

Fodor and Pylyshyn on Connectionism¹

MICHAEL V. ANTONY

Department of Linguistics and Philosophy, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Abstract. Fodor and Pylyshyn (1988) have argued that the cognitive architecture is not Connectionist. Their argument takes the following form: (1) the cognitive architecture is Classical; (2) Classicalism and Connectionism are incompatible; (3) therefore the cognitive architecture is not Connectionist. In this essay I argue that Fodor and Pylyshyn's defenses of (1) and (2) are inadequate. Their argument for (1), based on their claim that Classicalism best explains the systematicity of cognitive capacities, is an invalid instance of *inference to the best explanation*. And their argument for (2) turns out to be question-begging. The upshot is that, while Fodor and Pylyshyn have presented Connectionists with the important empirical challenge of explaining systematicity, they have failed to provide sufficient reason for inferring that the cognitive architecture is Classical and not Connectionist.

Key words. Cognitive architecture, computationalism, Connectionism, implementation, inference to the best explanation, Language of Thought.

1. Introduction

In a recent essay, Jerry Fodor and Zenon Pylyshyn (1988) offer a critical analysis of the Connectionist enterprise. Their attack focusses on a pervasive psychological phenomenon they call 'systematicity', and on purported difficulties Connectionism faces in accounting for it. To properly explain systematicity, they argue, a Classical architecture is required. But Connectionist and Classical architectures, they maintain, are incompatible. Therefore the correct account of systematicity is not Connectionist. The only possible role left for Connectionism, consequently, is as a theory of how the cognitive architecture is implemented in brains. It can have no role in psychological explanation. In a nutshell,

... all the reasons for thinking [Connectionist models] ... might be true are reasons for thinking that they couldn't be *psychology* (*sic*). (p. 66)²

Fodor and McLaughlin (1990) put more or less the same point in the form of a dilemma: If Connectionist devices cannot display systematicity, then Connectionism fails as theory of the cognitive architecture. If they can display systematicity, that must be because they have implemented a higher-level Classical architecture – in which case it is the Classical and not the Connectionist architecture that shoulders the psychological explanatory burden. Either way, Connectionism fails as theory of the cognitive architecture. At best, it provides an implementation theory of that architecture.

At the most abstract level, Fodor and Pylyshyn's argument against Connectionism has the following form:

Minds and Machines 1: 321–341, 1991.

© 1991 Kluwer Academic Publishers. Printed in the Netherlands.

- (1) The cognitive architecture is a Classical architecture
- (2) Classical and Connectionist architectures are incompatible
- (3) Therefore, the cognitive architecture is not a Connectionist architecture³

Clearly, the argument is valid. Premise (1) is argued for on the basis of Fodor and Pylyshyn's claim that Classical architectures provide the best explanation of systematicity; and premise (2) is meant to follow from the natures of Classical and Connectionist architectures.

Fodor and Pylyshyn's paper makes for characteristically exciting reading. One ought to be struck, however, by how extraordinarily strong their thesis is. They are insisting that an active and vibrant research program in the study of cognition, contrary to its own stated aims, has no role whatever to play in psychological explanation, at least where such explanation involves reference to entities with intentional content. Strong indeed.

What Fodor and Pylyshyn clearly have done in their essay is saddle Connectionism with the challenge of explaining systematicity, a challenge it ultimately may be unable to meet. However, empirical challenges are frequently advanced in the sciences by proponents of competing theories, methodologies, etc. Such theories or methodologies, after all, are usually competitors in the first place because of attractive features they possess that others lack. Pointing to such features, moreover, is often tantamount to issuing a challenge to the other theories to do "as good a job" in the relevant respects. Connectionism, of course, is no exception in this regard.

There are several ways proponents of theories can respond to empirical challenges. Difficulties can be altogether ignored, noted and put temporarily on the shelf, or tackled with varying degrees of urgency. What rarely occurs is that everyone infers on the basis of an empirical challenge alone that the challenged theory is false. Theories typically get abandoned when empirical challenges go unmet over time, and accumulate to the extent that the theories no longer seem worth pursuing, given more attractive alternatives.

These considerations suggest that if Fodor and Pylyshyn are to successfully argue that the cognitive architecture is Classical and not Connectionist, they must have something up their sleeve(s) other than the point, important as it is, that systematicity must be explained by a theory of the cognitive architecture. As I shall argue, however, nothing they say entitles them to the strong conclusions they desire: their defenses of premises (1) and (2) fail to stand up. In particular, their argument for (1) involves an appeal to *inference to the best explanation* which requires a defense of additional premises they do not supply, and their argument for (2) turns out to be question-begging. That the cognitive architecture is Classical and not Connectionist, therefore, is not established.

I should emphasize that my aim in this paper is a modest one. My objective is to show that Fodor and Pylyshyn's argument against Connectionism fails, and not

so much to determine whether premises (1) and (2) are true. (I do, however, provide some reason for thinking (2) is false.) For all I know, the cognitive architecture is Classical and not Connectionist. It strikes me as worthwhile, however, to show that has not yet been shown, contrary to what Fodor and Pylyshyn would have us believe. At a minimum, doing so will allow attention to be more profitably focussed on the genuine empirical and conceptual issues that concern Connectionist and Classical architectures, and their role in cognitive scientific explanation.

The remainder of this essay is structured as follows. In the next section, I discuss the notions of a functional architecture and architectural framework, as well as instantiation and implementation, so that Fodor and Pylyshyn's main argument can be properly understood, and to provide some necessary stage-setting for my criticisms of it. That done, I take up their arguments for premises (1) and (2), respectively, in Sections 3 and 4.

2. Understanding Fodor and Pylyshyn's Main Argument

ARCHITECTURES AND FRAMEWORKS

Cognitive science treats minds as computational devices. In characterizing a device computationally (rather than physically or biologically, say), one specifies its *functional architecture* – roughly, the device's *fixed resources*, functionally described, at a level involving reference to semantic entities. Thus one points to the device's primitive symbolic operations, its representational capacities, the features of its memory, its control structure, and so on. Such a characterization sets out the constraints within which all computation must occur, and in that sense precisely determines what the device can do.

Though the notion of a functional architecture is not entirely well understood, functional architectures can be conceived for our purposes as being the same as virtual machines. Alternatively, a programming-language manual can be regarded as providing a detailed description of some particular functional architecture. Functional architectures, virtual machines, and programming languages, accordingly, can be taken as roughly equivalent, and should be contrasted with the algorithms or programs that get executed in them.

