The Architecture of Belief

John Fitzpatrick

A Thesis Submitted
for the
Degree of Doctor of Philosophy
of
The Australian National University

October 1989
The preceding chapters are designed to provide a theoretical framework within which we may go about attempting to answer the question posed in the Prologue regarding the development of the cognitive sciences. That question, remember, concerns the propositional attitudes and their role in a mature science of the mind—The-One-True-Cognitive-Psychology. Perhaps, though, that answer can be had without recourse to the fancy theoretical framework thus far introduced. As we have seen in chapter 1, for instance, Functionalism, as commonly conceived, is designed (a) to avoid species chauvinism in the specification of mental states—if the possession of some mental state required the possession of neural states then that would exclude species without neurones from having mental states—and (b) to provide some way of specifying what the relevant mental state types are—in terms of actual and potential causal roles. There is a series of arguments in the literature which purport to show that these two putative beneficial properties of Functionalism cannot be achieved. If these arguments are sound, then it follows that any theory of cognitive structure which relies upon the Functionalist programme getting off the ground, such as intentional and cognitive state realisms, will falter in the absence of another, better theory of the mind. This might strike the reader as grist for my mill, given that I don’t believe that intentional realism will turn out to be correct. However, there are two reasons why I am not enthusiastic. First, even though Functionalism as a reductive enterprise might fail, I do believe that some Functionalist theory will hopefully provide a method for the individuation of brain states, and, second, I think that these arguments against Functionalism plainly do not work. So, this chapter is about why these arguments don’t work.
The arguments against Functionalism I wish to consider take three forms. The first denies that Functionalism solves the problem of species chauvinism, since whenever it does, it lets in too many systems as systems of mentation-cognition. According to this argument, Functionalism is impaled on the horns of a dilemma: either species chauvinism or liberalism prevails, Functionalism allows too few or too many systems to count as systems with mental states. The second denies that a specification of mental states in terms of causal roles will be possible, just because the conceptual data from which the causal roles are derived—common sense or folk psychology, or even a substantive psychological theory—will fail to provide the required specification of the mental states. The final argument is similar to the second, only it is less general. It claims that Functionalism's quest for the algorithm that specifies mental state types, where that algorithm itself is specified by reference to causal roles, will not be successful. We assess each of these arguments in turn. The general strategy in confronting these arguments will be to claim that on a suitable understanding of the Functionalist enterprise the objections fail. I do not claim that this reading of the Functionalist enterprise will be immediately accepted by many Functionalists; the interpretation of Functionalism which these arguments from the literature attack does seem to be accepted by many Functionalists, though.

1 Chauvinism or Liberalism?

The chauvinism or liberalism dilemma is classically put forward by Block (1978). The example used to illustrate the liberalism of Functionalism is that of the now famous China brain (1978 pp. 279). Pretend that a billion of China's inhabitants are provided with special purpose radios that allow them to connect to each other and some artificial body resembling ours, say, but which has no brain. In addition, all the connections to the body which would normally attach to a brain are attached to transmitters which are in turn connected to a subset of the radios. When one of the body's transducers fired a message is relayed from the body to one of the radios on to various other radios, eventually to be received by a receiver in the body which initiates some motor response on the part of the body. Let's even pretend that the functional organisation of the inhabitants of China and the connections between them mimic you for a certain time, i.e. it is functionally isomorphic.

\footnote{There are many other arguments against Functionalism as an account of mental states other than propositional attitudes, in terms of absent and inverted qualia, etc. However, for present purposes I am restricting my attention to features of Functionalism relevant to its providing an account of the propositional attitudes alone.}
with respect to you. Then, if Functionalism is true, that system is describable as possessing mental states, and indeed, the belief that this is a silly thought experiment, if that's what you are now thinking! But, the argument runs, such a system surely does not possess such mental states. Therefore, Functionalism must be false.

As Block himself admits (p. 281), the claim of this argument that the China brain could not possess mental states rests on only an intuition, and an intuition which runs perilously close to being question-begging at that. We need something extra in order to secure the point against Functionalism. That point is to be had from considering the fact that our neurophysiologically based functional organisation certainly does generate mentality. It is because the China brain lacks a neurological state description, and we know that in our case such a neurological state description generates mentality, that the onus should be on Functionalists to provide independent support for their intuition that the China brain generates mentality.

Not surprisingly, perhaps, I am going to offer some independent support for the Functionalist enterprise based upon some of the considerations of the previous chapters. For now, though, we need to look at chauvinism.

Block claims that the way to avoid the problems generated by the China brain and the Bolivian economy (1978 p. 315) would be to place some constraints upon the specification of the inputs and outputs to the functionally characterised system. One could specify the inputs and outputs in terms of neural impulses, movement of limbs or stimulation of transducers. The trouble with such descriptions, claims Block (p. 316) is that they are chauvinist. Moreover, if one tries to describe inputs and outputs in species neutral terms then Block claims that will bring on liberalism since the Bolivian economy has inputs and outputs, and they might correspond to the inputs and outputs of the cognitive system. What one would do in such a case is specify the inputs, outputs and states numerically: inputs $I_1...I_n$, states $S_1...S_k$ and outputs $O_1...O_m$ related according to the Functionalist theory. There is no guarantee, though, that such a neutral description is not isomorphic to the Bolivian economy! As we shall see below, the description of inputs is crucial, so crucial, in fact, that Functionalism can be saved by their proper description (see 1.3 below).

What is going on in the Block argument? A number of crucial things are going on, the most important of which can be summed up in the following two questions: (a) in the case of liberalism, to which systems are we trying to attribute mentality? And (b) what are the criteria of attribution we are employing? We can take these issues in turn.
1.1 Cognitive Systems and Agency

The case of the China brain is interesting because there are two systems at work in the example. The first system is the artificial body which is connected to the second system, the China brain. When considering this example we must keep in mind which system it is to which we are attributing mentality. It is, of course, the conjunction of the two systems to which we must attribute mentality. Taken by themselves though, we are certainly not going to attribute mentality to them. But when we are required to make a judgement about the status of that conjunction, our intuition about that conjunction will not necessarily be the same as that for the two independent systems.

Taken by itself, the ham-radio infested Chinese populus will not be attributed with mentality, in just the same way that the Bolivian economy ought not to be. It is the intuition about the China brain in isolation, not conjoined with the artificial body, that fuels the judgement that the conjoined system is not a cognitive system. The conjunction of the two systems, though, is one to which we might plausibly attribute mentality, since the only difference between the artificial system and us is that what is doing the causal work in the stimulation of the body is not contained within that body. If it really is the roles or functions performed that is important in the generation of mentality, then we should not demur from attributing mentality in this case. To ignore this point about roles is to beg the issue against the Functionalist.

This overlooks an important point, though, when it comes to the attribution of mentality. I think there is a principle underlying those judgements, one which seems to be contravened in the China brain case. We may call this principle the principle of agency. According to this principle we attribute mentality to a system when that system’s behaviour can normally be expected to be unintentionally caused by states (either Level One or Level Two) of the system itself. A classic case of the contravention of this principle would be that of our brains being mere transmitters to a superior species, controlling our actions like puppeteers. In such a case it is obviously not states of the system that are causally efficacious in the production of the system’s behaviour. But what if we move the puppeteers to inside the system? That would seem to generate the China brain example and yet not contravene the principle. It’s here that the unintentionality of the causes comes into play. In the puppeteering case the puppeteers are not filling the unintentionally mediated roles specified by a functional description. They themselves are deciding to make the system of which they are part perform certain actions, actions they decide the system should perform.
It is important to note that in the China brain case there is a marked
difference from the puppeteering case. Although the inhabitants of China are
intentional agents, themselves possessing mental states, they are not
performing the roles allocated to them as intentional agents in the way that
the puppeteers are.\(^2\) Even if the population of China knew what their own
task was in the cognitive economy of the system of which they are a part, and
acted out of their desire to keep that larger system running, say, they are not
deciding as agents the course of the system’s behaviour. By performing the
role that they are, they are not deciding to make the entire system move its
arm or make an utterance. If they did, then they would be performing some
other functional role in that system.

The China brain case does not, in reality, contravene the principle of
agency since the system in question is not the artificial body; that system
would contravene the principle. It is always in principle possible to avoid
contravening the principle in any given case by redefining that system so as
to incorporate some extraneous elements in the etiology of behaviour. In this
way a system can be made to conform to the principle. If we initially thought
that the artificial body, or us controlled by Martians even, were the system in
question, then that system certainly does not conform to the principle, and
there is no way that we would want to attribute mentality to it or us. What
this eventuality would require is that we re-evaluate what kind of system we
are—we would not think that we were cognitive systems.

Having redefined the kind of systems we are, we then still have to
decide upon the intentional status of the components which go to make up the
system. What I want to claim is that it is the possible transgression of the
principle of agency which underlies our intuitions about the mental status of
the China brain. So, if it’s true that the Chinese populus, themselves
possessing intentional mental states, are acting not out of their own
intentions, beliefs, desires etc., but performing the dumb work neurones can
do in virtue of filling the appropriate causal roles, then we must conclude that
the principle of agency has not been transgressed. If that’s the case, then I
think we should reject Block’s intuition and claim that the China brain does
possess mental states. In this way Functionalism will not be essentially
liberal in its attribution of mentality to complex systems.

---

\(^2\)This point is similar to one made by Putnam (1975 pp. 434-39) in which he
claimed that systems which decomposed into parts which are ascribable of
mentality, should not themselves be ascribable of mentality. I think this restriction
to be ad hoc. But it is a different restriction from that implied by the principle of
agency: According to this principle a system can decompose into sentient parts,
just so long as those sentient homunculi perform non-sentient roles, as the
population of China obviously do.
1.2 Criteria for Mentality

A crucial problem for any Functionalist-based psychology is the level of generality that one wishes to capture—that is, how broad should the domain of psychology be? This is evident in the discussion of autonomy and reduction in chapter 3. It is also crucial for the considerations regarding chauvinism.

The China brain example, in attempting to prove liberalism, assumed that it is supposedly sufficient for the attribution of mental states to a system that be functionally equivalent to us. There is also a putative necessity component to the argument for chauvinism. Block’s idea is that functional equivalence to us is a necessary condition for the attribution of mentality. Block admits that maybe functional equivalence is a condition on the recognition of mentality, but fails to see how it could be a condition on mentality itself. His reason is as follows. Suppose there are Martians with whom we develop extensive cultural and commercial intercourse. We learn about their science and philosophy, read their novels and go to their movies. We then discover that their underlying psychology is functionally different from ours. “Should we” therefore, asks Block, “reject our assumption that Martians can enjoy our films, believe their own apparent scientific results, etc.” (1978 p. 311). If we don’t reject our assumption then it would appear that a functional organisation of a certain type cannot be required for the attribution of mentality.

Lurking behind this example is the belief that there is some condition other than functional isomorphism with respect to us which must be met in order for a system to be described as a cognitive system. If there were no such alternative condition then Block would not be able to claim that “it would be perfectly clear that even if Martians behave differently from us on subtle psychological experiments, they nonetheless think, desire, enjoy, etc.. To suppose otherwise would be crude human chauvinism” (p. 311). The criterion lurking here is one which allows both Martians and us to be described as possessing of mentality. If some Level Two functional organisation is too specific a description, then it must be a higher level Level Two description or even a Level One description which is criterial of mentality.

Now Block gives us no idea what he takes the requisite criterion to be. Nevertheless, the Functionalist might well decide that it is that description, at whatever level, which is criterial of something’s being a system possessing mentality or not. The various functional organisations such as those of us and the Martians constitute realisations—or perhaps Pylyshynian functional architecture—of some more abstract descriptions, the possessors of which are attributed with mentality.
The point I am making here is the same as that made in chapter 4 regarding the level of description at which we decide when something counts as a cognitive representational system. I think it's an important point that Block fails to treat with enough respect. He thinks that specifying the functional architecture of a system at Level Two is criterial of our judging that something is a system which possesses mental states. He then shows how that criterion is inadequate by claiming that it is chauvinist. He can only do that, though, by employing a higher level criterion, Level One, say, which captures the relevant class of cognitively described entities in its net. I want to know what that higher level criterion is, and why the Functionalist cannot employ it in her programme.

Block recognises that one might be tempted to make the move that I prefer, that, maybe, Functionalism is a Level One enterprise, but claims that a simple example counts against the move. He then goes on to run the standard argument against the Turing Test. The machine imitating a human interlocutor seemingly possesses human conversational abilities, but works according to list-search principles. He claims that because the machine works according to these principles and it seemingly possesses the same linguistic inputs-outputs as us, we must claim that it has no mental states. We have already encountered this type of objection in the previous chapter and discarded it. It is not at all clear that the range of inputs and outputs, and the relations between them, are of a sort that is evident in a cognitive system to which we want to attribute mentality.

There is another difficulty with Block's attack on Functionalism which I wish to mention in closing this section. Block seems to assume in his attack based upon necessity conditions for mentality that the concept of mentality is robust enough for us to get criteria for it which will aid us in deciding whether Martians are attributable of that concept or not. However, maybe mentality is not such a concept. It may be the case that mentality is a highly graded and pragmatic concept whose conditions of application are vague and imprecise. It might be the case that we can decide that Martians possess mentality only to greater or lesser degree, when compared to us.

The situation confronting Block can be seen in the case of infra-verbal mentation. If the Functionalist paradigm is supposed to give us criteria of mentality then it should provide us with a means of deciding whether certain non-human animal species possess mentality, species with which we have some phylogenetic commonality. We don't, however, have any Functionalist-inspired way of doing this. We have yet to make the judgement as to whether or not cockroaches have beliefs and desires. There are bound to be functional similarities, to a degree, between cockroaches and ourselves—we both have
perceptual and motor control mechanisms, for instance—but how much similarity is required in order to claim that they have mental states, that infraverbal mentation is not a self-refuting concept?

1.3 Inputs and Outputs

We saw above that the description of the inputs and outputs of a functionally described system is crucial to the problems confronting Functionalism. Well, what descriptions of inputs and outputs must the Functionalist employ? There is a possibility of being species chauvinist in the specification of the Level One description which we take to be characteristic of mentality. Does a system have to possess linguistic capacities, mobility, reproductive capacities, etc.? As I have claimed in chapter 4, the properties of a complex system we take to be the mark of the cognitive are rather more abstract than a certain range of actual behaviours. They rather have to do with the different ways of responding to various ranges of stimuli. The trouble with the abstraction properties of relations between inputs and outputs is that those very properties which I have taken to be a mark of the mental might well be too liberal as well. One might think that the transitions of inputs and outputs evident in cognitive systems which are S-R abstract, say, can be exhibited by the Australian economy. This seems a fair bet since a Japanese import might well be related to a variety of outputs from the country: an export or an international bill of exchange.

Block and Owens (1983) rightly point out that this is a major problem for Functionalist-based accounts of the mental. How can its impact be reduced? One way would be to invoke the principle of agency. The intentional agents which of necessity make up an economy act out of their intentional states; somebody decides upon receiving a Japanese import that an export or bill of exchange gets output.

