, J. (1983). *The modularity of mind.* Cambridge, MA: MIT Press.

FODOR, J. & PYLYSHYN, Z. (1988). Connectionism and cognitive architecture: a critical analysis. *Cognition, 28,* 3–71.

GAZZANIGA, M.S. (2000). *The new cognitive neurosciences.* Cambridge, MA: MIT Press.

PREUSS, T.M. (2000). What's human about the human brain? In M.S. GAZZANIGA (Ed.) *The new cognitive neurosciences* (pp. 1219–1234). Cambridge, MA: MIT Press.

SAMUELS, R. (1998). Evolutionary psychology and the massive modularity hypothesis. *British Journal for the Philosophy of Science*, *49*, 575–602.

SEGAL, G. (1996). The modularity of theory of mind. In P. CARRUTHERS & P.K. SMITH (Eds) *Theories of theories of mind* (pp. 141–157). Cambridge: Cambridge University Press.

WEST-EBERHARD, M.J. (1992). Adaptation: current usages. In E.F. KELLER & E.A. LLOYD (Eds) *Keywords in evolutionary biology* (pp. 13–18). Cambridge, MA: Harvard University Press.