Insofar as the mind is treated as a computer, its functional architecture too must be characterized. The *cognitive architecture*, as it is typically called, is the functional architecture of the mind, and uncovering it, along with the algorithms executed in it, is the primary aim of cognitive science.⁴ Connectionism and Classicalism, therefore, are proposals about what the cognitive architecture is. More accurately, they are proposals about certain properties of the cognitive architecture, and thus are better understood as cognitive architectural *frameworks*, where considerable detail still needs to be fleshed out.

For a functional architecture to be a Classical architecture, according to Fodor

and Pylyshyn, it is necessary and jointly sufficient that there be representations with a *combinatorial syntax and semantics*, and *structure sensitive processes* that operate on those representations (pp. 12–13). To have a combinatorial syntax and semantics is to possess a *Language of Thought*, and to have that is to satisfy the following three conditions:

(a) there is a distinction between structurally atomic and structurally molecular representations; (b) structurally molecular representations have syntactic constituents that are themselves either structurally molecular or are structurally atomic; and (c) the semantic content of a (molecular) representation is a function of the semantic contents of its syntactic parts, together with its constituent structure. (p. 12)

Given a functional architecture with a Language of Thought, the remaining feature of the Classical framework, essentially, is that structure sensitive processes operate on the representations in virtue of the representations' structural or formal properties, and in such a way that appropriate semantic relations are preserved.

Connectionism also provides an architectural framework for the study of cognition. It offers a somewhat more detailed account of the cognitive architecture than Classicalism, however, being closer in degree of detail to *types* of Classical architectures like production systems, for example, than to the defining features of the Classical framework itself. Because Connectionism is an evolving research program, however, rather than a notion defined on the basis of a small set of properties, like Classicalism; and because there are serious difficulties with our current understanding of Connectionist computation;⁵ it is inappropriate at this time to try to define 'Connectionism', or provide necessary and sufficient conditions on Connectionist architectures. Instead I shall simply list the basic architectural features commonly mentioned in accounts of Connectionism. (This may come close to constituting a set of sufficient conditions on Connectionist architectures, though I shall not press the point.)

Architectures in the Connectionist framework consist of a set of simple processing units that take on various activation values, and a pattern of connectivity among the units. In general, individual units receive input from and send output to several other units. Each connection between pairs of units has associated with it a "strength" or "weight" that determines the effect one unit has on the other. At any given time, a unit's activation value is a function of its previous level of activity, and the degree of activity (excitatory or inhibitory) currently being received along its input connections. Finally, the representations in Connectionist architectures are either activated individual units – local representations – or patterns of activity across multiple units – distributed representations.⁶

INSTANTIATION, IMPLEMENTATION, AND THE MAIN ARGUMENT

Consider now the premises and conclusion of Fodor and Pylyshyn's main argument, beginning with premise (2).

(2) Classical and Connectionist architectures are incompatible.

Since Classicalism and Connectionism are frameworks for, or properties of, functional architectures, (2) amounts to the claim that the classical and Connectionist frameworks cannot both be instantiated in the same functional architecture. Put slightly differently, no Connectionist functional architecture can be an instance of the Classical framework (have the property of having representations with a combinatorial syntax and semantics, and structure sensitive processes), nor can any Classical functional architecture be an instance of the Connectionist framework (have whatever suffices for that: roughly, the properties listed above).

This notion of *instantiation* (where, e.g., a Connectionist architecture instantiates the Classical framework) must be distinguished from a quite different notion with which it can easily be confused – the notion of *implementation* (where a Connectionist architecture implements a Classical architecture, say). ‘Instantiation’ expresses a simple relation between individuals and properties: an individual *i* instantiates a property *P* if and only if Pi .⁷ Where individual computational models or functional architectures are under consideration, therefore, what is relevant is whether those models or architectures have certain properties – those, for example, that define the Classical or Connectionist frameworks. Thus one might ask of a particular Connectionist architecture whether it has the property of having a combinatorial syntax and semantics and structure sensitive processes, i.e., whether it instantiates the Classical framework. Or one might ask of a specific Classical architecture whether its properties also suffice for its being Connectionist, i.e., whether that same Classical functional architecture has interconnected processing units that take on various activation values, etc. In the case of instantiation, therefore, a *single* model or architecture is involved, and what is in question are its properties.

Where *implementation* is at issue, on the other hand, *two* functional architectures must be considered. A functional architecture FA1 is implemented, if at all, by the execution of a program in a distinct functional architecture FA2. Intuitively, the primitive operations, representational structures, etc. of FA1 get “made up” or “constructed” out of the resources of FA2. This is the relation that typically exists, for example, between assembly language functional architectures, on the one hand, and higher-level architectures like LISP or Pascal, on the other, when the latter are up and running on a computer. With regard to the question whether a particular Connectionist architecture implements a Classical architecture, therefore, what is being asked is not whether the Connectionist architecture has the property of having a combinatorial syntax and semantics, and structure sensitive processes (i.e., whether it *instantiates* the Classical framework). The issue, rather, is whether it is constructing the resources of some *distinct* functional architecture with that property.

The difference between instantiation and implementation can be brought out more clearly by noting that functional architectures cannot implement architectur-

al *frameworks*, for instance, the Classical and Connectionist frameworks. Implementation is a relation between functional architectures, and the Classical and Connectionist frameworks are not functional architectures. To say a device has a Language of Thought and structure sensitive operations, for example, is not yet to specify a functional architecture for that device. To say that implies nothing about what the device can compute. At a minimum, one must specify, in addition, *which* operations the device has at its disposal, what are its representational capacities, what is the nature of its memory, and so forth.

Notice further that in cases of implementation, lower-level architectures typically do not instantiate the characteristic properties of higher-level ones. An assembly-level architecture implementing LISP, for instance, does not also instantiate LISP: it lacks the necessary primitive operations (e.g., CAR, CDR), and has primitive operations LISP lacks (e.g., various operations on the contents of the accumulator). And, of course, it is also true that an instantiation of LISP need not implement any distinct, higher-level LISP architecture. The notions of instantiation and implementation, therefore, are mutually independent.

There will be a bit more to say of instantiation and implementation in Section 4. For now, the point to be emphasized is that it is instantiation, and not implementation, that is relevant to a proper understanding of premise (2) – the claim that Classical and Connectionist architectures are incompatible. For if the cognitive architecture is Classical, and Connectionist architectures can instantiate the Classical framework, it cannot be inferred that the cognitive architecture is not Connectionist. The validity of Fodor and Pylyshyn's argument, therefore, depends on construing the notion of incompatibility in the way I have suggested. And, at any rate, it is agreed on all sides that Classicalism and Connectionism are compatible in the sense that Connectionist architectures can implement Classical architectures, and vice versa.