A better way would be to specify the inputs and outputs of a system that avoids both liberalism and chauvinism. Consider the case of rocks. Why don't we attribute mentality to rocks? Quite often, rocks are deemed not to have mental states just because they don't behave—see Fodor (1987 p. 69). Complex systems such as rocks do, however, have outputs: erosion and heat radiation, for instance. The point is: why don't those outputs count as behaviour? Consider some not very complex organism, say a paramecium. Running the same kind of Block and Owens liberalism line should get the critic of Functionalism to say that the paramecium has some description in terms of inputs and outputs that makes it functionally isomorphic to us. However, that line is never run. Why? Because we know that the paramecium
is an organism responding to an environment in certain ways, ways that don’t allow us to attribute it with mentality—due, I claim, to its not possessing the collection of abstraction properties to the relevant degree. In describing the paramecium in these terms, we have fixed a certain level of abstraction at which to describe its transitions from input to output. It is \textit{that level} at which we judge that the paramecium is \textit{not} functionally isomorphic to us. There may be some other level of description of the inputs and outputs of the paramecium in terms of its absorption of sunlight and chemicals, and its outputting of waste and oxygen, such that the processes of photosynthesis are isomorphic to the processes of cognition. \textit{But that is not the level at which we make psychological judgements about paramecia.} If it were, then they too would count as cognitive systems.

It is at some level similar to that at which we describe the inputs and outputs of the process of photosynthesis at which we make the judgement about the transitions from input to output of rocks eroding. However, even if there is an isomorphism between the story we tell about the rocks, we are not going to judge that the rocks have mental states just because we realise that the level of description of those state transitions are not at the level of abstraction at which we make judgements about us or paramecia. It is for such reasons that we do not claim that rocks and paramecia are functionally isomorphic. Similarly, it is the reason why we demur from attributing mentality to economic systems or the Milky Way. Perhaps there is a gas cloud of galactic size which moved with such slowness that its time scale would be extremely slow by our standards (Putnam 1987 p. 88). We can grant that such a system might be a cognitive system not because we describe its inputs and outputs in astronomical terms, but in terms relevant to psychological theorising, as when we compare ourselves to paramecia and rocks.

What is this so-called level of abstract description at which we describe the inputs and outputs of a cognitive system? That’s the really hard question in the present discussion. Whatever it is, it’s the difference between describing the paramecium as functionally isomorphic to us, and describing it as functionally distinct from us. One might claim that it is the description of the inputs and outputs that are psychologically relevant: inputs that count as \textit{stimuli} and outputs that count as \textit{behaviour}. In short, the relevant level is the one at which folk psychology applies. When we are willing to describe the inputs as \textit{perceptions}, then we have arrived at the correct level. In explaining the outputs of rocks and paramecia we have no need to appeal to beliefs and desires, in order to frame the explanations or predictions of the outputs. We would describe the inputs to a rock as perception and its outputs as behaviour only if we were forced to attribute folk intentional states to the system. But in
Chapter 7

What's Wrong with Functionalism? 153

such cases we, obviously, do not have to: chemistry and geology will provide
the level of description at which we can state these explanations. Describing
the inputs in this way will not commit the Functionalist to any form of
chauvinism, since the general class of perceptions does not have to include
the visual or auditory perceptions of our species.

I admit that this is not much of an account of how inputs and outputs
should be specified. As I said, this is the really hard question, and at the
moment I don't have a worked out answer to offer. However, I think what I've
said is enough to placate Block and Owens.

2 Schiffer

Perhaps the easiest way to generate an argument against Functionalism is to
stipulate that it meet certain *prima facie* plausible, but in effect,
unreasonable desiderata, and then decry it for failing to meet those
desiderata. In essence this is what Block has done in his argument that
Functionalism is chauvinistic. This style of “argument” is also employed
against Functionalism by Stephen Schiffer (1987). Schiffer also claims that
Functionalism (again whether one is dealing with common sense or scientific
Functionalism) is required to postulate some necessary conditions in order to
generate the specifications of the mental states quantified over by the
Functionalist theory. Such necessary conditions will be, claims Schiffer,
perceptual input conditions and behavioural output conditions. An example of
an uncompleted perceptual input condition might be Schiffer's own example
already mentioned in chapter 4:

[P] If there is a red block directly in front of $x$ and ..., then $x$ will
believe that there is a red block in front of $x$.

It was argued in chapter 4 that it was the mark of a cognitive system that
there were indefinitely many ways in which that system might come to be in
a cognitive state. If that's so, then we should expect that no such perceptual
input conditions are going to be forthcoming, if the system in question is truly
a cognitive system. To insist that Functionalism must be able to come up with
such conditions is to insist upon the impossible. Now it might well be the case
that Functionalists *thought* that they might be able to come up with
conditions, and if it is this (what I claim to be a mistaken) belief which
Schiffer is calling into question then he is correct. The question is, though:
does the Functionalist have to come up with those strict conditions? If
Functionalism *does* have to provide such conditions then the considerations of chapter 4 suggest that the Functionalist’s task is an impossible one.

The demand that Functionalism must provide Schiffrerian perceptual input conditions is the demand that there be *definitions* of the mental states a Functionalist theory of the mind quantifies over—for our purposes, the propositional attitudes. This can be seen in Schiffer’s attack upon what he calls *commonsense Functionalism*. Commonsense Functionalism is the view that our propositional attitude concepts are defined by reference to the common knowledge (either explicit or implicit) of the agents that possess them. In other words, it is commonsense platitudes regarding propositional attitudes which give propositional attitude concepts their meaning. Schiffer complains about this conception on two fronts. The first has to do with who has the access to these platitudes. He says: “If the meaning of ‘believes’ is determined by a folk psychology expressed by its use, then that theory must be one implicitly held by everyone who has the concept of belief” (1987 p. 31). But, claims Schiffer, it is clear that those who possess the concept of belief have no idea about defining that concept, however implicit the theory might be. This is especially evident, he claims, in the case of machines, extraterrestrials, and even Helen Keller and Ray Charles. How could they possibly have any idea of the *perceptual* input conditions defining belief?

The second front has to do with the likelihood of coming up with conditions which will be strong enough to define mental state concepts. Schiffer claims that even if there were some corpus of knowledge possessed by all those with the concept of belief, that knowledge would not be of a kind to generate definitions.

On both fronts, Schiffer imposes standards that Functionalism ought not to have to meet. As to the first front, take the case of ordinary grammatical competence. Many speakers of a language have no idea of the formal arrangements of the language even though the grammars descriptive grammarians generate are determined by the use of the speakers of that language. We don’t say that there is no theory of the language just because Bruce Layman cannot articulate such a theory. The same goes for Functionalist definitions. The response to Schiffer’s second front of objections is to deny that Functionalists must, in offering their theory, provide “definitions” in terms of sets of necessary and sufficient conditions. We saw in chapters 4 and 6 that in the cognitive system case to which Functionalist theories are going to apply, there are not going to be any such necessary and sufficient conditions because of the abstraction properties. So, whatever account the Functionalist is going to give, it won’t be in terms of the definitions alluded to by Schiffer. It’s impossible; we cannot get them. Nor
should we even contemplate for a moment that we could get them. Remember that the causal roles alluded to by the Functionalist are not only actual roles but potential roles as well. The trouble with specifying counterfactuals in the roles determinant of mental states is that they give us another reason to deny that there will be a necessary and sufficient set of roles which are constitutive of mental statehood, since there are indefinitely many counterfactual roles that could so feature.

An alternative strategy for dealing with Schiffer's objections is to re-examine the way in which Functionalism hopes to individuate mental states. In the discussions of Functionalism so far, and in the discussion of Putnam in the next section, it is assumed that the Functionalist wants to give a very fine grained taxonomy of mental states where the belief that $p$ is differentiated from the belief that $q$. It might be the case, though, as we saw in chapter 1, that the Functionalist wants to taxonomise mental states more coarsely only, so that we individuate believing that $p$ as opposed to the desire that $p$. In effect the Functionalist would be giving identity conditions (in the loose sense of 'condition' I have just been advocating) for the so-called "intentional boxes". We can think of these intentional boxes as boxes in a flow chart representing what Pylyshyn (1984) has called the functional architecture of the cognitive system. If that's right, then it will avoid Schifferian style of objections. There simply will not be a Functionalist individuation of the belief that there is a red block in front of me. Consequently, there will be no need for a perceptual input condition of the kind required by Schiffer.

As mentioned in chapter 1, Fodor recognises that there are potential problems for Functionalism's going fine-grained, and steers clear of a primarily Functionalist-based semantics for those reasons (although he thinks that functional role might have some minor part to play in the determination of content). Of course, having some way of fixing the contents of the intentional boxes is crucial. For that one is going to need some sort of semantic theory which will secure the intentional status of the boxes' contents. I am not going to have much to say regarding what is the right semantic theory; perhaps a causal theory similar to Fodor (1987) or Dretske (1981) or Millikan (1989) will suffice. In this work I am not crucially concerned with that enterprise, for reasons cited in chapter 6. However, I will have a little something to say about semantics and content, as promised, in the next chapter.
3 Putnam and Multiple Realisability

The most recent attack upon Functionalism can be found in Putnam's recent *Representation and Reality* (1987). Putnam has finally betrayed the doctrine he helped spawn; he has aborted his own conceptual child. Putnam offers three lines of argument against any Functionalist programme. The first is an argument from meaning holism. We considered that argument in chapter 1. The second is what we may call the argument from broad content. That will be discussed in the next chapter. The last argument, which is the subject of this section, we may call the argument from the multiple realisability of functional states. Whether one believed in some form of "Turing machine" formalism of the computational states quantified over by Functionalist theory, or relied upon a David Lewis (1972) style (or even Putnam style—as of "Philosophy and Our Mental Life" (1975f)) of formalisation in terms of an implicitly held theory such as folk psychology (or in terms of a substantive psychological theory in Putnam's case), it is required by Functionalist theory that each mental state reduce to a computational state. Putnam claims, in a similar vein to Block, that it is false that there is one computational state shared by all physically possible systems to which we want to attribute the same collection of mental states.

3.1 The One True Algorithm and Levels

Putnam's idea is this. Some Functionalists (although not Fodor and Pylyshyn) think that a functional description of a system will generate a taxonomy of mental kinds fine-grained enough to differentiate those states according to their content. If system A is in some computational state $x$, and another system is in a computational state $y$, those systems are in the same mental state provided there exists some synonymy relation between $x$ and $y$ such that both $x$ and $y$ mean $p$. If it's functional role that determines the meaning of mental states, then any two agents which differ in their belief set will turn out to mean different things by any state defined by the overall computational-functional model (1987 pp. 85-87). Or consider an agent within some linguistic-cultural context. Putnam claims that the "functional organisation" of any two individuals may not be exactly the same (p. 82). Suppose that there is "belief fixation" component to an inductive logic which is hardwired into us. It might be part of our functional architecture. Inductive logics can differ in their assignment of probabilities, and so belief fixation can vary across individuals. The upshot is that when two agents are deemed to believe
that there are kangaroos in the neighbourhood, there will not be anything
functional-cum-computational-cum-physical in common (pp. 81-84).

There is an important addendum to Putnam’s argument. As he argues
in an Appendix, it turns out that there are, in effect, too many realisations of
any system, the workings of which we specify by some computational formal-
ism—machine tables or (folk) psychological theories (pp. 121-25). If this is
true, then it follows that

the assumption that something is a “realisation” of a given automaton
description (possesses a specified “functional organisation”) is
equivalent to the statement that it behaves as if it had that description.
In short, “Functionalism”, if it were correct, would imply
behaviourism! If it is true that to possess given mental states is simply
to possess a certain “functional organisation”, then it is also true that
to possess given mental states is simply to possess certain behaviour
dispositions! (p. 124)

This is an important result for Functionalist theory. Where as Putnam takes
it to be a reductio, I will argue below that the conclusion should be embraced.
For now, though, back to the main argument.

Putnam goes on a great deal about “interpretation”. Functionalism is
described, in Putnam’s words, as the Master Algorithm for Interpretation
(p. 91). In this role attributed to it, Functionalism is a panacea for all one’s
psychological and even semantic ailments (p. 92), designed to provide not only
an account of propositional attitude types (such as believing and desiring) but
also a general account of meaning and reference! While I think that this is
imposing too much theoretical work onto Functionalism, we can grant, for the
sake of argument, the Functionalism-as-interpretation metaphor.3 Even
though Putnam does not believe that there is such a master algorithm for
interpretation, he is certainly no eliminativist with respect to propositional
attitudes either (see 1987, chapter 4). He thinks that we do make
propositional attitude ascriptions, and we make those ascriptions in the
course of some form of interpretative practice.

3Having said that, I think contemporary philosophy of language does depend upon
the Functionalist programme bearing a fair degree of theoretical weight. If the
meaning of our words derive their meaning from the semantic properties of our
psychological states, then the psychological story takes on a responsibility reaching
further than mere psychological interests. However, I’m not sure that any
Functionalist has thought that there was going to be a Functionalist theory of
reference.
With this interpretative practice in place, Putnam is then able to claim the multiple realisability of computational states postulated by Functionalism theory:

...we are not going to find any physical state ... that all physically possible believers have to be in to have a given belief, or whatever. But now it emerges that the same thing is true of computational states. ... Physically possible sentient beings just come in too many “designs”, physically and computationally speaking, for anything like “one computational state per propositional attitude” Functionalism to be true. (p. 84)

We have seen in Part I that multiple realisation is a relation which holds only between levels—between a lower level and a higher level to be exact. Since functional organisations are multiply realisable with respect to Putnam’s interpretative practice, that interpretative enterprise constitutes a level of explanation-description. This point is crucial for two reasons. Firstly, we need to know at what level of analysis this putative interpretative analysis is supposed to be; and secondly, if some level of explanation is being employed in order to decide when to attribute propositional attitudes, one should be explicit about it, because, as we saw in examining Block’s argument, maybe it is a level that can be employed by the Functionalist so as to avoid the current objection. We examine these points in turn.

I can only presume that the explanandum of the interpretative practice of Putnam’s is the action of agents. In other words, the systems being interpreted are human agents, or intentional systems, as Putnam himself calls them—no doubt following the lead of Dennett (1987 & 1979). These individual intentional systems might form, through their interactions, some larger system—a society or culture—which might even include the natural kinds of the agents’ environment. However, it is the individual agent within that larger system which is the object of study under Putnamian interpretation. We might well have to locate that system in its environmental and cultural context, but it is that individual to which we are going to attribute mental states such as propositional attitudes.

Now the computational states—either boxes featuring in the description of the functional architecture or a fine-grained taxonomy of mental states which distinguished various beliefs, say, from each other—which the Functionalist wanted to attribute to such a system in order to explain the presence of propositional attitudes were presumably the highest level Level Two states of the system. If the system were a black box, the functional organisation would be postulated to explain the capacities of
the system. But as Putnam has just argued, that kind of state attribution cannot account for the propositional attitudes. Those states capturable by some Level Two flow chart representing the functional architecture are multiply realisable. The only alternative left, therefore, is that Putnam's interpretation must be a Level One analysis.

This result is significant, I think for two reasons. The first is that if the arguments adduced by Putnam are sound, and to possess propositional attitudes is to be capable of being attributed with states under a Level One analysis, then those arguments support the conclusions of chapter 4, in that the criteria of the cognitive, where we attribute cognitive states of which propositional attitudes are a species, are had from Level One. The second reason why the result is significant is that the result Putnam proves in the Appendix is just the result that it is the Level One properties of a system that count when one is considering whether the system realises some functional model. Putnam thought the results of his Appendix constitute a reductio of the Functionalist position. I think what that result shows is that many functional specifications are really Level One analyses of complex systems, when Functionalists mistakenly thought that they were doing Level Two analyses. That's a Level confusion if ever there was one.