I turn now to premises (1) and (3).

(1) The cognitive architecture is a Classical architecture.

(3) The cognitive architecture is not a Connectionist architecture.

To say that the cognitive architecture is a Classical architecture is to say this: that the cognitive architecture instantiates the Classical framework; that among all of its properties, one is that it has a combinatorial syntax and semantics, and structure sensitive processes. And to say that the cognitive architecture is not Connectionist is to deny it instantiates Connectionism's defining features. These claims are best considered in the context of the familiar picture according to which there are multiple levels of functional architecture, with each architecture being implemented by the architecture one level down, the highest level being the level of cognitive architecture, and the lowest being realized directly in hardware.⁸ Premise (1), accordingly, asserts that the level of cognitive architecture is

Classical, quite aside from whether functional architectures at any other levels are. And (3) claims that the properties that define the Connectionist framework are not instantiated at the level of cognitive architecture, even though, for all we know, they may be at some lower level.

So the issue between Classicalism and Connectionism is over the properties of one particular level of functional architecture – the cognitive architecture. That is the level of psychological explanation, the level it is the aim of cognitive science to characterize. And the proponents of both frameworks want to claim that level as their own.

On now to Fodor and Pylyshyn's main argument.

3. Is the Cognitive Architecture Classical?

Consider this thesis: Given the natures of the Classical and Connectionist frameworks, and of systematicity, it follows that a correct account of systematicity *must* be Classical and *cannot* be Connectionist. It is easy to get the impression that that is what Fodor and Pylyshyn have argued. To my knowledge, however, their paper contains no argument that a Classical architecture is necessary for systematicity; only that the correct explanation of systematicity (in humans, and perhaps other species) is in fact Classical. That argument is their defense of premise (1), and is the principal topic of this section.

Their essay does, however, include an argument that Connectionist architectures cannot explain systematicity (as contrasted with the weaker thesis that the correct explanation is not Connectionist in fact). That argument, along with the assumption that systematicity is a genuine psychological phenomenon, allows for a more direct attack on Connectionism than their main argument, in that Classicalism need not be invoked at all:

- (1') Any adequate theory of the cognitive architecture must explain systematicity.
- (2') Connectionism cannot explain systematicity.
- (3) Therefore, the cognitive architecture is not a Connectionist architecture.

The argument for (2') is presented briefly, almost as a side issue. However, its inclusion, I believe, helps explain one's strong feeling that Connectionism is meant to be ruled out somehow *in principle*. I discuss that argument at the end of this section. For now, I turn to the main argument, and, in particular, premise (1).

Fodor and Pylyshyn's defense of (1), that the cognitive architecture is Classical, is meant to be empirical and nondemonstrative. They describe the phenomenon of systematicity (discussed presently), and argue it is best explained on the assumption that the cognitive architecture is Classical. Then, by inference to the

best explanation (ITTBE), they conclude the cognitive architecture is Classical.⁹ Since Fodor and Pylyshyn do not explicitly state their argument is an instance of ITTBE, some evidence that that is indeed what they have in mind is in order.

Fodor and Pylyshyn's systematicity argument for Classicalism is meant to "precisely parallel" their systematicity argument for the constituent structure of sentences (p. 39). However that argument clearly is an instance of ITTBE. They write,

There is . . . a straightforward argument from . . . the systematicity of language capacity to the conclusion that sentences must have syntactic and semantic structure: . . . *in effect systematicity follows from the postulation of constituent structure . . .*

On the view that sentences are atomic, the systematicity of linguistic capacities is a mystery; on the view that they have constituent structure, the systematicity of linguistic capacities is what you would predict. *So we should prefer the latter view to the former.* (p. 38) (Emphases mine.)

Also, in *Psychosemantics*, Fodor (1987) offers much the same argument for (1) as Fodor and Pylyshyn's. Focussing on the need for a Language of Thought, Fodor says,

I take it that what needs defending here is . . . *not* the idea that the systematicity of cognitive capacities implies the combinatorial structure of thoughts. I get . . . [that] claim for free *for want of an alternative account*. [Fodor (1987), p. 151] (Second emphasis mine.)

I assume this suffices to show that Fodor and Pylyshyn have an ITTBE argument in mind. I propose now to examine it. Here is the ITTBE argument for (1):

- (.5) Systematicity, a phenomenon to be explained by the cognitive architecture, is best explained by a Classical architecture.
- (1) Therefore, the cognitive architecture is a Classical architecture.

Prima facie, this looks like a perfectly legitimate utilization of ITTBE. I shall argue presently that it is not. First, however, I consider the phenomenon of systematicity itself, and how Classicalism explains it.

CLASSICALISM AND SYSTEMATICITY

Fodor and Pylyshyn introduce the notion of systematicity thus:

The easiest way to understand what the systematicity of cognitive capacities amounts to is to focus on the systematicity of language comprehension and production . . .

What we mean when we say that linguistic capacities are *systematic* is that the ability to produce/understand some sentences is *intrinsically* connected to the ability to produce/understand certain others. You can see the force of this if you compare learning languages the way we really do learn them with learning a language by memorizing an enormous phrase book . . . [The] point is . . . that you can learn *any part of a phrase book without learning the rest* . . . [But you don't], for example, find native speakers who know how to say in English that John loves the girl but don't know how to say in English that the girl loves John. (p. 37)

And a bit further on,

What does it mean to say that thought is systematic? Well, just as you don't find people who can understand the sentence 'John loves the girl' but not the sentence 'the girl loves John', so too you don't find people who can *think the thought* that John loves the girl but can't think the thought that the girl loves John. (p. 39)

So the systematicity of linguistic capacities, and cognitive capacities more generally, seems to involve a kind of *relation* among capacities, one Fodor and Pylyshyn illustrate with conditionals of the following form:

If one has the capacity to say/understand/think that *Rab*, one also has the capacity to say/understand/think that *Rba*.

Fodor and Pylyshyn describe this relation as an "intrinsic connection", but do not say what that is. Presumably, conditionals of the above form are not meant to be true by logic, for then no explanatory work would be left for Classicalism. Nor are they material conditionals, since they would be counted as false where the antecedent and consequent are both true by accident, for example. It appears, therefore, that the relation must be understood as a *nomic* one, where the capacities to think that *Rab*, and to think that *Rba*, for example, always co-occur as a matter of natural law.¹⁰

Now how is Classicalism supposed to explain systematicity? The basic story goes like this. Consider a system that displays systematicity in that *Rab* can be represented if and only if *Rba* can, as a matter of psychological law. Suppose the system's functional architecture instantiates a Language of Thought in virtue of the following: The representations '*Rab*' and '*Rba*' are made up of the constituents '*R*', '*a*', and '*b*', and the meanings of '*Rab*' and '*Rba*' are a function of their structure, and the meanings of the constituents. If we now imagine a structure sensitive operation or process that can construct '*Rab*' from '*R*', '*a*', and '*b*' if and only if it can do the same with respect to '*Rba*', we are home. Such operations, however, are easy to come by. An operation whose only constraint on the construction of complex representations is that they have the structure *predicate-name-name* is a case in point. The systematicity of the system's representational capacities is thus explained.

THE TROUBLE WITH FODOR AND PYLYSHYN'S ITTBE ARGUMENT

My criticism of Fodor and Pylyshyn's ITTBE argument is straightforward. We have seen that Classicalism can account for systematicity. Let us assume it provides the *best* account. Still, it cannot correctly be inferred that the cognitive architecture is Classical.

Here is why. In the domain of cognitive science, there are many phenomena other than systematicity that must be explained by the cognitive architecture and the algorithms running in it. Fodor and Pylyshyn must consider those phenomena

too, and for the following reasons. Suppose there were some non-Classical architecture that explained a particular phenomenon (or range of phenomena) better than any Classical architecture. Surely, under those circumstances, it would be wrong to infer that the cognitive architecture is Classical *just because Classicalism best explains systematicity*. For proponents of the non-Classical architecture would be equally entitled to conclude on the basis of its explanatory success(es) that the cognitive architecture instantiates *it*. Assuming a single cognitive architecture must explain both systematicity and the other phenomenon,¹¹ however, it cannot be both that the cognitive architecture is Classical and that it is non-Classical. That is a contradiction. Now because the above sort of scenario is possible, it follows that inferring (1) on the basis of (.5) alone is a mistake. Further premises are required.

More generally, here is what has gone wrong. Consider any theory T , the aim of which is to explain a domain of phenomena $D = \{P_1, P_2, \dots, P_n\}$. And suppose T provides the best explanation of the phenomenon P_1 . From that supposition alone, ITTBE can be used to infer the truth of T with respect to D *only if P_1 is the only element in D* . For otherwise additional premises are required: like that (i) there is no other theory T^* in competition with T to explain the phenomena in D ; or (ii) if there is such a theory, there is no phenomenon P_i in D that T^* explains better than T ; or (iii) if there is such a phenomenon, T 's explanation of P_1 is sufficiently better than T^* 's explanation of P_i to warrant concluding that T is true and T^* false; and so on. The worry, of course, is that if there is more than one element in D , some other theory T^* may do an "overall better job" than T in explaining the phenomena in D , and thus would deserve to be counted as true, *even though T best explains P_1* . In the absence of premises like (i), (ii), or (iii), therefore, ITTBE can result in blatantly false conclusions, as well as contradictions if applied to incompatible theories.

In saying further premises are required, one ought not to suppose I am demanding a proof for a conclusion that was meant to be argued for only *nondemonstratively*; or that I am requiring more than is reasonable from an ITTBE argument. Even if (i), (ii), or (iii) are true, it still does not follow that T is true. A proof, therefore, is not what is needed. The additional premises, rather, are to ensure that a theory that must explain a *cluster* of phenomena is not deemed true just because it best explains a single phenomenon in that cluster. It is conceivable, after all, that a theory of magic goblins might provide the best explanation of some natural phenomenon for which there as yet exists no alternative account. But that, of course, is no reason to believe in goblins. Non-goblin theories are too much better at explaining other things. To be accepted, therefore, a goblin theory would have to prove itself with a much wider range of phenomena. And that surely is a reasonable demand to make of it. Now Classicalism is significantly more appealing than Goblinism, but the point is the same.

The upshot is that Fodor and Pylyshyn need something more than (.5) to get

(1), since there are phenomena other than systematicity in the domain of cognitive science to be explained. And because there is a theoretical framework in competition with Classicalism – Connectionism – Fodor and Pylyshyn require a premise like (ii) or (iii) above. They must argue either that (a) there is *nothing* (non-Classical) Connectionist architectures can do better than Classical architectures; or if there is, that (b) when all the points are tallied, Classicalism comes out sufficiently ahead of Connectionism to warrant accepting Classicalism as true of the cognitive architecture, and Connectionism as false.

The prospects for establishing either (a) or (b) at this time would appear dim. On the surface at least, it seems there *are* things Connectionist models do better than Classical ones (i.e., Connectionist models that are not themselves Classical). Some examples: error-tolerant, content-addressable memory; graceful degradation; flexible and computationally affordable reasoning in terms of prototypes and emergent schemata; and so on.¹² If Fodor and Pylyshyn are to argue for (a), therefore, they must argue that none of the above phenomena, or any others, are best explained by Connectionist architectures.¹³

If, on the other hand, Fodor and Pylyshyn agree there are things Connectionism does better than Classicalism, and attempt instead to argue for (b), their prospects for success would seem dimmer still. For unless one of the two frameworks is obviously the superior, it is hard to see how Fodor and Pylyshyn could calculate the relative strengths and weaknesses of Connectionism and Classicalism in the absence of a general theory of how to select among competing scientific theories. And of course there is no such theory. Indeed, it would seem the best gauge we have of the relative merits of competing theories is in terms of the amounts of support they get in the relevant scientific communities. With regard to Classicalism and Connectionism, however, there is nothing even approaching a general consensus among cognitive scientists. It is too early. The prospects for establishing (b), therefore, seem not very promising at best.

In any event, the burden falls squarely upon Fodor and Pylyshyn to provide an argument for either (a) or (b) if they wish to infer that the cognitive architecture is Classical. That Classicalism best explains systematicity simply is not enough. I conclude that Fodor and Pylyshyn's ITTBE argument for (1) fails. That, of course, is not to say (1) is false, or even that no sound ITTBE argument for (1) is possible. It is just to say that Fodor and Pylyshyn have thus far failed to provide sufficient reason for thinking (1) is true. While they have presented Connectionists with the important challenge of explaining systematicity, that is no different in kind from empirical challenges facing Classicalism. Accordingly, Classicalism gains points, but not victory.