Putnam assumes that his result could not be embraced by a Functionalist because it would make such functional descriptions behaviourist. However, behaviourist analyses of complex systems really are just a species of Level One analysis. Not all Level One analyses need be riddled with the problems associated with behaviourism. What Putnam's result shows is that Functionalist analysis is a species of Level One analysis, not that it is Behaviourist analysis. Describing the kind of analysis the Functionalist should be pursuing as behaviourist is mere mud slinging, trying to prove theoretical guilt by association only. Why is Level One analysis not mere Behaviourist analysis? For a start, Logical or Analytical Behaviourism wanted to define mental state concepts. I've argued in the previous section that a Level One Functionalist analysis of mental states will not be seeking definitions in terms of the inputs and outputs of cognitive systems. These forms of Behaviourism also wanted to define the mental state concepts in terms of behavioural dispositions. In so doing behaviourists wanted to eschew any reference to states of the system in their analyses. So, they would, for instance, analyse attributions of the form 'A believes that it is raining outside' as equivalent to 'If A were to go outside then she would carry an umbrella'. Level One analyses do not demand that states of a system be analyses away in this way. As we saw in chapter 2, one can advert to states of
a system in Level One analyses provided that they are individuated at Level One.

We should now look at Putnam's argument applied to David Lewis' version of Functionalism. Remember that on Lewis' account, the mental states get specified by reference to causal roles between sensory stimuli, motor responses and mental states, where the causal roles can be gleaned from the platitudes of folk psychology. Even though mental states are included in this specification, and they can even be “internal states”, it does not follow that the specification is not made under a Level One analysis. As claimed in chapter 2, a system can go through state transitions in the production of output; but these are Level One state transitions. The Lewis story should, I think be interpreted as a Level One description of a system. If a Lewis style functional specification gets realised just when its predictions about the system's behaviour come out to be true, as Putnam claims (1987 p. 96), which would be the case if Lewis' Functionalism were pitched from Level One, then that is the notion of realisation with which the Functionalist is going to have to live (despite Lewis' supposed reluctance).

Stephen Schiffer interprets, correctly I think, Functionalism in the way I have been advocating. He could not be clearer as to his interpretation:

We might have a black-box problem: we are given an input/output system (the black box) whose outputs are a function of its inputs and it's internal, physical states; we have access to the inputs and outputs but know nothing about the nature of the internal states or of the causal laws governing them. Nevertheless, we seek a theory that will be explanatory and predictive of the outputs. To provide such a theory is to solve the black-box problem.

We might be able to solve the problem by devising a correct functional theory of the system: we might theorise that there are so many internal state-types the system might be in, which are related to one another, to inputs, and to outputs in such-and-such causal or transitional ways. If this theory is correct and detailed enough, it could enable us to predict the system's outputs on the basis of its inputs, just as knowledge of a computer program may provide us with the ability to predict its outputs, even though we know next to nothing about its internal hardware. (1987 p. 24)

---

4I think also that the minimal Functionalist account of the propositional attitudes employed by Jackson and Pettit (Forthcoming) reads Lewis and Functionalism in the way I am suggesting here. According to them folk psychology is a higher level of explanation than The-One-True-Cognitive-Psychology or the neurosciences since the various ways these lower level enterprises might turn out will be consistent with the Functionalist account of beliefs and desires.
According to this interpretation, a functional theory employed by a Functionalist will be unashamedly Level One. Given this interpretation, it is not surprising that Schiffer does not employ the multiple realisability of algorithms argument used by Putnam against Functionalism.

That just about concludes my criticisms of the Putnamian objection to Functionalism. Before moving on to some of his minor objections, I want to look at what I think are some of the implications for the reading of Functionalism I am adopting. If Functionalism with regard to propositional attitudes should be interpreted as a Level One enterprise in order to avoid the multiple realisability and chauvinism arguments, then it will follow that The-One-True-Cognitive-Psychology, which is a self professed Level Two enterprise, is going to be chauvinistic. This is, I think, an advantage, since it comports well with the considerations of chapter 3. We saw there that maybe domain-specific reductions might take place between various (Level Two) levels of explanation-description. If The-One-True-Cognitive-Psychology has already contravened the Functionalist's ideal of nonchauvinistic multiple realisability, then a chief obstacle in the path of reduction is blocked, and the way is opened for a domain-specific reduction. Also, we saw that The-One-True-Cognitive-Psychology should not be thought of as either developmentally or confirmationally autonomous from lower enterprises such as the neurosciences. Since cognitive psychology is already chauvinist then allowing the neurosciences to play some developmental or methodological role will not detract from the multiple realisability desideratum of the Functionalist account of mental states such as the propositional attitudes.

Having said all of this, though, there still might be the theoretical possibility that the Level One Functionalist story I have been advocating is chauvinistic. This, again, is a point about the robustness of concepts such as MENTALITY. Block and Putnam want to attribute mentality to a system which could be said to possess mental states such as beliefs and desires. Now it might well be that systems to which we attribute beliefs and desires are only a subclass of all the possible “intelligent” creatures to which we might want to attribute mentality, or, at least, “intelligence”. What goes for Level Two states, in their being only one possible realisation of a Level One belief-desire model, might well happen to the Level One belief-desire model. Maybe there is some higher level than my proffered Level One analysis at which we attribute concepts such as mentality. It might well be the case that we would want to attribute mentality or intelligence to creatures to which we would not normally attribute propositional attitudes.
What this chapter shows, I think, is that traditional conceptions of the Functionalist programmes are wrong. If cognitive state realism and intentional realism require that mental states are conceptually analysed at Level Two, then I think the objections presented in this chapter have some force. But what I hope to have shown is that the functional description of mental states can be interpreted as being at Level One. It is then an empirical matter as to whether or not there is an isomorphism between the states at Level One which are definitive of mental states, and the states at Level Two over which The-One-True-Cognitive-Psychology quantifies.

That leaves us in the position of having to decide whether a Level Two analysis of our cognitive systems are going to turn out along the lines envisaged by cognitive state realism. That thesis, now, amounts to the view that there is such an isomorphism between Levels One and Two. Functionalism might be a Level One analysis of propositional attitudes, but nevertheless, there will have to be a functional Level Two analysis of cognitive systems in order to explain the capacities of those systems. The question, therefore, is whether that functional decomposition of a cognitive system matches up with the Functionalist Level One analysis usually employed by philosophers. Our task, then, is to take a look at some candidate Level Two decompositions. That's the task in chapter 9.

Where we are in Part III is this. We have just employed some of the results of the previous sections in order to stave off some quick ways of cutting the theoretical ground from under the cognitive state realist—we now know that it doesn't depend upon Functionalism as popularly construed. We have also seen in this chapter that the question of the so-called semantic properties of mental-cognitive states is a hot one, and many of the problems for intentional realism seem to stem from this very property. So before proceeding to chapter 9 and our look at Level Two cognitive architectures, we must first address the topic of intentional semantics.
Chapter 8

Broad Content

One traditional conceptual argument against intentional realism has its roots in the philosophy of language. Putnam (1975a), in his now famous "The Meaning of 'Meaning'", argued for the slogan that "Meanings ain't in the head". What this amounts to is the claim that what our words mean depends upon the way the world is rather than just, say, the descriptions associated with those words in the minds of the speakers of those words. Two speakers could be in the exact same collection of physical states (or brain states, even), and yet refer to different things. Since terms with the same meaning must refer to the same things, the meaning of those words of the two psychologically identical speakers must have different meanings.¹

How does this relate to intentional realism? On the intentional realist story, intentional states are representational in just the same way as linguistic items. Intentional states have content, and it's in virtue of how they connect up to the world that they have the content that they do. In short, beliefs and desires have a "meaning" in just the same way as linguistic items. However, the intentional realist posits that intentional states are states of one's brain. If those states are in one's head, as it were, and meanings are not in the head, as it were, then how can intentional states be in one's head given that they have meaning? That, in a nutshell, is the challenge to intentional realism in terms of the semantic properties of the propositional attitudes.

¹Putnam (1975a) uses his now famous "twin earth" experiment in arguing for this position. The idea is that there is a molecule for molecule replica of earth somewhere except in one respect: where water (H₂O) is present on earth, some other substance (XYZ) is present on twin earth, called 'water' by its inhabitants. This means that two molecule for molecule (except for H₂O and XYZ) replicas of Putnam could be in exactly the same brain states, and use the exact same set of descriptions (the story is set before the advent of an advanced chemistry) for picking the clear liquid filling their oceans while Putnam on earth refers to water whereas his Doppelgänger refers to XYZ. Since words with the same meaning must refer to the same thing, Putnam concludes that the meaning of 'water' on earth and twin earth must be different.
The efficacy and influence of this challenge has been entrenched by, most notably, Burge (1979 & 1986). As well as explicitly extending the Putnam line to intentional states, Burge also argues that social facts such as one's community's linguistic practices in part determine the content of one's intentional states.

We may call the Putnam and Burge type argument the argument from broad content. Why 'broad'? Because linguistic items and intentional states refer, mean, or represent in terms of their relations to the world "outside the head or skin" of the speaker or agent. Things going on within the speaker or agent are, contrastingly described, “narrow”. More precisely, there are really two types of broadness, what we may call synchronic broadness and diachronic broadness. The classic cases of broadness are synchronic. In these cases, some representational item only counts as a representation in virtue of its synchronic relation to the environment. An example of this might be an organism's perceiving a red apple. One can also be related to the environment diachronically. One's belief that there is a beer in the fridge, although its truth value depends upon how the world is now, is a belief about that beer because its diachronic connection to that beer, just like my belief that there is water in the lake is diachronically connected to water rather than Putnam's twin water. Quite often diachronic broadness consists in a causal history of some meaning or representational item, and an example to follow in latter sections of this chapter will employ this form of broadness.

There are basically two responses to the argument from broad content. The first is to reject the intentional realist enterprise on the grounds that its commitment to mind-brain supervenience fails. This is the tactic adopted by Burge, the later Putnam, and Pettit, among others. The second response comes from the intentional realists, most notably Fodor (1987). He accepts the broad challenge but answers that broad content is of no use to scientific psychology, and advocates the adoption of what has become known as "narrow content". Roughly, the narrow content programme urges that the contribution of the context determining environment upon mental states should be ignored for the purposes of scientific psychology.

In this chapter I want to argue that even to the extent that the argument from broad content correctly hits a target, the hitting of that target in no way counts against the intentional realist programme. To that extent, I don't think that the intentional realist needs to opt for the narrow content programme. For that reason, an examination of the narrow content programme is just redundant from the point of view of this work, and so I don't care whether that programme can be coherently spelled out. Basically, I am going to argue that the argument from broad content and its slogan
“beliefs aren’t in the head” does not count against one of the relata of the meaning or representational relation being “narrow”. That is, although the mental representations which bear the meaning relation possess that meaning according to some broad criterion, it does not follow that that bearer of the meaning is broad. Or, again, in terms of intentional states, contentful beliefs might be the beliefs they are just because they feature as relata in a relation to a broad environment, but it does not follow from that, that as relata they cannot be “in one’s head”. Spelling out this idea is the content of section 1. In section 2 I distinguish between two roles that broadness might play when it comes to taxonomising representational or meaning bearing entities. I argue that one of them is assumed by the argument from broadness, but it must be the other role that is required in the current debate. Finally, in section 3, I offer an analysis of the putative puzzle about how broadly individuated states can seem to supervene upon narrow states. It is here that the issue of the causal relevance of broadly individuated states is raised. I argue that the considerations of section 1 will provide a means for dealing with this issue.

I hope it will then be all too clear why the narrow content programme is unnecessary, and why the slogan “beliefs ain’t in the head” is an obscuring red herring.

1 Can “Beliefs Be In The Head”?

Words have semantic properties. In that way they are substitutable for ‘x’ in the relations: x means y, x refers to y and x represents y. When Putnam tells us that “meanings ain’t in the head” what he must mean is that any theory of such relations, as opposed to either of the relata, cannot be given solely in terms of the properties of one of the relata. That would seem to be impossible, since it would seem to misidentify what is to be explained; it gets the explanandum of our theory wrong. Semantic properties such as meaning, reference, and representation are paradigmatic relational properties, and relational properties are specified in terms of the relation and its relata (how many relata being dependant upon the adimity of the relation). If this is so obvious, why does the argument from broad content strike so many as a surprising and new point in the philosophies of language and psychology? It should be just obvious that it could not be solely the intrinsic properties of one of the relata of an intentional relation that could provide an account of its relational properties. Putnam says:
reference is socially fixed and not determined by conditions or objects in individual brains/minds. Looking inside the brain for the reference of our words is, at least in the cases of the kind we have been discussing, just looking in the wrong place. (1988 p. 25)

This seems obviously right; but I don’t think anyone has ever claimed that one could look inside the brain for the reference of our words, precisely because reference is a relation.

1.1 Relational Properties

What's gone wrong here? Perhaps people have mistakenly thought that the properties which are in fact relational are not relational. It's hard to imagine that one could have thought that given the seemingly obvious relationality of semantic properties. It does seem obvious that a word’s meaning, for instance, does not depend upon the intrinsic properties of the word: different words, after all, can have the same meaning. However, it might have been thought that if representational mental states are identical with or supervene upon the states of one’s nervous system, then those mental states might have been thought to be intrinsic nonrelational properties, since the states of my nervous system are presumably intrinsic states (whether or not this is so is the content of section 3).

Perhaps, though, this is to mistake the object of the surprise. Even though an account of the representation relation is undoubtedly broad, the slogan “meaning ain’t in the head” might be interpreted as the claim that the bearer of the meaning, the ‘x’ in the relation, cannot be narrow. This is most clearly seen in the intentional state case, the slogan being “beliefs ain’t in the head”. To ensure the isomorphism with the meaning case, we must read the slogan as claiming that beliefs have the content they do because of some of their relational properties; the slogan tells us that some mental state x represents some state of affairs y. That seems fairly uncontroversial.

However, there seems to be an alternative reading of the slogan in which it might be thought that the broadness of a belief’s representational status implies that that belief qua relatum of the representation relation—the bearer of the meaning or representational property—must also be broad, and hence “not in the head”. Now this is a surprising and interesting claim, and one that I think is assumed to follow from the former claim about the broadness of the representation relation. If you think that the former claim implies the latter, then there really is going to be something for the intentional realist to worry about, since not only must she give a broad
account of the semantic properties of intentional states, the bearers, *qua bearer*, of those semantic properties themselves cannot supervene on, or be identical with states of one's brain.

Well, does the implication hold? I am going to argue that it does not, and, consequently, the intentional realist should not be overly worried. However, one might object here that the implication must hold, since if the states are individuated broadly according to their relational properties, then the bearers of those relational properties must also be relational entities. This is a move made by Garfield (1988 ch.5), for instance. Because, for example, linguistic tokens have their meaning essentially, in terms of their relational properties, those tokens are "things in a conventionally constituted context" (Garfield 1988 p. 100). The thought seems to be that the implication does hold because there is only a bearer of the meaning because of its relationality. This seems to be the claim that the theory of meaning underlying the argument from broad content also determines the ontology of the linguistic or the representational domains.

I think that in one sense this is right, and yet in another sense it is wrong. There would seem no doubt that it is relational or broad properties which determine the ontology of the *representational* or the linguistic. However, to the extent that the tokens we describe as linguistic, and hence having meaning, or the mental representations which we describe as representational are individuated linguistically or representationally in terms of their relational properties, the bearers of those intentional properties might be ontologically individuated by reference to some of their other properties. How is this possible? The considerations which follow suggest the answer.