AN ARGUMENT AGAINST CONNECTIONISM IN PRINCIPLE

I mentioned earlier that Fodor and Pylyshyn's paper contains, in addition to their main argument, the following argument against Connectionism:

- (1') Any adequate theory of the cognitive architecture must explain systematicity.
- (2') Connectionism cannot explain systematicity.
- (3) Therefore, the cognitive architecture is not a Connectionist architecture.

In this paper, I assume that Fodor and Pylyshyn have drawn our attention to a genuine psychological phenomenon – systematicity – that must be explained by any adequate account of the cognitive architecture. Accordingly, I concern myself in what follows only with premise (2').

The argument that Connectionism cannot explain systematicity as a matter of principle is presented in the context of Fodor and Pylyshyn's positive ITTBE argument for Classicalism.¹⁴ Here are the relevant passages:

... as far as Connectionist architecture is concerned, there is nothing to prevent minds that are arbitrarily unsystematic. But that result is *preposterous*

[Now it's] possible to imagine a Connectionist being prepared to admit that while systematicity doesn't *follow from* – and hence is not explained by – Connectionist architecture, it is nonetheless *compatible* with that architecture

[But] it's not enough for a Connectionist to agree that all minds are systematic; he must also explain *how nature contrives to produce only systematic minds*. (pp. 49–50)

The embedded argument seems to be this:

- (A) For all theories *T* and phenomena *P*, a necessary condition for *T* to be an explanation of *P* is that *P* follows from *T*. ('... systematicity doesn't *follow from* – and hence is not explained by – Connectionist architecture').
- (B) Systematicity does not follow from Connectionist architecture, since both the presence and absence of systematicity are compatible with it.
- (C) Therefore, Connectionist architecture cannot explain systematicity (2').

Like Fodor and Pylyshyn's main argument, this too is valid. And it is against Connectionism *in principle* in the sense that the conclusion follows just from claims about the nature of explanation, and the nature of Connectionist architecture; excluded is the possibility that a Connectionist account of systematicity might be discovered some time in the future. As I now shall argue, however, depending on how 'Connectionist architecture' is interpreted in (B) and (C), either the argument is sound but uninteresting, premise (B) is false, or (B) cannot be evaluated. Either way, then, it fails.

'Connectionist architecture', as it appears above, can be understood to refer either to the general features that define the Connectionist architectural framework, to features of particular Connectionist models, or to features of subclasses of such models. If it is the Connectionist framework that is meant – the properties in virtue of which models or functional architectures are or are not

Connectionist – then the argument is sound¹⁵ but harmless. For any Connectionist account of systematicity will rely on features of a particular Connectionist model, or of a subclass of such models. But then arguing that Connectionism cannot explain systematicity because the presence and absence of systematicity are compatible with the Connectionist *framework* is like arguing the laws of physics do not explain why a leaf falls as it does since it is consistent with those laws that it might have fallen some other way. The criticism is misdirected.¹⁶

If, instead, Fodor and Pylyshyn mean any possible Connectionist model by ‘Connectionist architecture’, then premise (B) is false. For, as a rule, each Connectionist model will be such that it either displays systematicity or not, and so it will be false that both the presence and the absence of systematicity are compatible with it. Finally, by ‘Connectionist architecture’ Fodor and Pylyshyn might mean subclasses of Connectionist models. If it is any possible subclass they have in mind, then (B) is false, since there are subclasses containing single models as their sole member. If they mean particular subclasses of Connectionist models, however, then since they do not tell us which those are, we have no way of determining whether the presence and absence of systematicity is compatible with them.

I conclude that the argument fails. There is no sound argument against Connectionism in principle to be found in Fodor and Pylyshyn’s essay.¹⁷

4. Are Classicalism and Connectionism Incompatible?

Although Fodor and Pylyshyn’s failure to support (1) shows their main argument is unsound, it is worth considering independently their defense of (2) – the claim that Classicalism and Connectionism are incompatible. For it appears their argument does not hit the mark there either, and so falters on two counts. I shall first criticize Fodor and Pylyshyn’s defense of (2), and then offer some reasons for thinking (2) might be false. I end with a few remarks on what the falsity of (2) would imply for the Connectionism-Classicalism debate.

FODOR AND PYLYSHYN’S DEFENSE OF (2)

To claim that Classicalism and Connectionism are incompatible, once again, is to claim that they cannot both be instantiated in the same functional architecture; or what is the same, that there can be no Connectionist architecture that is an instance of the Classical framework, and vice versa. Now this claim is a logical or metaphysical one. It is two sets of properties – those that define the Classical and Connectionist frameworks respectively – that are being said to be incompatible, just as being a square is incompatible with being a circle. The way one argues for such incompatibility, then, is by showing that nothing could simultaneously satisfy the properties of both sets.

Fodor and Pylyshyn begin in the second section of their paper by defining ‘Classicalism’ in the manner discussed earlier: in short, instantiating the Classical

framework amounts to having a Language of Thought and structure sensitive processes. Next, they describe a simple Connectionist network that draws inferences from A&B to both A and B, and contrast that with a Classical machine that does the same. Their argument for the incompatibility of Connectionism and Classicalism, if indeed they have one, is contained somewhere in their discussion of those models. For immediately following that discussion they embark upon a diagnosis of how their conclusion might have been missed: they discuss four ways one might mistakenly think Connectionist models are Classical.

I am able to find no explicit argument for (2) in that discussion. However, perhaps there is an argument implicit in what they say. If there is, then since they are arguing from a discussion of particular models, it would seem the argument must have more or less the following form:

From our knowledge of the defining properties of Classicalism, with the aid of the particular Classical model before us, we see that the Connectionist model in the example is not an instance of the Classical framework. That we do, presumably, by noting that the Connectionist model has a property P that is inconsistent with the defining properties of Classicalism. Finally, we ask whether it is plausible that P is a necessary condition on Connectionist architectures. If and only if it is, (2) follows.¹⁸

The Connectionist network employed in Fodor and Pylyshyn's discussion is illustrated in Figure 1. Though the diagram looks simple enough, its interpretation is in fact not straightforward. The reason is that instead of the circles representing units, as they would in a conventional Connectionist network diagram, they represent what Fodor and Pylyshyn call 'nodes', a node being anything that is either a unit or an aggregate of units. The lines connecting the circles, accordingly, represent either individual connections between units, or multiple connections between aggregates of units. Fodor and Pylyshyn adopt their

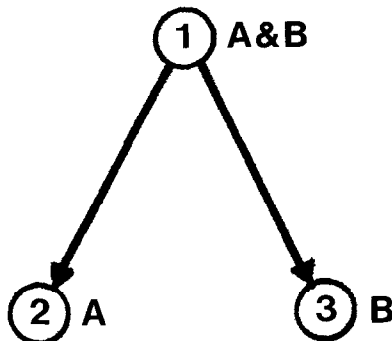


Fig. 1. Fodor and Pylyshyn's Connectionist network.

terminology, we are told, so their remarks can apply generally across models with either local representations or distributed representations (p. 5, fn. 1).