1.2 What Can One Dollar Buy The Intentional Realist?

We can learn something about the intentional case by looking at an uncontroversial non-intentional relational case. Consider the case of dollar coins. Suppose that I am visited by A who comes from a Central American tribe where there is no monetary exchange or the like. She sees me put a dollar coin into my pocket and asks: "What is that in your pocket?" I of course answer that it is a dollar coin. "Well, what is a dollar coin?" A persists. A couple of answers are going to be possible here. One answer will describe the physical properties of the dollar coins: shape and markings, chemical analysis, weight *etc.*. We may call this the "narrow" explanation. But something can possess all those narrow properties and *not be a dollar coin*: replicas for display or forgeries, for instance. An alternative answer would be
to list the properties of the narrow explanation (perhaps not in as much
detail, but at least enough to differentiate a dollar coin from other coins and
random pieces of metal) and, in addition, a history of the proper production
(by mints) of units of monetary exchange, combined with some historico-
economic story regarding the practices of economic exchange. We may call
this the "broad" explanation, where the additional features contained therein
constitute the context. It seems that the broad answer to A's question is the
better of the two, since it will enable A to use a dollar coin within the
monetary system, and when she comes across the special edition dollar coin
with a different set of intrinsic properties, or indeed some other unit of
monetary exchange, she will also have an explanation of why that object is
also a dollar coin.

It's easy to see that there will be no strict narrow account of coins; they
differ vastly in size, shape and constitution. In fact, it might be that it is
not even coins, or notes that function as the appropriate units of monetary
exchange. Perhaps locks of hair could have been used.

Given that our answers to A's questions are appropriate, the kind of
question A was asking was a conceptual one. It tells us what makes
something a dollar coin, or, more generally, a unit of currency. In these cases
it is certain relational properties of that something which make it the
something that it is. The broad explanation is intended to give some idea of
the criteria of application of the concept DOLLAR COIN or UNIT OF
CURRENCY to some entity. Now each of the entities to which we apply these
relationally construed concepts seems to possess a set of narrow or intrinsic
properties—it's that entity determined by narrow properties which we
describe as a dollar coin. It is because of these properties that A was able to
ask about the object, the dollar coin, in my pocket. Even though the
determinants of that object's being a dollar coin are broad, it does not follow
that that coin is not in my pocket, or that A's question is nonsensical. Why?
Because the location of the coin which is the bearer of the relational
properties is determined solely by another set of properties of the coin which
are not constitutive of its being a dollar coin.

My claim is that the case of intentional states is just like the case of
the dollar coin. Both possess relational properties, but these relational
properties can be born by some entity which is specified by other, often
nonrelational, properties.

It might be thought that since a dollar coin is a unit of currency in
virtue of its relations to The Royal Australian Mint and our community wide
socio-economic practices, the coin is not really in the pocket, since the dollar
coin somehow literally spans the intrinsically identified thing in my pocket
and the rest of the surrounding context. I think this is a crazy view, and I think that for same reasons that I think that a bearer of a relational property can possess certain properties (such as "being in the head") in virtue of their intrinsic properties. Let's see why.

1.3 Contextual Conferment

How an intrinsically specified entity can possess broad properties can be seen once the doctrine of what I call context conformation is understood. Contextual conferment is the bestowing of certain properties upon an individual entity or entities in order to facilitate certain goals or ends.

To see why dollar coins are contextual conferees consider a slightly different economic practice. In this practice the unit of exchange must be taken back to its point of origin or to a "point of exchange" in order to be verified as genuine. Since it can be verified, possession would not constitute ownership since the token unit of exchange's exchange history could be kept. For this reason one might as well not have some token which one transports to possible cites of exchange (where someone might want to buy or sell something) since one would have to take the token to a point of exchange anyway. This system would undoubtedly be very secure; it would minimise forgery, and it would hard to establish a currency black market. However, the system would be extremely slow and cumbersome. It would only make sense if it were possible to have the verifying and recording points of exchange just about everywhere. Under this scenario, the point of having dollar coins would be lost, since they were designed to allow exchange to occur wherever they were exchanged.

The main point to notice about this tale is that the units that get exchanged possess their properties of genuine units of exchange only in virtue of being processed by an official point of exchange. It is these processes occurring that provide a context in which those units are legal units of exchange, and for that reason the system of exchange becomes cumbersome and slow. The system can be speeded up if one confers upon the units, in this case what are to become known as dollar coins, the powers of exchange independent of being processed by the mint or points of exchange. In other words, if one confers the properties which the processing system used to apply to an individual or entity, when those processes are absent—ie. one claims that dollar coins are "legal tender"—then the system can be freed up.

Who knows, maybe the electronic banking revolution will enable these practical problems to be solved in the manner of the spread of EFTPOS (electronic funds transfer at point of sale).
It is this conferring of the properties originally provided by the context onto some entity which I call contextual conferment. What it does is to confer certain relational properties onto some object which is individuated, in part, by certain intrinsic properties such as being a coin of a certain size shape etc. It is such requirements for contextual conferment, and it is a mere practical requirement, that makes it possible that there are forgeries. To make a forgery all one needs to do is create a fake conferee which does not have the relevant contextual connections. This possibility is one of the costs associated with contextual conferment, but it's a cost worth living with given the advantages of the context conferment programme.

One should notice that context conferment of relational properties upon an entity seems to confer a certain type of relational property; not all relational properties can be context conferred. Consider the relational property of location. An object has its location just because of its relations to other objects or a set of space-time co-ordinates. It would seem that location is not a relational property that can be context conferred since that would mean that an object could be at some location when, in fact, it was not in that location. Whatever the characteristics are of the relational properties which can be context conferred which distinguish them from relational properties such as location, I don't know. The point is that there is a difference between various relational properties, and we must keep that in mind when considering the relationality of intentional states. I will be claiming, of course, that intentional states possess context—conferrable relational properties.

One important difference between non-conferrable and conferrable relational properties is that the need for conferring certain properties upon an object in some sense determines some of the intrinsic properties of the object conferred upon. I possess the height that I do for reasons not much to do with my relational property of being the same height as Bruce. However, in the dollar coin case, the objects must be easily transportable and easily exchanged. A ten tonne rock will not satisfy the demands required by the conferring of context. Similarly, when a collection of objects have context conferred upon them, normally there will be a relatively small number of types of objects conferred. If the number were too great, then that would also militate against the reason why conferring was introduced; if all dollar coins were different, then we would have trouble determining if we were being offered legal tender or not.
1.4 Beliefs and Context Conferment

I want to claim that context conferment is ubiquitous. Being a dollar coin, a Holden car, made in Japan are all conferred relational properties. Linguistic tokens also possess conferred relational properties. Consider a linguistic analogue of the above fictitious dollar coin story, in which the language is an “os·ension only” language or, if that’s impossible, a language in which tokens can be used only the presence of their referents. Such a language would be clumsy and unworkable, and really misses the point of what a language does. It’s because the items in a language refer, that the referents of the terms need not be around when the term is employed. In this case, the linguistic item refers or means because of the context in which it is used, where that context must always be immediate. In this language one could not refer to Australia unless one were in Australia. In order to free up the language, in the same way that exchange needed to be freed up, the tokens of our language are context conferred. In this way, it is not the intrinsic properties of a token that determine what it means, but the relational properties conferred upon the token which determine its intentional status. Notice again that some of the intrinsic properties of the token have been especially chosen with an eye to the ends desired by context conferring: the tokens are constructed out of a limited alphabet, for example.

And ditto for intentional states. Imagine a complex organism which literally does not represent its environment—it has no memory capacity. Perhaps all it could do is believe and desire what it was currently perceiving. I suspect that a species with such limited capacities would not survive long. There is a limited sense in which such an organism does represent its environment: it represents its occurrently perceived environment. But it’s a form of representation which is inadequate since its “representations” are literally tied too closely to their contextual objects. In order to count as a true representor the organism’s cognitive structure would have to have context conferment upon its internal states. In this way some internal state of the organism will represent according to certain broad or contextual determinants when those determinants are not in evidence. I stress here that it is the internal states which are the relata of the representation relation. That this can be so is guaranteed by the presence of context conferment.3

3Each of the earlier examples I have cited has had the intervention of an intentional agent in the process of deciding that context conferment is required in order to achieve the desired ends. In the intentional state case the forces guiding the development of cognitive systems which possess intentional states will have to play the intervening role. Those forces, I presume, will be fitness enhancing or Divine.
As with the previous examples of context conferred relational properties, the intentional state case will also have certain of the conferred upon object's intrinsic states determined by the relational properties. Let's assume, contrary to the best evidence available, that individual neurones are context conferred. That is, each neurone represents some thing in the environment—one's grandmother, say. In the process of being context conferred some of the intrinsic properties of that neurone might change. The excitation threshold of its synaptic connections might change; they change so that it fires in the presence of grandmother or when she is thought about.

This apparent effect of context conferment upon the entities which are conferred upon is crucial to the current concern. The theoretically puzzling feature of the intentional realist programme, which the argument from broadness plays on, is the one of determining how can context get into one's head or onto a dollar coin. What the discussion of context conferment is supposed to do is give some idea as to how this is possible. We will return to the explication of how context conferment achieves this in section 3.

2 Individuation vs Constitution

I think that the intentional waters have been muddied by a mistaken view of what the argument from broad content shows. One way to take the conceptual data evident in that argument is to think that context plays a role in the individuation of certain entities. It is only after consulting the context that we decide that we have dollar coins, since those hitherto intrinsically specified objects now have a set of additional relational properties. On this reading, the entities are individuated just because they have been context conferred. I will call this broad individuation.

The second reading of the argument from broadness takes the context as somehow constituting part of the relationally specified object. Broadness generates ontology, in other words. This is roughly the view attributed to Garfield above. The view seems to be that entities with relational properties "carry their context around with them", in much the same way that an object with location carries its context around with it, except in our case that context is somehow part of the constituency of the object. This reading ignores the distinction between context conferred and non-context conferred relational properties.

By construing broadness as an individuation criterion I am claiming that intentional states, and, a fortiori, content bearing mental states generally, are, and can be, states within individual organisms. At the same time, these states are not individuated by a non-broad narrow semantical
theory. This view, which postulates a narrow ontology but broad individuation, has been dubbed, rather pompously, Naturalistic Individualism by Garfield. He describes the view thus:

Although the facts responsible for determining interpretations are on this account naturalistic (comprising relations of the organism to distal stimuli and objects), the phenomena that get interpreted (only internal states and processes), and the generalisations over them, obey the individualistic supervenience principles of the Individualistic [narrow—J.F.] Theory of Meaning. The meaningful phenomena that are the theoretical entities of such a psychology are still, on this account, individualistic states and processes within individual organisms. ... But Naturalistic Individualism tries to have its cake and eat it too. Having established the role of naturalistic evidence in the interpretation of events or states as representational, the Naturalistic Individualist turns around and asserts that, despite the essential role that the nonindividualistic properties of these phenomena play in their individuation, their nature—quasi psychological phenomena of particular types—can be specified individualistically. (1988 p. 94)

This is a fair description of the way I think the intention realist should spell out her theory. The only questionable point in the description is the claim that the individualistic supervenience base is adhered to by the view I am advocating. Garfield's point is that if intentional states are individuated broadly, then it would seem that those states must supervene on the conjunction of the intrinsically specified bearers of the relational properties and the context. Dollar coins, it might seem, supervene not only the intrinsically specified bit of metal, but upon context which has been conferred upon it. I think this is not quite right.

Garfield goes on to argue that this balancing act cannot be maintained, since the arguments against intentional states being individuated narrowly also count against Naturalistic Individualism. The reason he gives is the very reason I objected to his description: broadly individuated states will supervene on context. From this he draws the ontological moral that "the correct way to describe these phenomena ontologically would be as relations between their bearers and their environments" (1988 p. 106). This is the exact move advocated under the constitutive broadness strategy. Notice, though, that the intentional state is now described by Garfield as a relation. Constitutive broadness ends up confusing intentional relata with the intentional relations in the way
Chapter 8

Broad Content

174

described in section 1.1. Broad individuation with its commitment to context conferment is designed to avoid this confusion.

There is a puzzle here though. If an entity which has been conferred with context supervenes upon that context, then why isn’t that entity ontologically or constitutively broad? The whole point of introducing context conferment was to get an ontology that was not broad so that certain practical goals could be achieved. The sense in which such an entity supervenes on its context is that *qua* representer, one of the relata of the representation relation, it supervenes upon that context. An individualistic or intrinsic ontology combined with conferred broad properties gives us those goals. But how do we generate that individualistic ontology? How can a broadly individuated entity ontologically supervene on the individualistic?

3 Supervenence

As mentioned in the introduction to this chapter, the argument from broad content is supposed to cause problems for mind-brain supervenience. Since X supervenes on Y iff there is no change in X without a change in Y, then psychological states can only differ if brain states differ. But the twin earth cases show that agents with identical brains can possess different propositional attitudes. The argument from broad content, as stated, is an argument about how we type-individuate mental states. If mental are broadly type-individuated as the argument claims they must, then it follows that any mental state token will be individuated broadly, just in virtue of being tokens which fall under that type. That means that the argument from broad content may be construed as an argument against *token* mental states being individuated narrowly or non-relationally. Supervenience is often thought to yield token identity. Therefore, the argument from broadness is actually an argument against mental state *tokens* being supervenient upon, or token identical to, brain state tokens, since brain states are thought not to be relationally individuated.

3.1 Levels of Explanation

Fodor asks (1987 p. 30) if there really is a problem about supervenience to be solved in the face of the argument from broadness. Ultimately, he thinks there is some challenge to answer, but thinks that the problem arises only if we ‘believe in the narrow or nonrelational individuation of brain states. So, the seemingly easy way to avoid the problem for supervenience would be to
argue that brain states are relationally individuated. He also claims that relationally individuating brain states is just plain silly (p. 31). Now Fodor does not give any explicit arguments why he thinks this line is that silly. Even ignoring the possibility (mentioned in chapter 3) of functionally individuating brain state types, where functional individuation is a paradigm case of relational individuation, in fact, the relational individuation of brain states follows from the mind-brain token identity: if a mental token is a relationally individuated type of mental state, and that mental token is identical with some brain state, then we have relationally individuated a brain state token state.

This is still just plain silly, according to Fodor, since it means that the individuated brain states will then fail to supervene on molecular states. If one then makes the same move as in the mental state case, it will follow that molecular states must also be individuated relationally. And ditto for all the other levels of description down to the very basic level (if there is one).

Where the hell does it all stop?

The answer is that it both does stop and doesn't depending upon what one is talking about. Relational individuation stops at the psychological level (or some neurofunctional level) when it comes to the individuation of the kinds at that level. Mental states types are individuated relationally with respect to context, if the argument from broad context is correct. However, neurobiology, molecular biology, chemistry, and physics might not individuate its kinds relationally with respect to context (they may, of course individuate kinds relationally, but not relationally with respect to context).

Relational individuation does not stop when we have a token mental state which is token identical to a state of the brain, a collection of molecules, and collection of the basic stuff of the universe. To the extent that this mental state token consists of, say, molecules, then that collection of molecules is individuated relationally for psychological purposes. Of course, that collection of molecules does not constitute a chemical kind, but does fall under a psychological kind just because the token identities which exist.

We can see this more clearly in the dollar coin case. That clump of minerals does not itself form a chemical kind, nor does the collection of physical particles constitute a physical kind. However, those collections do fall under one of the kinds of units of currency. It is these units of currency as a type that is individuated relationally. So, the collections of molecules or particles which make up the token unit of currency have been individuated relationally.
3.2 Causal Powers

If the claims of 3.1 are correct then the puzzle regarding the supervenient base of broadly individuated entities has been dissolved. There are, however, two more stumbling blocks in the path of broad intentional realism, both of which have to do with the causal powers of mental states. The first has to do with how mental states can cause behaviour as result of possessing the intentional content that they do, given that intentional content is individuated broadly. We can understand how putatively narrow brain states cause behaviour, but if mental states are necessarily broad, how can those contextual features affect the causal powers of internal causes of behaviour? In other words, how can broadly individuated states be causally relevant?

The second stumbling block has again to do with causal powers. Fodor claims that individualism is a thesis about individuating mental states only in ways that affect causal powers. Both me and my Doppelgänger perform the same actions, so we should attribute to both of us mental states with identical content. But the argument from broad content tells us that we should type the contents of our mental states as distinct.

We take these stumbling blocks in order.

3.2.1 Intentional Causation

The reply to this problem has been partially covered already in this section. Remember that what gets individuated according the principles of the argument from broad content combined with the idea of context conferment, is a state of the brain. Stated in this way it's difficult to see where the problem lies. Perhaps the problem is more evident with some different terminology. Block (forthcoming) describes the situation as a paradox, generated by the following three premises:

1. The intentional content of a thought (or other intentional state) is causally relevant to its behavioural (and other) effects.

2. Intentional content reduces to meanings of internal representations.

3. Internal processors are sensitive to the "syntactic forms" of internal representations, not their meanings.

The most obvious response to this set of premises is to claim that mental representations have their syntactic properties just because of their semantic
properties. We saw in chapter 6 that to a large extent semantics affects syntax. As we have also seen, Fodor, for instance, recognises that in order for intentional realism to work, syntax must at least mirror or “mimic” semantics. But it’s no accident that this mirroring takes place in a mental representational system. The syntactic properties of the representations derive from the links to the object of the representational system, viz. the world which provides the context. We have already seen above that context conferment can bestow certain properties on the relatum of a broad relation, in much the way of the grandmother neuron’s firing potentials being so affected. The broad/semantic-narrow/syntactic dichotomy is surely a misleading and false dichotomy.

To see why there should be no problem here, consider again the dollar coin. There is no analogous problem in the dollar coin case. The properties of the dollar coin that are analogous to that of a mental representation are its being used in legitimate economic transactions. In the same way that context ensures that the mental representations feature in the etiology of behaviour, the dollar coins feature in exchange transactions just because of their relations to context. However, in the case of the dollar coin we see no mystery, so we should not see any mystery in the case of intentionality either.

3.2.2 Broad Individuation

There is still a problem lurking in 3.2.1. The problem arises in cases where mental representations seem to have different meaning but their syntactic properties are identical. Both me and my Doppelgänger have mental states whose syntactic properties must count as identical since our behaviour is identical, but we represent different things. We thus arrive at the second stumbling block.

If the line I am running here is correct, then it follows that the intentional realist should adopt a broad conception of content in her programme. There is a problem with this though, inasmuch as intentional realism, to the extent that it is a view of how scientific psychology will develop, attempts to map the etiology of behaviour. However, the etiology of behaviour seems to be the same across me and my Doppelgänger, whereas accepting a broad criterion for the individuation of content leads us to distinguish between etiologies. How can such a tension be resolved?

One way to dissolve the tension would be to argue that the behaviours of myself and my Doppelgänger are different. It is possible to deny that our behaviours are the same, by claiming that I pick up a glass of water, where as my twin picks up a glass of XYZ. So, the mental states will have different
causal powers since they elicit different behaviours. However, as Fodor objects (1987 pp. 41-2), that would mean that context could affect the causal powers of our mental states without affecting our brains since token identity is assumed to hold, and, ex hypothesi, our brain states are the same. The claim here is that to type-individuate behaviours as distinct would be to give up the local supervenience of causal powers. Asks Fodor: “How could differences of context affect the causal powers of one’s mental states without affecting the states of one’s brain?” (1987 pp. 41-2).

I take it that “affecting the states of one’s brains” is to alter the properties of those brains. Now does my brain and my Doppelgänger’s share all their properties in common. Obviously not, since my brain is possessed by me and not my Doppelgänger. There is another difference: the information bearing states of my brain have transacted with water where as the states of my Doppelgänger’s brain have transacted with XYZ. That’s surely a difference, but the issue is whether or not that difference constitutes a difference in causal powers. The following suggests that it does.

I said above that we could think of the causal powers of a cognitive state as being given by its set of syntactic properties. Now as we saw in the outline of Stich’s STM above, the syntactic properties of a state will be derived from a functional specification, where that specification will be spelled out in terms of causal roles. Something has causal powers to the extent that it can fill certain causal roles. Now the causal roles that determine something’s causal powers should include both actual and potential causal roles (as should be the case with functional role semantics from chapter 1). While me and my Doppelgänger would seem to possess mental states which have the same causal powers in 1750, or whenever the pre-molecular time was in which the Putnam story was set, once our chemistry was developed and it was discovered that water was on earth but XYZ was on twin earth, then the causal roles of those states across me and my twin come apart. I would say: “When I said in 1750 that there was water in my cup, I was talking about H\textsubscript{2}O” whereas my Doppelgänger would say: “When I said in 1750 that there was water in my cup I was talking about XYZ”.

The immediate response to this is to claim that this does not show that the causal roles are different since if you switched me and my Doppelgänger just as the chemical discoveries were made, then I would have uttered what my Doppelgänger uttered, and vice-versa. All one has to do here, though, is to alter the counterfactual case to one in which the switch has been made but I and my Doppelgänger know of the switch. The causal roles of a state depend upon other states with which it interacts such as my believing that I am on
the same planet that I was one minute ago. The situation seems to be this. There all too many stories one can come up with where the causal powers do not differ. However, if we find an instance where they do, the supporter of broad content and individuation can maintain his view in the face of the Fodorian objection, since she will have a case of causal powers differing.

Moreover, I think there is a strong independent motivation for going against Fodor, when one considers the issue of the meaning change of contentful states. Remember that in the case of the original Putnam story, we might have been willing to grant that me and my twin had different beliefs after the state of chemical knowledge increased. What Putnam had to convince us of was that the meaning of ‘water’ did not change subsequent to those discoveries. One of the advantages of the broad individuation is that we can see how the meaning of ‘water’ in both my language and the language of Doppelgänger remains the same across the chemical discoveries post 1750.

Let us consider the question of meaning change under the narrow individuation of Fodor. Looking at me and my Doppelgänger synchronically, as it were, in 1750 there does seem to be a pull to the view that the causal powers are the same (although I have just given some reasons why I don’t think they are). So we can claim that they have the same narrow content. However, what happens to the narrow content of our respective mental states over time? When the chemical discoveries were made it seems that the causal powers of one’s earlier mental states have changed, since we will now be disposed to make different utterances about the referents and truth values of our long held mental states. What this means is that Fodor can have his individuation if he wants, but if he does, he will be forced into claiming that the narrow content of beliefs and desires change depending upon certain discoveries regarding context. Since contentful mental states have their content essentially, this means that the mental states over which scientific psychology quantifies will change. Fodor has said nothing about how the narrow content of mental states can change over time. He needs to tell such a story. In sum: it might well seem counterintuitive to some that we type the broadly individuated mental states of me and my Doppelgänger as distinct in 1750, but it is going to be equally as counterintuitive that some of our mental states change meaning after 1750, according to the narrow content programme.

While I think these considerations ought to be telling against the narrow individuation opted for by the intentional realist, there is another line which also seems compelling. Fodor gives the broad individuator two options: she can either give up mind-brain supervenience (yuk!) or go broad on the
individuation of brain states. What he does not allow is the broad individuator to split the difference:

If supervenience be damned for individuation, it can't be saved for causation. Burge says that "local causation does not make more plausible local individuation" (p. 16), but he's wrong if, as it would seem, "local causation" implies local supervenience of causal powers. Local causation requires local individuation when so construed. You can have contextual individuation if you insist on it. But you can't have it for free. Etiology suffers. (1987 p. 42)

I suppose it is the point of the doctrine of context conferment being advocated here that one can, as Fodor puts it, split the difference. The reason is that context conferment can affect some of the locally supervenient causal powers of the object conferred upon, or more generally, in context conferring, we choose an entity that has a range of actual and potential causal powers that suits the need from which we context conferred in the first place. There is also a sense in which there are two senses of causal power at play here. In the brain state case, we know that the states conferred upon will have certain causal powers at the neurological level of description. We can even assume that they have those powers because of their intrinsic properties. However, because those brain states are token mental states, they also possess certain causal powers with respect to the psychological level just because they fall under a psychological kind.4 Now it's context conferment that can bestow such psychological level causal powers onto mental states which are token identical to brain states, and those brain states are chosen just because of some of the causal powers they possess.

Signpost

The thrust of this chapter is the claim that since there is no broad-narrow currency distinction in the case of the dollar coin, to the extent that the two cases are analogous, there should also be no need for the broad-narrow content distinction. To that extent there is going to be no problem for intentional realism in terms of mind-brain supervenience. To be sure, broad individuation will lead to a taxonomy of mental states different from one's

---

4I'm assuming here that causal relations can be exhibited at many different levels of explanation-description. That in itself is a controversial thesis, and argued for by Peter Menzies (1988) and argued against by Braddon-Mitchell (Unpublished). For the sake of argument I'm assuming it here since Fodor also assumes it in his (and Pylyshyn's) (1987).
initial expectations. However, if certain counterintuitive results are to be avoided, and we are not going to confuse intentional relations and relata, this unexpected taxonomy generated by broad individuation will be all too easily tolerated.

Having, hopefully, accomplished that task, it is time—at last—to find out what I think is wrong with intentional realism. Basically, I think it is the cognitive state realism component of the view that is the culprit. So, it is an examination of that thesis to which we now turn.
Chapter 9

Two Theories of Cognitive Architecture

If the traditional objections to intentional realism don't work, as I have claimed, then what line of criticism should we adopt? The strategy we should employ is to argue against intentional realism by rejecting the thesis implied by intentional realism viz. cognitive state realism. What kinds of argument might be summoned against cognitive state realism? I don't think there are any apodeictic a priori arguments that will work. Ultimately, the truth or falsity of this thesis, as with any in the discipline, will be empirically established. So, as well as examining some of the more conceptual style of arguments against cognitive state realism, we may also look at what might constitute empirical arguments against it, i.e. what type of evidence might count against cognitive state realism.

Cognitive state realism is a thesis about cognitive architecture. It postulates cognitive mechanisms which range over representational states with the properties listed in chapters 1, 4 and 5. One way to cast doubt on whether the cognitive state realist account is the correct form of architecture is to survey some other theories of cognitive architecture. In this chapter I wish to compare and contrast the architecture of cognitive state realism with another theory of cognitive architecture. The alternative theory we are to examine is based upon the idea of a vertical faculty as introduced in chapter 5; we may call it vertical faculty theory (VFT, for short). In order to get the clearest idea of the underpinnings of that programme, I think we should ideally examine in detail the work of the originator of the idea of a vertical faculty, Franz Gall. That, however, is a basically historical task, tangential to our direct concerns. I, therefore, direct the reader to Appendix A for those underpinnings. Instead, we can move on to the major tenets of vertical faculty theory, some of which we have come across in earlier chapters.
Chapter 9 Two Theories of Cognitive Architecture 183

1 Vertical Faculty Theory

Jerry Fodor has dubbed Gallean faculties “vertical faculties” and the traditional philosopher’s faculties—ones consistent with cognitive state realism—“horizontal faculties”. Gall agrees with the philosophers that the mind is composed of functionally isolatable subsystems. He disagrees as to how the division should be drawn. As has already been intimated, Gallean vertical faculties are distinguished by reference to their subject matter—i.e. they are *domain specific*. As we saw in chapter 5, domain specificity, as such, will not provide a taxonomic edge in the way Fodor envisages. Instead, vertical faculties are distinguished by their taking as their domain of operation, domains which are relatively *fine* grained. Cognitive state realist architecture, on the other hand, ranges over mechanisms whose domain of operation is coarse—they answer a wide range of questions in the performance of their functions. Vertical faculty theory and cognitive state realism are, then, alternative ways of slicing the same cognitive pie.

By way of formalising this alternative approach to cognitive architecture, we may follow Fodor and summarise the main ingredients of Gallean slices of the cognitive pie thus: vertical faculties are *domain specific* (have a fine grained domain of operation), *genetically determined* and realised by *distinct neural structures*. Fodor adds a fourth ingredient: he claims that they are *computationally autonomous*. That vertical faculties are realised by distinct neural structures is central to the historical Gallean programme; but is not a component of vertical faculty theory. The same goes for the innateness of vertical faculties. For present purposes domain specificity and computational autonomy are the crucial properties of vertical faculties we need to explore. As we have already examined domain specificity in some detail (in chapter 5), we may take a look at computational autonomy in the next section. Following that I will say why the questions of neural realisation and genetic determination are not relevant to our current concerns.

1.1 Computational Autonomy

What Fodor means by computational autonomy seems to be that not only are vertical faculties distinct in the function they perform, but they are also independent in the performance of those functions (Fodor 1983 p. 21). He says: “That we can, most of us, count and chew gum at the same time would have struck Gall as a fact that offers significant perspectives upon our mental organisation” (p. 22). But Gall also wanted vertical faculties to *interact*. He says: “We have to discover the fundamental powers of the mind, for it is only
these that can have separate organs in the brain" (Quoted in Marshall 1980 p. 23). That there are separate powers should not lead us to think that they cannot interact. Non-fundamental powers or capacities result from the interaction from the fundamental and distinct faculties. He says: “Indeed, we believe that the total nervous system is a combination of many; that all these individual systems differ in their office; ...that more or less of a bond, and therefore of reciprocal influence, exists between all the individual systems” (Marshall 1980 p. 24). The crucial point here is how the individual faculties “differ in their office”. Gall would not want the destruction of one faculty to impede the performance of any other fundamental faculty, and yet he wants interaction between the faculties.

We can take the idea of the autonomy of vertical faculties two ways. The first is the claim that vertical faculties are autonomous in the sense that a faculty is an actual functional mechanism, with some specific function to perform. In Pylyshyn’s terms, the vertical faculty would have to appear as part of the functional architecture of the system under analysis. Or, in terms more in keeping with the terminology of this work, the vertical faculty will feature as a module within the Level Two functional description of the system.

I suppose ultimately we will require more complete individuation conditions for what constitutes a cognitive mechanism than these very abstract characterisations. Such conditions, though, are required not only by the vertical faculty theorist but any theory of cognitive architecture, including cognitive state realism. So, not much hangs on the formulation of those conditions for our current concerns. I read Gall as claiming that faculties’ “differing in their office” amounts to autonomy in this sense. That vertical faculty theory is committed to this sense of autonomy is incontrovertible; any cognitive theory—including cognitive state realism—must be so committed. However, the other sense of autonomy, the sense intended by Fodor, should not be thought of as part of the vertical faculty programme, for the following sorts of reasons.

The sense in which a mechanism can be “independent in the performance of its function” is grossly unclear. Suppose that some vertical faculty processes outputs from some perceptual system and our faculty fails to operate because of damage to either that perceptual system or the faculty's connection to that system, then should we claim that our putative faculty is not computationally autonomous? The functioning of our faculty would seem to “depend” upon the perceptual system. Or perhaps computational autonomy amounts to a vertical faculty's not having to “share computational resources” whatever exactly they are. Such resources might be information about the
world, where two faculties compete for that information stored in some memory. The case for this sense of computational autonomy of vertical faculties is not prima facie obvious given certain psychological data. The kinds of data I have in mind are attentional studies. The idea is this. If faculties really are computationally autonomous, then attentional studies in one modality should generate results independent of results in other modalities. This, however, is far from being obviously the case. According to some studies there is a correlation in the ability to execute auditory and visual attention demanding tasks (Earl Hunt 1985 p. 20). There also exist data from studies of interference patterns affecting simultaneous task execution. Even when tasks would seem to involve different faculties, interference is almost always observed. Admittedly, there are some exceptional cases. Balancing, for instance, interferes with visual memory but not verbal memory. Interference would also seem to be reduced by learning.¹

1.2 Localisation and Equipotentiality

Central to Gall's programme was a commitment to cerebral localisation. Historically (see Appendix A), the debate between vertical faculty theory and the antecedents of cognitive state realism took the verticality-horizontality dispute to be if not the same dispute as the localisation-equipotentiality dispute then, at least, to be dependent upon it. If it turned out that equipotentiality of the brain ruled the day, then the vertical faculty theory was a goner. Hence, a great deal of time and effort was, and still is, devoted to determining the extent, if any, to which cognitive functions are realised in particular regions of the brain. However, these two disputes really are independent disputes; and for that reason VFT is not committed to the sort of localisation of function its historical antecedents had in mind.