The network depicted in Fodor and Pylyshyn's diagram has three nodes, representing the propositions A&B, A, and B, respectively. There are connections from the A&B node to the A node and to the B node. The idea underlying the network's operation is that when the A&B node becomes activated, activation is caused in both the A and B nodes. Thus the network correctly infers A and B from A&B.

Fodor and Pylyshyn provide a brief description of how a simple Turing machine would make the same inferences, and they spell out the differences between the Connectionist and Classical machines thus:

In the Classical machine, the objects to which the content A&B is ascribed (viz., tokens of the expression 'A&B') literally contain, as proper parts, objects to which the content A is ascribed (viz., tokens of the expression 'A'). Moreover, the semantics . . . of the expression 'A&B' is determined in a uniform way by the semantics of its constituents. By contrast, in the Connectionist machine none of this is true; the object to which the content A&B is ascribed (viz., node 1) is causally connected to the object to which the content A is ascribed (viz., node 2); but there is no structural (e.g., no part/whole) relation that holds between them. In short, it is characteristic of Classical systems, but not of Connectionist systems, to exploit arrays of symbols some of which are atomic . . . but indefinitely many of which have other symbols as syntactic and semantic parts . . . (p. 16)

It would seem Fodor and Pylyshyn's argument for (2) must be in that passage, if anywhere. For immediately following it is the diagnosis I mentioned above. For the sake of clarity, rather than discussing their network of nodes, I shall discuss, in turn, the localist and distributed versions of their network. It will be easier that way to evaluate whether Fodor and Pylyshyn have made their case for (2).

First the localist version. On that interpretation there is one unit which, when activated, represents A&B, and causes two other units to become activated, one of which represents A, and the other B. Now in this case it is obvious the model does not instantiate the Classical framework since it lacks a Language of Thought. And it lacks a Language of Thought because it has no complex representations. Each of the three representations in the network is realized in an individual unit, and individual units are structurally simple. But what follows? Just that Classicalism is incompatible with the localist variety of Connectionism, since every possible localist model is such that each representation in the model is realized by a single, structurally simple unit.

Consider now the distributed version of the network. Here things immediately break down. For, in the distributed version, the details of the model are far too underspecified to determine whether or not it instantiates the Classical framework. In the localist model the matter is clear: its failure to instantiate Classicalism is entailed by the fact that each representation is realized by a single unit. So when Fodor and Pylyshyn tell us the semantics of the A&B representation is not determined by the semantics of its parts, we know exactly why: it has no parts. But how do we know the same holds when the A&B representation is understood

as realized in a pattern of activity across an aggregate of units? Is it obvious the units of the aggregate could not have semantic values that determine, along with the structure of the aggregate, the semantic value of the whole? And it is obvious the behavior of the network could not be understood as involving processes that are sensitive to the representation's structure? Fodor and Pylyshyn say that is so, but they do not show it.¹⁹ In the absence of such a demonstration, therefore, Fodor and Pylyshyn have simply begged the question; they have not established the incompatibility of Classicalism and Connectionism.

IS (2) TRUE?

Fodor and Pylyshyn's discussion of the incompatibility of Classicalism and Connectionism is concerned primarily with showing that Connectionist architectures cannot have Languages of Thought. Though they also argue that Connectionist architectures cannot have structure sensitive processes, that argument merely rests on their claim that there are no structured representations to be sensitive to. I shall now argue that Connectionist architectures can instantiate Languages of Thought, as Fodor and Pylyshyn have defined that notion. I shall not discuss structure sensitive operations in any detail, however. Nonetheless, in arguing that Connectionist models can have Languages of Thought, some reason will have been given for thinking Connectionism and Classicalism might be compatible after all.

Imagine a distributed representation in some network that is composed of three units which, when activated, have the meanings *John*, *loves*, and *Mary*. If it is granted that this is a possible scenario, then to determine whether a Language of Thought is involved, the crucial question that must be asked is this: Do the three units have a *structure* which, along with the units' meanings, determines a unique meaning for the distributed representation as a whole (e.g., that John loves Mary rather than that Mary loves John)? If they do, then it would seem there is everything one needs for a Language of Thought. There is a complex representation, the meaning of which is a function of the meanings of its parts and their structure.

I would submit that the three units along with their meanings *can* have structure in the required sense. For all it is for representational parts to be structured so that a unique meaning *M* is determined is this:

The parts have nonsemantic properties that play a role in causing the machine to behave in ways that *make sense*, given that the complex representation means *M*.²⁰

However, that can easily be made to hold of the three units in the network we are imagining. The relevant properties of the units would be (some of) their particular

connections, and their weights. And the network could be designed so that those properties are causal determinants of certain inferences computed by the network, inferences that make sense (e.g., are valid) on the assumption that activation of the three units means that John loves Mary.

Suppose, for example, that as a consequence of the three units being activated, three others with the meanings *John*, *loves*, and *someone*, from some other distributed representation, become activated; as do three units from a third representation with the meanings *Mary*, *is-loved-by*, and *someone*. Given that the meanings of the units from the three distributed representations are fixed,²¹ we can make sense of what the network is doing if we assume the initial distributed representation means that John loves Mary: it is validly inferring that John loves someone, and that Mary is loved by someone. If we assume it means that Mary loves John, however, the network's behavior is unintelligible.

It seems quite natural, therefore, to say that the structural properties of the three units, along with their meanings, determines that the distributed representation means *John loves Mary*; and that because the network is causally sensitive to that structure, it is able to make the valid inferences it does. This is just the sort of talk that gets generated with respect to more standard Language of Thought instantiations, like Turing machines, for example. However, I see no relevant difference between the two kinds of machines in the senses in which they have structured representations.

Consider the sense in which complex representations in Turing Machines are structured. Suppose that upon three squares of a Turing machine's tape there is written, from left to right, 'John', 'loves', and 'Mary'. And suppose the machine is causally sensitive to the fact that the John-square is to the left of the Mary-square. Clearly, this does not suffice to determine that the complex representation means John loves Mary, rather than Mary loves John, or vice versa. It is still required that the *effects of the machine's sensitivity to that fact* be considered in order to determine what the representation means, just as it is with the simple Connectionist machine I have described.