The crucial thing to realise here is that there are various localisation theses. The most general is that which claims mind-brain supervenience or identity. Gall certainly needed to push this line (see Appendix A). Another thesis is that cognitive mechanisms are localised in neuroanatomical states. Language comprehension is sometimes thought to be localised in the first temporal gyrus (Wernicke's area) and language production localised in the third frontal gyrus adjacent to the Rolandic motor strip (Broca's area). It's this localisation thesis to which VFT is uncommitted. If the contention of chapter 3 is correct, then, although vertical faculty theory is construed as a psychological level model, it is going to be strictly realised in some high level

¹Hunt 1985 p. 20). Actually, these studies concern input systems or Fodorian modules and not faculties as such. These modules will be described below.
neurofunctional states—or perhaps even organs—of the brain. However, these states are not anatomical states in the mould of the traditional localisation-equipotentiality debate.

The reason why VFT is not committed to anatomical localisation of function is that whatever story one comes up with regarding cognitive architecture, that story, is going to be, in principle, consistent with the equipotentiality of the brain holding. That's because the localisation-equipotentiality dispute is a dispute at the implementation level. What holds for VFT here, also holds for cognitive state realism. The cognitive state realist's metaphorical intentional boxes might well be localised in particular anatomical regions of the brain as well. This is in fact just what the old faculty theorists claimed, although they localised horizontal faculties in the ventricles rather than the cortex. What is at issue between vertical faculty theory and cognitive state realism is the nature of the cognitive mechanisms which may or may not be realised in anatomical organs, the former opting for fine grained mechanisms, and the latter coarse grained mechanisms.

Granted the conceptual independence of these disputes, just what is the relation between them? I take it that there is an evidential relation between the disputes. For suppose it turned out that the localisationists win the day, then the functions localised will support either VFT or cognitive state realism, depending upon the functions. Suppose, on the other hand, that the equipotentialists win. That won't in itself support cognitive state realism over the vertical faculty account. It just means that the implementation level is not going to provide a way of deciding between the two theories of cognitive architecture. The reason for this is that if localisation turns out to be true then it follows that we must have some idea of what the functions are that are localised, while it's not clear that, assuming equipotentiality, we have to have a completed story of the functions of the brain.

So, the fuss about localisation consists in the hope that by examining the disruption of neural structures we might get some idea as to what functions should feature in our cognitive model. This is exactly the task of the cognitive neuropsychologist. Of course, the task of the neuropsychologist is not an easy one. A correlation between a lesion in one area of the brain x and loss of a particular cognitive function does not imply that area x is the neural centre responsible for that cognitive function. X might perform some necessary but not sufficient role in the utilisation of a cognitive capacity (Patricia Churchland 1984 p. 143). For example, one might have the capacity to verbally communicate even though one cannot do so due to the impairment of the motor control mechanisms which feature in speech production. It might also be the case that cognitive functions are not correlated with any one area
of the brain. So a lesion to one area automatically causes deficiencies in other cognitive abilities. Hence, the brain lesion-behavioural deficit correspondence does not simply guarantee that we are determining the locality of a cognitive capacity. Despite these theoretical constraints upon the cognitive neuropsychologist's programme, all seems well for the disconfirmation or confirmation of our two theories of cognitive architecture.

1.3 Innateness

It is assumed by many proponents of VFT style theories that vertical modules must be innate. The following considerations suggest that they ought not be. Referring back to footnote 7 in chapter 5, we saw that Fodor changed his mind about there possibly being a chess playing module. In his (1983) he claimed that there is no way that a modularity theorist would want to claim that there is a chess playing module. One reason for why there might not be a chess playing module is that it is hard to see how a mechanism specific to chess playing could result from innate structure. In his (1987a), however, he allows for there being "brute force" modularity in addition to modularity "in the nature of things". If there is a module ("in the nature of things") for chess, then it is not a standard Fodorian module. In addition to not being innate (as standard Fodorian modules are supposed to be), a chess playing module is not a perceptual mechanism at all. For that reason, there not even a prima facie case to be made for its being a module of the standard Fodorian type. If anything counts as a higher cognitive process, chess playing would seem to. Chess board perception might well have an accompanying specialised mechanisms but chess playing would seem to require more intelligent processing than mere recognition of positions.

In virtue of what might the chess playing module develop? The answer to this question lies somewhere in the expert-novice literature from psychology—eg. Larkin, McDermott, Simon and Simon (1980) and Chi, Glaser and Rees (1982). Perhaps what separates the expert from the novice in some cognitive task (playing chess or solving physics problems) is the emergence of certain complex productions consisting of condition-action pairs. When that condition is met then the action is automatically triggered without the need for vast amounts of processing which would slow down the agent's processing time of the problem under consideration (Larkin et al. p. 1337). More metaphorically, it is suggested that the expert's cognitive system executes instructions in a compiled rather than an interpreted manner. If this is right there is no need for an interpreter—whatever it is—to execute one instruction at a time; execution after compilation has the advantage of running many
instructions together rather than separately. These productions might be the mechanisms underlying the phenomenon of chunking. Chunking is thought to be in part responsible for expertise in domains such as problem solving. If a chunk is generated in STM (short term memory), then that one symbol stored actually represents many pieces of information about that task, which in the novice would have to be stored separately—that's why it is a chunk (Newell and Simon 1972 p. 781). The advantage of chunks for the expert is that reaction time is decreased and amount of information stored is increased.

Assuming this conclusion, at what level of grain do chunks feature? Now chunking is a task relative phenomenon—in fact, extremely so. The mechanism of chunking only operates in the domain of expertise, to the extent that while the expert’s performance is enhanced by chunking in typical situations, (remembering the positions of chess pieces from an actual or possible game, say) that performance drops off dramatically when the situation becomes atypical (the chess pieces are arranged at random) (Larkin, McDermott, Simon and Simon 1980 p. 1336). This suggests that there is no general cognitive capacity evident here that cuts across various problem domains, contrary to what one might think if there were generalised cognitive mechanisms (a central processor, say) responsible for higher cognitive processes. Since problem solving ranks as a higher cognitive process, and if there is the close connection between information and cognitive mechanisms claimed in chapter 5, then the facts about chunking suggest that the mechanisms responsible for those cognitive processes can be fine grained in the manner required by VFT.

Fodor himself suggests that it is a form of informational encapsulation which generates this “in the nature of things” sense of modularity:

Chess playing, by contrast, is modular in the sense that only a restricted body of background information (call it chess theory) is relevant to rational play even in principle. This second kind of modularity...is interesting to the engineer, however, since informational encapsulation makes for feasible simulation regardless of what the source if the encapsulation may be. (Fodor 1987a p. 36)

There is no requirement that the process of the formation of vertical modules goes the way Fodor suggests; although it is consistent with the example model from the expert-novice literature, where the formation of productions in chunking allows processing without other parts of the cognitive system needing to be consulted; and that amounts to encapsulation.

All this suggests that there are non-perceptual cognitive tasks that can be performed by mechanisms which can be described as modular; and
that is all the VFT requires. The mechanisms are modular to the extent that they are required in our functional decomposition of the cognitive system in order to account for the data of cognition, in this case the differences between experts and novices. Because such mechanisms are not brute force, they are not innate. Even if it turns out that there is no chess module, the above discussion has provided the conceptual possibility of non-innate vertical modules; since I am currently listing the essential elements of VFT, that possibility is all we currently require. The claim of VFT is simply that whatever the set of mechanisms postulated by cognitive theory, those mechanisms will not be coarse grained.

1.4 VFT and Fodorian Modularity

We have seen that in The Modularity of Mind Fodor offers a tripartite functional taxonomy of cognitive processes. Transducers provide the interface between an organism and its environment. Distal environmental objects cause the impingement of energy at various surfaces of the transducers. These patterns of proximal stimulations are then converted in some lawful way into neural code—the transducer's output. Input systems "mediate between the transducer outputs and central cognitive mechanisms by encoding the mental representations which provide domains for the operations of the latter" (Fodor 1983 p. 42). What this means is that input systems perform an inference making role: they take the transducer outputs which specify proximal stimulation patterns and "deliver representations that are most naturally interpreted as characterizing the arrangement of things in the world" (Fodor 1983 p. 42). However, the output from input systems is only going to give the organism the way the world looks, sounds or feels, rather than the way the organism believes the world actually is. The organism's beliefs about how the world is should be constructed out of (inter alia) these input system outputs in conjunction with background information about how good the seeing, hearing or touching is. So the operations of the input systems should not be identified with belief-desire fixation since what we believe depends upon this very background information. According to Fodor, then, it is central processes, those mechanisms which access the informational outputs from input systems which perform belief fixation.

The properties of these input systems, or "modules" as Fodor calls them, need not be spelled out in detail here.\textsuperscript{2} The point about them is that

---

\textsuperscript{2}For the record they are: domain specificity, informational encapsulation, mandatoriness, speed, being associated with distinct neural structures, exhibit characteristic and specific breakdown patterns, their ontogeny exhibits characteristic
they make input systems a member of the class of vertical faculties. To the extent that Fodor is committed to the verticality of modules, he is an adherent of VFT. However, as we have seen Fodor also thinks that there are central cognitive processes whose domain of operation is cognitive states. Now central processes are domain inspecific; so to this extent Fodor adopts a mixed cognitive taxonomy: partly vertical and partly horizontal. VFT, as I envisage it, in fact committed to the verticality of much more than perceptual input systems; VFT claims that higher cognitive processes are performed by domain specific cognitive mechanisms, that is by mechanisms which do not operate over cognitive states. The arguments for this position feature in the section 3. By way of rounding off our discussion of the nature of vertical faculty theory, I wish to give some idea of what the implications of accepting the VFT are going to be.

2 Implications of VFT

Let's return to Fodor. He says:

We have suggested that the characteristic function of modular cognitive systems is input analysis and that the characteristic function of central processes is the fixation of belief. If this is right, then we have three ways of taxonomising cognitive processes which prove to be coextensive:

FUNCTIONAL TAXONOMY: input analysis versus fixation of belief.

TAXONOMY BY SUBJECT MATTER: domain specific versus domain neutral.

TAXONOMY BY COMPUTATIONAL CHARACTER: encapsulated versus Quinean/isotropic.

I repeat that this coextension, if it holds at all, holds contingently. Nothing in point of logic stops one from imagining that these categories cross classify the cognitive systems. If they do not, then that is a fact about the structure of the mind. Indeed, it is a deep fact about the structure of the mind. (Fodor 1983 p. 112)

pace and sequencing, consciousness has limited access to them and they have shallow outputs.
If this coextension does hold, then that would be a deep fact about the mind. VFT is committed to the coextension not holding. The VFT claims that there is no actual interesting taxonomy by subject matter: all cognitive processing can be accounted for by domain specific mechanisms. VFT can remain fairly neutral about how interesting the taxonomy by computational character turns out; maybe there are encapsulated input systems. Now if the VFT is right in all of this, then the functional taxonomy cannot be made. Why? Because in order to be made, that taxonomy conceptually depends upon the other two taxonomies holding, and, as we have just seen, VFT denies that taxonomy by subject matter holds.

Why is there a conceptual dependence? It is here that we can discharge the results from chapter 6 regarding the abstractness properties of cognitive states (cognitive states in our technical sense). Remember that cognitive states, of which beliefs are a species, are stimulus, S-R, and domain abstract, etc. What that means is that cognitive states can take just about anything as their content. So, the mechanisms responsible for the generation of cognitive states must be domain insensitive, or rather take domains of operation that are very coarse grained. This means that one can only get the functional taxonomy off the ground once one has a taxonomy by subject matter flying; and VFT denies that the latter taxonomy has wings. If VFT is correct, then cognitive state realism is in trouble, therefore, so is intentional realism.

A lot of work is being done here by the notion of the degree of grainedness of domains of operation and the abstraction properties. Whether that work is interesting from the point of view of cognitive theory is dependent upon how plausible VFT turns out to be. It is time to turn to the consideration as to whether VFT is to be preferred over cognitive state realism. Ultimately, the decision we, or rather future cognitive theorists, make will be an empirical one. However, I think we should begin looking at the evidence now. In that way, cognitive state realist cognitive architecture might be seen to be “not the only game in town”.

3 A Priori Arguments Against VFT

We have seen that VFT is committed to higher cognitive processes being performed by domain specific, or more precisely, fine grained cognitive mechanisms. This claim has two components: the first is an a priori claim to the effect that it’s conceptually possible to account for higher cognitive processes under VFT. The second component is an empirical claim to the effect that in all likelihood, the data can support VFT. We tackle the a priori
in this section, and the empirical in section 4 where evidence from neuropsychology will be examined.

Why might one opt for a cognitive state realist, or at least a mixed cognitive architecture, part VFT and part cognitive state realist? One might think that we are conceptually compelled towards cognitive state realism. Both Fodor and John R. Anderson think this way. We may look at Anderson first and then Fodor.

3.1 Anderson

In *The Architecture of Cognition* John R. Anderson says:

> The most deeply rooted preconception guiding my theorising is a belief in the unity of human cognition, that is, that all the higher cognitive processes, such as memory, language, problem solving, imagery, deduction, and induction, are different manifestations of the same underlying system. This is not to deny that there are many powerful, special-purpose "peripheral" systems for processing perceptual information and co-ordinating motor performance. However, behind these lies a common cognitive system for higher-level processing. ... The view that the mind is unitary is certainly not universally held; it may not even be a majority opinion. ... This faculty approach holds that distinct cognitive principles underlie the operations of distinct cognitive functions. The unitary approach holds that all higher-level cognitive functions can be explained by one set of principles. (Anderson 1983 pp. 1-2)

He gives three arguments for accepting the cognitive state realist style architecture over that of VFT.

The first is an argument from evolutionary development. The VFT postulates special faculties or organs whose operations are restricted to a particular content domain such as mathematics, chess, computer programming or sculpture. But these are content domains for which there was no possibility of anticipation in our species evolutionary history. So, we shouldn't postulate domain specific, innate mechanisms but rather a domain cross-cutting mechanism that while being operative in particular domains such as chess and computer programming do so only due to their general properties consistent with the species evolutionary history (Anderson 1983 p. 3).
The main problem with this argument is that no post-phrenological faculty theorist is going to postulate innate faculties that are as content specific as computer programming or chess. Not even a rabid nativist like Fodor claims that there is a chess module. Obviously, these cannot be anticipated in our species’ evolutionary history, and hence no innate genetically determined faculties ought be postulated for these abilities. Just because there seems to be some isolatable cognitive ability, it does not follow that there is an actual mechanism to be postulated by VFT. Even Gall thought that there were a collection of fundamental faculties which had to be discovered.

This objection also assumes that VFT is committed to the innateness of vertical modules. As argued in the previous section, VFT requires only that there be isolatable cognitive mechanisms. How we come to have them is another issue. My bet is that many of the mechanisms identified by VFT will be innately specified; but this does not follow from VFT itself. If the contention of section 2 is correct, then it might well be the case that there are chess playing or computer programming modules.

The second argument concerns the localisation of function. He says:

There is a tendency to regard the existence of “language areas” and other localisations of function in the brain as strong evidence for faculties. However, there is nothing necessary about this inference, as shown by a computer analogy: two programs can occupy different areas of computer memory, much as two different cognitive abilities might lie in two separate areas of the brain. However, the programmes may have identical principles. For instance, I can have one ACT simulation doing language and another doing geometry. Thus, there need be no connection between distinct physical location and distinct cognitive principles. The real issue concerns the uniqueness of the structure and processes underlying cognitive functions, not their physical location. (Anderson 1983 pp. 3-4)

Two points need to be raised regarding this passage. Firstly, Anderson is correct in claiming that it is not the physical location of cognitive function that is important. However, the important issue does not involve the uniqueness of the structure and processes underlying cognitive function. Instead the issue is about the grain of domain of operation of whatever structures and processes we want to postulate. The point VFT is committed to is not one of localisation of function, but that a function corresponding to a task domain gets operated upon by its own mechanism which performs that function rather than some other(s).
Chapter 9  
Two Theories of Cognitive Architecture 194

The second point follows from the first and shows what I take to be Anderson's chief misunderstanding regarding the VFT programme. Anderson's use of 'distinct cognitive principles' hints at his mistaken belief that the "states and processes" underlying a domain specific mechanism will vary across domains and hence faculties; but this is not what the VFT requires—although it may well contingently be the case. Domain specificity does not imply distinct cognitive principles as Anderson seems to assume. All that is implied is the distinctness of the content of the information being processed by the faculties—and not the distinctness of the principles by which that information is processed.

If we take Anderson's own analogy between programmes and faculties then just as two programmes can have identical principles (built up out of list structures etc.), two faculties might possess similar principles (input systems, memory or storage capacity etc.). This is the idea behind vertical faculties each possessing general horizontal properties such as memory and judgement mentioned when discussing Gall above. Thus, if Anderson has one simulation doing language and another geometry, although the underlying states and processes are the same, they are different applications of the one system of states and processes. Different applications of the one set of principles are just: two mechanisms which process different informational content but with similar structural features. In effect, then, Anderson himself has content specific mechanisms operating which are different applications of certain cognitive principles which is perfectly consistent with the VFT. In short, in this argument it is assumed that the VFT requires cognitive domains to exhibit disparate cognitive processing principles. But this is a false assumption—as was the case with the idiosyncratic processing strategies of Fodor (see chapter 5). Anderson seems to confuse the question of mechanism with the question of process when he speaks of "distinct cognitive principles". Higher functions might be performed by one big central processor in which the one processing strategy that processor uses is common to the performance of all those functions. Alternatively, there might be a collection of vertical mechanisms each operating according to their own idiosyncratic principles; this seems to be what Anderson is arguing against. However, there is a third alternative viz. that there is a collection of vertical mechanisms each operating according to more or less the same principles. This scenario is also consistent with VFT.3

---

3If the processing strategies of the vertical mechanisms are the same then why have them instead of the one big processor? See chapter 1.
The final argument is that faculties should not be postulated for the reason that the boundaries between these organs cannot be drawn *a priori*. He says:

> It is pretty clear where the activity of the lung leaves off and that of the circulation system takes over, but this really cannot be said for cognitive faculties. The lung and the heart are both involved in an activity such as running, but it is possible to identify their distinctive contributions. It has been proposed that there is a language faculty, a number faculty, a deduction faculty, and a problem solving faculty, but if there are such faculties, their activities are terribly intertwined in a task like computer programming. When we look at an expert programmer creating a programme, we cannot separate the contributions of the various faculties. Indeed, if we applied any reasonable criterion for individuating faculties, we would have conclude that computer programming was a separate faculty. This is because some of the core principles for this skill organisation, such as strategies for creating recursive programmes, apply across the entire range of programming behaviours and are seldom if ever evoked elsewhere. Since it is nonsense to suggest a programming faculty, we should be more skeptical of other proposed faculties. (Anderson 1983 p. 4)

I'm not sure why Anderson thinks that there is an *a priori* difficulty here. It would not appear obvious that distinguishing the respective contributions of the lungs and circulatory system in running is in any sense *a priori*. Surely, such a distinction is an *empirical* matter. I grant that it is difficult to make such a distinction in the case of which faculties contribute to an activity such as computer programming. Of course, at this stage we are nowhere near to coming up with anything like a complete list of faculties. So determining what mix is required in order to perform computer programming is just not possible. This is an epistemic shortcoming, given our present state of knowledge. The same state of ignorance would probably have been the case for the lungs and circulatory system with respect to running a few hundred years ago. It is not an *a priori* shortcoming.

What this objection should be claiming is that there is no way to individuate actual cognitive mechanisms in the way VFT requires. Now that *would* be a problem; but it is general problem for all cognitive theories which require the individuation of cognitive mechanisms. When does the activity of input systems leave off and central processes start? We need some answer to that question (at the boundary between perception and cognition Fodor must say, but has no account to give of how to draw that boundary), but if so, then
that is not an objection against VFT; it is instead a general methodological problem for cognitive science.

3.2 Fodor

In considering higher cognitive processes, Fodor asks two questions: Are there domain inspecific cognitive mechanisms? and if there are, are they modular (i.e. also informationally encapsulated, mandatory, fast, etc.)? His answers are that, yes, there must be domain inspecific mechanisms which are informationally unencapsulated. The arguments for the two answers are independent, so I will take them in turn.

3.2.1 The Argument for Domain Inspecificity

The argument for the necessity of domain cross cutting cognitive mechanisms is an a priori one. He says:

The general form of the argument goes back to at least Aristotle: the representations that input systems deliver have to interface somewhere, and the computational mechanisms that effect the interface must ipso facto have access to information from more than one cognitive domain. (Fodor 1983 pp. 101-2)

Three instances are given of this general argument. (a):

To a first approximation, we can assume that the mechanisms that affect this process work like this: they look simultaneously at the representations delivered by the various input systems and at the information currently in memory, and they arrive at a best (i.e., best available) hypothesis about how the world must be, given these various sorts of data. But if there are mechanisms that fix perceptual belief, and if they work in anything like this way, then these mechanisms are not domain specific. Indeed, the point of having them is precisely to ensure that, wherever possible, what the organisms believes is determined by all the information it has access to, regardless of which cognitive domains this information is drawn from. (Fodor 1983 p. 102)

(b) The mechanisms responsible for speech production must have access to what we see, hear, remember or think. Hence, such mechanisms effect an
interface amongst the vertical faculties and are not domain specific. (c) One of
the points of perception is that the world often turns out to be other than the
organism expects given background information in the possession of the
organism; but an interface between this background information and
perception must take place somewhere if that information is to be used in the
fixation of belief and production of consequent behaviour. Again, it seems that
there must be mechanisms which cut across the domains of the input systems
(Fodor 1983 pp. 102-3).

The above fail to support the necessity of domain inspecific
mechanisms. Remember that Fodor himself says:

domain specificity has to do with the range of questions for which a
device provides answers (the range of inputs for which it computes analyses); whereas encapsulation has to do with the range of
information the device consults in deciding what answers to provide. A
system could thus be domain specific but unencapsulated (it answers a
relatively narrow range of questions but in doing so it uses whatever it
knows); and a system could be nondenominational but encapsulated (it
will give some answer to any question; but it gives answers off the top
of its head—i.e., by reference to less than all the relevant information).
If, in short, it is true that only domain-specific systems are
encapsulated, then that truth is interesting. (Fodor 1983 pp. 103-104)

The argument deriving from Aristotle makes no mention of the questions that
mechanisms processing the outputs from input systems must answer. Each
instance of the argument concerns the amount of information used in the
processing of the outputs from those perceptual systems. So, these arguments
cannot yield the conclusion that higher mechanisms are domain inspecific. At
most, they support the conclusion that these mechanisms are unencapsulat-
ed. It is still possible given these examples that the questions posed are quite
specific, but are answered with respect to all the information available to the
organism. The only way this argument would work was if there were a
premise to the effect that all encapsulated mechanisms are domain specific;
but there is no such premise. While VFT can agree that the mechanisms
responsible for higher cognitive processes need to be unencapsulated, VFT
can be compromised only if it is shown that there need be coarse-grained
mechanisms at work.
3.2.2 The Argument for Informational Unencapsulation

The structure of this argument takes the form of an analogy between scientific confirmation from the philosophy of science and the nature of belief fixation in cognitive science:

central systems look at what the input systems deliver, and they look at what is in memory, and they use this information to constrain the computation of “best hypotheses” about what the world is like. These processes are, of course, largely unconscious, and very little is known about their operation. However, it seems reasonable enough that something can be inferred about them from what we know about explicit processes of nondemonstrative inference—viz., from what we know about empirical inference in science. (Fodor 1983 p. 104)

Fodor goes on, firstly, to suggest that nondemonstrative fixation of belief in science is unencapsulated; it is unencapsulated because it exhibits two properties: scientific confirmation is both isotropic and Quinean. He then goes on to suggest that central processes are also isotropic and Quinean, and hence, unencapsulated. He has independent reasons why central processes should exhibit these properties: assuming central processes to be unencapsulated, then a major problem confronting AI modelling of higher cognitive processing, viz. the “frame problem,” is exactly what we would expect under the assumption that those higher processes involve isotropic and Quinean computations (Fodor 1983 pp. 105-15). There are various conceptions of the frame problem. It can probably be best summed up as the problem of how to update an AI system’s database in the course of the system’s operation, such that only the data structures relevant to those operations are affected. A robot, for example, in moving from one room to another should update data relevant to only its position, and not data regarding, say, the value of pi, the name of its creator or the colour of the walls in the room.4 More will be said about the frame problem in 3.2.4.

Notice that these arguments attempt to show that central systems are informationally unencapsulated only. They are not designed to show the domain specificity of those systems. That, at this stage, is assumed to have been argued for under the Aristotelian argument considered in the previous section.

---

4For more on the frame problem see Dennett (1984a) and McCarthy & Hayes (1969).
Having thus outlined Fodor's argument, I do not intend to examine it in any detail. As might be already obvious, I agree that higher processes are plausibly considered to be unencapsulated. It is their domain specificity that is at issue when considering VFT. In particular, I want to show how domain specificity can be made consistent with the properties of being isotropic and Quinean. In order to do this, in the remainder of this section I will consider these two properties in some detail.

3.2.3 On Being Isotropic and Quinean

If the fixation of belief is isotropic, that means that the information relevant to the processes of fixation may be drawn from anywhere in the organism's cognitive economy. By "anywhere" is meant information from any content domain. In the case of scientific confirmation, then, "everything the scientist knows is, in principle, relevant to determining what else he ought to believe. In principle, our botany constrains our astronomy, if only we could think of ways to make them connect" (Fodor 1983 p. 105). It's fairly obvious that isotropic systems are also unencapsulated systems par excellence. By referring to any and all the information available to the organism, there is no question of the belief fixation occurring as a result of access to limited information.

Because the information available to an isotropic system is domain inspecific, it should not be thought that isotropic systems are domain insensitive as well; isotropic systems may well utilise information from a variety of content domains, but only answer a very limited range of questions in which case it must be domain specific.

Closely related to a system's being isotropic is a system's being Quinean. In scientific confirmation it might be the case that the isotropic system puts up two theories that generate predictions about every relevant thing covered by the theories, but that one of the theories is better confirmed than the other on grounds of simplicity, conservatism or whatever. The point about being Quinean is that these grounds for confirmation "are properties that theories have in virtue of their relation to the whole structure of scientific beliefs taken collectively. A measure of conservatism or simplicity would be a metric over global properties of belief systems" (Fodor 1983 p. 108). Again, the theories generated by a Quinean cognitive mechanism may be in answer to a very limited range of questions even though which theory is

---

5I refer the reader to Philip Cam (1987) if s/he is interested in Fodor's argument.
accepted is determined by such global properties. There is no reason why those questions need be in any sense global.

One might want to object against VFT that it's just obvious that our higher cognitive processes answer a wide range of questions, and are therefore domain inspecific. I find this not at all obvious. What I do find obvious, though, is that our cognitive system as a whole answers a wide range of questions. That, however, tells us nothing about the mechanisms at Level Two which go to make up that cognitive system; to make that inference would be to commit a Level mistake—see chapter 1. In order for VFT to be correct, it will have to account for this apparent obviousness of our being coarse-grained representers. What the supporter of VFT will have to do, therefore, is make a move analogous to that which we saw in chapter 4. Remember that in the case of the at worst connectionism scenario, we had to retreat to Level One in order to count the system as a representer. In the present case, the adherent of VFT will have to retreat to Level One also, so that systems such as ourselves can count as coarse-grained representers.

One reason why the issue of informational encapsulation has been run together with domain specificity, is that the two might not be as independent as Fodor assumes. As we have seen, content would seem to be an informational notion. That is, after all, the presumption behind works such as Dretske's *Knowledge and the Flow of Information* (Dretske 1981). So the content domain specificity of a cognitive mechanism might well be conceptually linked to questions of encapsulation.

### 3.2.4 The Frame Problem and CSR

As we have seen above, Fodor's argument that central processes are Quineian and isotropic rests to a large extent upon an analogy with scientific confirmation. We also saw that he has independent motivation for thinking that central processes have those properties: if they are isotropic and Quineian then the frame problem is just the sort of problem we should expect to arise. Now the frame problem (and, indeed, the problem of scientific confirmation) is considered by Fodor to be a fairly intractable problem. In fact, Fodor casts doubt upon many of the approaches by AI theorists to get around the former problem (1983 pp.114-17). Both "heuristic" approaches and the employment of scripts or frames fail to get around the problem, at least, according to Fodor.

From this supposed intractability, Fodor then casts a net of gloom over the prospects for a cognitive science of central processes. He claims that all the progress made in cognitive enquiry thus far has been about the modular
input systems, and its the holistic and global nature of belief fixation which has led to a relatively dead end in the case of central processes. Consequently, Fodor characterises what he hopes will one day be called “Fodor’s First Law of the Nonexistence of Cognitive Science”. This “law” states that the more global and isotropic a cognitive process, the less we can understand it, and very isotropic processes can’t be understood at all (1983 p.107).

This gloomy conclusion might strike some readers as surprising, given Fodor’s stance on intentional and cognitive state realism. Since, central processes are where belief fixation occurs, there will be a cognitive science utilising intentional/cognitive states, only if there is a cognitive science of central processes. So, if Fodor’s pessimistic epistemological story is correct, then so much the worse for intentional realism.

Actually, it seems that either Fodor’s pessimistic scenario does not follow, or intentional realism is going to be in trouble. Consider the status of the frame problem with respect to us. The frame problem seems to be a problem only from the point of view of AI, and not cognitive psychology proper. Why? Because, if Fodor is right, although we possesses central processors which are inherently Quineian and isotropic, we do not suffer from the frame problem; we don’t suffer the sorts of problems confronting robots of AI research. Two alternatives follow from this, it seems to me. First, since nature, seemingly, has solved the frame problem in our case, then there should be, in principle, no reason why we should also not be able to solve it in the case of AI; the frame problem would not, then, be as intractable as Fodor suggests. In this case, Fodor’s pessimistic epistemological conclusion does not follow. Or secondly, if it is true, as Fodor claims, that central systems are Quineian and isotropic, and that these properties do invariably lead to intractable frame problems, and it is also true that we do not exhibit the frame or similar problems, then that must constitute evidence that we do not possess central processes. If we do not possess central processes then CSR fails. Either way, it seems to me that Fodor’s argument employing the frame problem is going to be problematic for him.