Consequently I see no reason to deny that the Connectionist machine instantiates a Language of Thought, as Fodor and Pylyshyn have defined that notion. There is some reason, therefore, to think (2) is false. One must be careful, however, not to infer from these considerations that there can be Connectionist architectures with Languages of Thought that account for systematicity. For, as we have seen, having a Language of Thought does not in itself suffice for systematicity; at a minimum, appropriate structure sensitive processes are required in addition (or at least they are in the standard sorts of Classical models). Perhaps there are also other requirements Connectionist architectures cannot meet. The model I discussed, for example, is extraordinarily simple. However, it is plausible that Languages of Thought with a certain degree or kind of *complexity* are needed for systematicity, a complexity Connectionist architectures are unable to realize. All this, however, remains to be seen.

CONCLUDING REMARKS

Since Fodor and Pylyshyn's defense of (2) is unsatisfactory, and since there is some reason to believe Connectionism and Classicalism are compatible, we are left with the epistemological possibility that a Connectionist instantiation of the Classical framework will constitute the best theory of the cognitive architecture. If so, the cognitive architecture will be both Classical and Connectionist. I would like to end by briefly considering what the implications of such a scenario would be for the debate between Connectionism and Classicalism.

Consider first whether there is any sense in which Connectionism will have suffered a defeat if the cognitive architecture turns out to be a Connectionist instantiation of the Classical framework. For, after all, a supporter of Classicalism might reason, the cognitive architecture is Classical, so Connectionism cannot really have anything new to offer cognitive science, all claims to the contrary notwithstanding.

Such reasoning, however, is confused. If the cognitive architecture turns out to be Classical, and the best Connectionism can do is *implement* that architecture, then Connectionism *will* have suffered a defeat. (Recall the implementation/instantiation distinction discussed in Section 2.) For such a Connectionist implementation will be a *mere* implementation of the cognitive architecture in two respects: first, it will be irrelevant for the purposes of psychological explanation, in that it can be more or less ignored;²² and second, though it may be true of how the cognitive architecture is implemented in fact, it is still possible that the cognitive architecture can be implemented in many different ways, just as LISP or Pascal can.

But the issue at hand has nothing to do with whether Connectionist models can at best *implement* the cognitive architecture (which we are assuming to be Classical). The issue, rather, has to do with whether the cognitive architecture *is* a Connectionist architecture, one which happens also to be Classical. If it is, then Connectionism is crucially relevant to psychological explanation. The explanations of psychological phenomena will be Connectionist explanations through and through; not explanations in terms of some distinct, implemented Classical architecture, like a production system, for example. Connectionism, therefore, will hardly have been defeated.

There is one sense, however, in which Connectionism, or more precisely the views of certain Connectionists, will have suffered a defeat under the circumstances we are imagining. I have in mind the many claims in the Connectionist literature of psychological phenomena that can be explained by Connectionist models without appeal to symbol systems, complex representations, Languages of Thought, etc. If Classical architectures are needed for those explanations, such claims will of course be false. This kind of defeat, however, is not serious. For, as we have seen, it does not imply in the least that Connectionism can be ignored for the purposes of psychology, that it has nothing new or interesting to offer, that it should be denied funding, or anything else of the sort.

It should be clear by now that even if Connectionism attempts to model the cognitive architecture by instantiating the Classical framework, that in no way excludes the possibility that Connectionist architectures might differ in interesting and important ways from all other Classical architectures developed thus far. It should also be clear just how abstract an architectural framework Classicalism is. Even if the cognitive architecture turns out to be Classical, the vast majority of psychological phenomena will properly be explained by far more specific details than those embodied in the definition of the Classical framework.

Notes

¹ Thanks to Ned Block, Rob Cummins, Martin Davies, Eric Lormand, Gabe Segal, Bob Stalnaker, and two reviewers from *Minds and Machines* for many helpful comments and criticisms. I also gratefully acknowledge the Social Sciences and Humanities Research Council of Canada for funding part of this work.

² All page references are to Fodor and Pylyshyn (1988) unless otherwise noted.

³ I call this their 'main argument'. Their arguments for (1) and (2) appear in the third and second sections of their paper, respectively. Though many seem to doubt whether Fodor and Pylyshyn have the main argument in mind, Fodor and Pylyshyn are clear on this: "... Second, we try to say what it is that makes Connectionist and Classical theories of mental structure incompatible. Third, we review and extend some of the traditional arguments for the Classical architecture." (pp. 6-7).

⁴ See Pylyshyn (1984) and Haugeland (1985) for elaboration.

⁵ For example, the standard conception of computation in terms of *symbol manipulation*, which involves both symbols and distinct operations on them, does not naturally apply to Connectionist machines. Certain Connectionists, e.g., Smolensky (1988), even go so far as to deny that symbol manipulation occurs in Connectionist machines at all. Getting clear on these matters, it seems to me, is an important philosophical task. For some interesting work along these lines, see Cummins (forthcoming) and Lormand (1990).

⁶ There is some question whether to count activity in so called "hidden units" as representational, and also how to determine what the representational content is, if such activity is so counted. Moreover, in addition to unit activity, the weights of connections are often said to represent some of the system's knowledge, memories, etc. These are just a few of the murky issues I alluded to in the previous paragraph. I shall ignore them in this essay.

For more detailed accounts of the Connectionist framework, see, e.g., Rumelhart, McClelland, and the PDP Research Group (1986), McClelland, Rumelhart, and the PDP Research Group (1986), and Smolensky (1988).

⁷ I use 'individual' in a logical sense to mean, roughly, anything to which properties can be attributed. Individuals, therefore, are not limited to physical objects, but include abstract objects like functional architectures as well.

⁸ This picture (see Pylyshyn (1984) and Haugeland (1985) for details) should be taken with a grain of salt, since it is oversimplified, and perhaps false in its details. Though the highest level of functional architecture plausibly is cognitive (a level at which reference is made to entities with intentional content), it is not obvious that it must be so. Nor is it altogether obvious that there is only one level of cognitive architecture. (This is a common Connectionist response to Fodor and Pylyshyn's claim that Connectionism will at best implement the cognitive architecture, and so can be ignored for psychology. For if there is an implementation level that is cognitive too, psychology cannot ignore it.) For our purposes, however, the situation can be assumed to be more or less as I have described. Doing so will simplify the discussion, and not prejudice matters in any way.

⁹ Fodor and Pylyshyn also argue for Classicalism on the basis of their claim that thought is *productive*. But because Connectionists are free to deny thought is productive, Fodor and Pylyshyn "propose to view the status of productivity arguments for Classical architectures as moot..." (p. 36). I shall follow suit.

¹⁰ This analysis is confirmed by a discussion of Fodor and McLaughlin's (1990, sec. 1) in which they explicitly state that what must be explained are the lawful connections among capacities. However, their comments still leave open whether systematically itself requires that the capacities be related nomically, or whether mere co-occurrence suffices. This, of course, is a verbal matter, and so is of little consequence. I shall assume in what follows that the nomic component is part and parcel of systematicity.

¹¹ This is another harmless simplification. Though it is conceivable that distinct cognitive architectures might explain distinct phenomena (across modules, say), the Classicalism-Connectionism debate loses all its bite if it is not assumed there is at least one domain of phenomena over the explanation of which Classicalism and Connectionism are *in competition*. That domain should be understood to be the one under discussion in this paper.

¹² For detailed treatments of these and other features of Connectionism, see, e.g., Rumelhart and McClelland (1986), McClelland and Rumelhart (1986), and Smolensky (1988).

¹³ In Section 4 of their paper, Fodor and Pylyshyn review many of the features of Connectionist models that have been taken to favor Connectionism over Classicalism, including some of those just mentioned. Their general line of response is to insist that each such feature is compatible with Classicalism. Well that may be so. However to escape the difficulty raised above, it is not enough that Classicalism *might* do as well as, or better than, Connectionism in the relevant respects. It must be shown that it *does* as well or better (in enough cases). And that Fodor and Pylyshyn have not shown.

¹⁴ The same argument, essentially, is found in Fodor [(1987), p. 151] and Fodor and McLaughlin [(1990), pp. 201–203].

¹⁵ Assuming (A) is true, and the presence and absence of systematicity are compatible with the Connectionist framework.

¹⁶ Note that where 'architecture' is taken to refer to the defining features of an architectural framework, the argument against Connectionism applies equally well to Classicalism. For one can easily imagine a functional architecture with a Language of Thought and unconventional structure sensitive processes that can construct the representation '*Rab*' from its constituents, but not '*Rba*'.

Fodor and Pylyshyn say that in order to construct such a Classical model, one would have to "go out of one's way," though one would not in the case of Connectionist models (p. 49). Even if that is so, however, it is hard to see what that buys them. Fodor and McLaughlin (1990) go further, and deny such Classical models are even possible. They write: "... in the Classical architecture, if you meet the conditions for being able to represent *aRb*, *YOU CANNOT BUT MEET THE CONDITIONS FOR BEING ABLE TO REPRESENT bRa*; the architecture won't let you do so..." (p. 202).

That claim simply seems false, however, so long as primitive operations are included in the specification of the architecture, as they typically are.

¹⁷ A more charitable interpretation of what Fodor and Pylyshyn have in mind in the passages quoted might be this: In the Connectionist framework, or even with respect to subclasses of Connectionist models, there are models that display systematicity and those that do not. These two sorts of models, however, do not differ from each other in any deep, structural way, but only superficially – for example, in how contents are assigned to activated units. (E.g., in one case, but not in the second, the theorist ensures that the model has a representation of the form *Rab* when and only when it has one of the form *Rba*.) The point, then, on this interpretation, is not that Connectionist models cannot explain systematicity because its presence and absence are compatible with those models. It is that Connectionist models that exhibit systematicity merely by assigning contents to preserve systematic relations are *ad hoc*. And since accounts that are not *ad hoc* (like the Classical account) are preferable to those that are, the Classical account should be preferred. Now if this is what Fodor and Pylyshyn have in mind, there is no argument against Connectionism in principle in the quoted passages (appearances to the contrary); just some considerations in favor of the claim that Classicalism currently offers the best explanation of systematicity, a claim assumed true in this paper.

¹⁸ As I have said, I am hesitant to specify necessary or sufficient conditions on Connectionist architectures. Accordingly, it will suffice to establish (2), so far as I am concerned, if a *plausible candidate* for a necessary condition on Connectionist architectures is offered that is inconsistent with Classicalism. A property all currently existing Connectionist architectures share, for example, will do.

¹⁹ To emphasize the point, the kinds of details needed for the distributed version of the model are like the details we have for the localist version. For instance, suppose the distributed model were such that individual units do not have semantic interpretations; only patterns of activity across aggregates do. If

that were a property of the distributed model Fodor and Pylyshyn have in mind, it would *follow* that the model does not instantiate Classicalism, since the semantics of the distributed representations would not be functions of the semantics of their parts. And the reason would not be that the representations have no parts. They do. Rather, it would be that the parts have no semantic content. Now this property, of course, applies only to a restricted class of distributed representations, and so cannot suffice to secure (2) for Fodor and Pylyshyn. But it is the sort of property that is required.

²⁰ I intend this only as a rough characterization of what it is for representational parts in computational systems to be structured, but as one precise enough for my purposes.

²¹ The legitimacy of this assumption is not in contention here, just whether the meanings of the individual parts can be structured.

²² But see Note 8.

References

- Cummins, R. (forthcoming), 'The Role of Representation in Connectionist Explanations of Cognitive Capacities', to appear in D. Rumelhart, W. Ramsey, and S. Stich (eds.), *Philosophy and Connectionist Theory*.
- Fodor, J. A. (1987), *Psychosemantics*, Cambridge: MIT Press.
- Fodor, J. A. and McLaughlin, B. (1990), 'Connectionism and the Problem of Systematicity: Why Smolensky's Solution Doesn't Work', *Cognition* 35, 183–204.
- Fodor, J. A. and Pylyshyn, Z. W. (1988), 'Connectionism and Cognitive Architecture: A Critical Analysis', *Cognition* 28, 3–71.
- Haugeland, J. (1985), *Artificial Intelligence*, Cambridge: MIT Press.
- Lormand, E. (1990), *Classical and Connectionist Models of Cognition*, Ph.D. thesis, MIT, Department of Linguistics and Philosophy.
- McClelland, J. L., Rumelhart, D. E., and the PDP Research Group (eds.) (1986), *Parallel Distributed Processing*, Vol. 2, Cambridge: MIT Press.
- Pylyshyn, Z. W. (1984), *Computation and Cognition*, Cambridge: MIT Press.
- Rumelhart, D. E., McClelland, J. L., and the PDP Research Group (eds.) (1986), *Parallel Distributed Processing*, Vol. 1, Cambridge: MIT Press.
- Smolensky, P. (1988), 'On the Proper Treatment of Connectionism', *Behavioral and Brain Sciences* 11, 1–74.