4 Empirical Arguments

In the previous section we took a brief look at some arguments which purported to show why an alternative architecture to that of cognitive state realism will not be feasible. Since cognitive state realism is an empirical thesis we now look at some empirical data to look for support for either it, or the alternative architecture I have been advocating viz. vertical faculty theory. I stress again that this section should be considered more as a
preliminary step to the disconfirmation or confirmation of cognitive state realism rather than a final word. The present work should be considered as the theoretical foundations for a substantive empirical research project. With this in mind we turn to the data.

4.1 The Neuropsychological Argument

What does the evidence presented in Appendix B tell us about the modularity of cognitive mechanisms? Most of this evidence could be used to support the existence of Fodorian modular input systems. We, however, are interested in the architectural principles which generate the higher cognitive processes underlying behaviour.

Remember that the cognitive neuropsychologist makes the following sort of inference: the impairment of some function F in isolation from other functions is evidence for that function's being performed by a module (in the brain state sense from chapter 6), and hence being a psychologically confirmed function. Now, what about the converse inference, from lack of impairment to lack of modularity? One might think that in the absence of such impairment we should seriously entertain the possibility that there is no such module. The evidence from neuropsychology would tell us that there are no coarse grained impairments; all the impairments cited are fine grained. This suggests the possibility of there not being coarse grained modules that the cognitive state realism requires to account for higher cognitive processes.

Martin Davies and Max Coltheart (Unpublished) cast doubt upon this inference since there are two possible explanations of the lack of dissociation: one is that modularity is lacking, while the other is that certain non-psychologically important factors are coming into play. Such factors might be that the brain states in question which realise the psychological functions are not modular in the brain state sense of modularity introduced above in chapter 6.

The D & C line is correct as far as it goes; lack of impairment might just mean that central processes, say, are not modular in that brain state sense. However, we are interested in the other sense of modularity in which a task is modular if it features as part of the Level Two functional analysis of the system. So, there is another stage to the neuropsychologically based inference: lack of impairment implies lack of modularity which, further, implies lack of psychologically real function. Davies and Coltheart take as their example two hypothetical tasks, A and B, and ask if they are modular or not. Now if either breaks down or continues in the absence of the other task, then we would say that they were modular tasks. However, how many actual
tasks do we want to claim are performed? If it's the case that the nonpsychologically interesting factors ensure that these tasks never get performed in isolation, then why do we want to claim that there are separate modules within our Level Two analysis of the system? It would seem that we ought to say that there is some general task being performed here, a task one might describe as the conjunction of those two subtasks. We ought do this unless we have some other reasons, independent of the neuropsychological evidence, to suggest that there are two tasks here instead of one, even though those tasks are not modular in the brain state sense. In other words, the neuropsychological evidence does not support, in this hypothetical case, the postulation of two tasks being performed, and we must look elsewhere for evidence to support the existence of our two psychologically real tasks.

Where might such alternative evidence come from? Well, from anywhere in the psychological disciplines, really. We have already seen in chapter 1 Pylyshyn's criteria for psychological reality. It will be data from areas such psychometrical studies which will provide evidence for the Level Two modularity of some psychological function. In the absence of such alternative evidence, then the lack of modularity in the brain state sense should lead one to infer lack of modularity in the functional description sense.

Memory is probably the choice example of a higher cognitive function which, seemingly, fails to count as a module in our sense, just because it would appear to be non-modular in the brain state sense. Says Gross:

No one has ever selectively destroyed memory and only memory. Even lesions in the hippocampal region, which leave a devastating effect on many different types of memories, leave memories of motor skills intact. (Corkin 1968)

To see the importance of this type of claim we can look at Stich's now famous example of Mrs T., an elderly woman who suffered a progressive loss of memory due to the degeneration of brain tissue (Stich 1983 pp. 54-6). Mrs T. presumably lost the concept of ASSASSINATION since she would continue to utter "McKinley was assassinated" but would not realise, inter alia, that this would mean that McKinley must consequently be dead. Now the usual conclusion drawn from this example is that it is unclear that Mrs T. believes that McKinley was assassinated. If we grant that she has lost the concept ASSASSINATION just what cognitive differences does that make to Mrs T., and are those differences the result of some coarse grained disruption to her cognitive system?

Now although Mrs T. has lost the concept ASSASSINATION something within her cognitive system is playing various roles in the running of
Chapter 9  Two Theories of Cognitive Architecture  204
cognitive processes. For instance, she can still utter the sentence ‘McKinley
was assassinated’. Moreover, she utters the sentence in response to questions
about McKinley. So, because she lacks the concept but can still exhibit
behaviour taken to be partly indicative of belief, then it follows that coarse
grounded cognitive states are not necessary for the production of those
behaviours.

But what does that tell us about vertical modules? There would seem
to be two options. The first is to claim that there are no central processes and
that although Mrs T. has somehow lost the concept she still possesses some
fine grounded modules that have stored information about McKinley—thus
explaining the continued ability to answer questions about him. On this
option it is possible to explain the various deficits and continued competences
of Mrs T. Strictly speaking there are no concepts featuring at Level Two
according to VFT. What Mrs T. has lost is perhaps the meaning of
‘assassination’ (from her language module’s lexicon of course!).

In claiming that according to VFT there really are no concepts at Level
Two, I am not saying that we do not possess a conceptual repertoire. The
claim is, rather, that there are no Level Two states that play the causal roles
that cognitive state realism requires. Perhaps concepts are attributed to a
system at Level One. Alternatively, perhaps concepts are just lexical items
featuring in a language module. If either of these options turns out true so
much the worse for cognitive state realism.

Even in the face of this ignorance of the mechanisms at work
according to the VFT programme, things get worse if one opts for the
following second option. Assume that there is a central processor at work
which accesses concepts. If Mrs T. has lost that concept then it is pretty
difficult to see how she can believe that McKinley was assassinated. Okay;
but if she didn’t believe that McKinley was assassinated it’s also difficult to
see why she would respond to questions about McKinley in the manner she
does given the loss of the concept. It would be no good for the believer in
central processes to claim that Mrs T.’s continued capacities resulted from,
say, the continuance of some sentence token in a belief register; since it is
causal role that is constitutive of something’s being a belief, and it is exactly
the causal role that has broken down in Mrs T.’s case, by hypothesis, then it
does not make sense to postulate a mental sentence. So, on this option we are
going to be at a loss to explain Mrs T.’s behaviour.

The Mrs T. example was used by Stich to bring out the vagaries
associated with the attribution of content to mental states. I think the

---
6Well, I really don’t have the faintest idea what Mrs T. has lost, since the range of
cognitive mechanisms postulated by VFT is unknown.
example shows us something about causal roles. What Mrs T. shows us, like the rest of the neuropsychological data, is that breakdowns fractionate the mind at a fine level of grain. There is obviously some state of Mrs T. that plays some causal role in her answering of the interlocutor's questions. Given that this causal role is a partial constituent of beliefhood, the state in question would have to be a partial belief of some sort. I'm not sure of what such a partial belief could be—a mental sentence of the form 'McKinley was #$%^*!'. That would still not explain, though, why Mrs T. gave the answer that she did. We can have degrees of belief that form the basis of our actions according to one's favourite decision theory, but that cannot be the sense of partial belief required here. In order to postulate coarse grained Level Two cognitive mechanisms a cognitive state theorist will have to account for such cognitive anomalies as well as the normalities—this being the claim the Churchlands have been making for years—by given some account of the etiology of Mrs T.'s response. I don't see where such an account is going to come from on the cognitive state realist story.

What the neuropsychological evidence and Mrs T. tell us is that really there are a number of smaller tasks that are performed, and most importantly that due to dissociations and breakdowns, the modularity of many of those tasks can be inferred. Since those modules are fine grained mechanisms, we get support for the VFT over cognitive state realism. Now this seems to be the state of play as seen from the neuropsychological literature only. The argument as just presented depends crucially on there not being other independent reason for postulated coarse grained mechanisms. Granting that my knowledge of the empirical data leaves quite a deal to be desired, I haven't seen any data that should lead us to confirm the existence of coarse grained mechanisms, which is necessary in order to confirm cognitive states realism. That might be surprising to some given some well known experimental data, which we should now examine. I don't think, however, that the data shows what it is purported to show.

4.2 Propositional Representation

It is without doubt that from the perspective of Level One, our cognitive processes seem to range over coarse grained representations in the way that cognitive state realism (and intentional realism) requires. This is most easily seen in our use of language: in our production and understanding of linguistic items we manipulate coarse grained information. One way to interpret this capacity which is consistent with cognitive state realism is to claim that rather than these coarse grained representations featuring within the
language module, say, there is another representational system which not only translates into a natural language but which also underlies cognitively mediated behaviour, with linguistic utterances being examples of such behaviour. Such propositional representations as they have been called (Stillings et al. 1987 p. 21) can be construed as cognitive states. The question is: what kind of evidence is adduced in support of the existence of such states?

4.2.1 Introspective Arguments

One style of argument is based upon the introspected distinction between having an idea and putting that idea into words. We often, so the story goes, have a concept "clearly in mind" but cannot retrieve the word that stands for the concept. This is the so-called "tip-of-the-tongue" phenomenon, and is supposed to indicate that concepts are not identical with linguistic items. Similarly, there is also the problem of giving the definition of a some natural language word which stands for a concept. Very often we find that the giving of the definition of a word takes a considerable amount of processing and mental effort. This is supposed to suggest that the definitions of those words are expressed in a separate internal representational medium which gets associated with the word (Stillings et al. p. 22).

It seems, though, that the former example's working depends upon how we might possibly "have a concept clearly in mind" without being able to retrieve the word for the concept. Maybe one cannot have a concept in mind in the way required without being able to come up with the word. Such euphemistic expressions, which often feature in such anecdotal evidence, don't do much to clear the already muddied cognitive waters. Of course, a proper cognitive theory is going to have to account for such introspected experiences. If one assumes that words cannot be concepts, then this problem can be explained by that fact. However, if words are concepts, then one can equally explain the problem by reference to occasional difficulties in accessing lexical items by the language module, say. Why there are such occasional difficulties we would hope to eventually explain.

Far from showing that the definitions associated with linguistic items are expressed in a separate propositional system, the second example can be easily explained by questioning whether there is a strict definition associated with words. If there are not, then it will take some cognitive processing to come up with the criteria of application for the word. Perhaps we are not very good at giving definitions because we have only a rough prototype to draw from and when asked to give an actual definition we try to flesh out the prototype as exhaustively as we can. The point is that the processing that
seems to be required might well be within the linguistic module, and not consist in the translation from a propositional representational system into the language system.

Another reason offered for postulating a propositional representational system separate from the language system is that the existence of a propositional network putatively explains how children can begin acquiring concepts even though they were not "born knowing a natural language" (Stillings et al. p. 22). As stated, this reason will fail because it might be that being born with a natural language is irrelevant, just so long as one acquires concepts when acquiring that language. That would count against there having to be separate coarse grained representational network. The response to this is to claim that children can learn concepts that they are unable to explain until they have been talking for quite some time. Again, this does not show that concepts are formed independently of language, but just that in the acquisition of concepts we do not acquire strict definitions of those concepts so that we can explain what the concepts are.

4.2.2 A Recognition Memory Test: The Retention of Meaning

All of us at some stage have had the experience of remembering the overall point of something said to us, while not retaining the actual words used to communicate that point. It is argued that this phenomenon supports the view that there is an underlying propositional representational system independent of language. There is, supposedly, experimental evidence in support of this anecdotal evidence. Here is a description of a classic experiment by Sachs (1967).

In the experiment subjects listened to paragraphs on various topics, such as Galileo's work on the telescope. The passages were interrupted, and subjects were given a recognition memory test for a sentence that had occurred 0, 40, or 80 syllables earlier in the passage. In the test the subject was presented with a sentence and asked to judge whether it was the same as a sentence in the passage or changed. For the Galileo passage the subject could be presented with any one of the following four sentences. The actual sentence was sentence 1.

1. He sent a letter about it to Galileo, the great Italian scientist.
2. He sent Galileo, the great Italian scientist, a letter about it.
3. A letter about it was sent to Galileo, the great Italian scientist.
4. Galileo, the great Italian scientist, sent him a letter about it.

Notice that sentences 2 and 3 have the same meaning as sentence 1, although they have different grammatical forms. Sentence 4 has a different meaning. At a delay of 80 syllables (20-25 seconds) subjects rarely made the error of thinking that sentence 4 has occurred in the passage. However, they frequently and mistakenly thought they remembered sentence 2 or sentence 3. Thus, under the conditions of Sach's experiment subjects were able to remember the meaning of a sentence without remembering its linguistic form. A reasonable explanation of the finding is that the meaning was represented in an underlying nonlinguistic form that could be translated into a variety of linguistic paraphrases, which subjects confused in the recognition test (Stillings 1987 pp. 22-3)

This case certainly shows that certain components in the analysis of linguistic items, syntactic and lexical details, can be lost very quickly. However, what I fail to see is that it shows us anything about there being some nonlinguistic propositional representational system. I think it's fairly uncontroversial that in this case there has been an abstraction away from certain linguistic properties of an utterance; that should not be surprising given the distinguishing marks of cognitive systems put forward in this work. However, there does not seem to be any requirement that the abstract entities which somehow carry the meaning of the utterances be nonlinguistic. Fodor has called these meaning bearing entities “interlevel representations” (Fodor 1983 pp. 55-60), and they are representations which may be internal to the language module. In that way, they will not be nonlinguistic. Whatever the mechanisms are which mediate the processes of verbal comprehension they will have to deliver the meaning of linguistic utterances, where that meaning abstracts away from the details of syntax in just this way. Perhaps the claim being made by the proponents of propositional representation is that just as soon as the particular details of syntax are lost then the system has gone propositional. If that's so, then there cannot be an argument here for cognitive state realism. That's because the representations which are now described as propositional feature within the language module, and that means that they will not be coarse grained enough to qualify as cognitive states in the sense of the act.

---

7There must be a huge caveat hanging over this example to the effect that very often such details are not lost in comprehension. The passing of dictation tests at school and the existence of stenographers points to this fact.
4.2.3 Scripts

The trouble with the arguments presented so far is that, even if they do show that there is a propositional representational network, and so there are coarse grained representations, that still does not provide evidence for cognitive state realism's being true, since those representations might not possess the causal roles that are partly constitutive of cognitive states. We now turn to arguments from cognitive psychology which purport to show the existence of coarse grained representations which play the kinds of causal roles outlined in chapter 5. We turn then to questions of scripts.

A script is hypothesised to be “a declarative knowledge structure that captures general information about a routine series of events or a recurrent type of social event, such eating in a restaurant or visiting the doctor” (Stillings et al. p. 31). They are thought to be constituted by such things as names or themes (visiting the doctor), typical roles (doctor, nurses and patient) entry conditions (feeling ill) and goal directed scenes (checking in the receptionist, waiting, undressing).

The idea behind scripts seems to be this. We act out of our knowledge about the world, given certain desires we possess. Now scripts are supposed to provide the representational states which interact with those desires to bring about action. So, it might be scripts which will contain some of the cognitive states which will play the appropriate causal roles. Two questions are raised by the literature on scripts: are there scripts? and if so, are they structures composed of cognitive states?

As to the first question, the evidence for the existence of scripts is especially weak. The main argument for their existence runs something like this: a knowledge structure such as scripts are postulated. One then hypothesises certain behavioural consequences of their presence that should show up under experimental conditions. It then turns out that the behavioural consequences occur, and that constitutes evidence for the scripts. The weakness in the argument is that very often that behavioural consequence might be explicable by some other feature of cognitive processing.

A classic experiment used to support the existence of scripts was performed by Bower, Black and Turner (1979). They hypothesised that subjects would confuse material they had read in the experiment with material they had inadvertently filled in from their script knowledge. Their prediction was “borne out”. The trouble with such experiments is that the evidence vastly underdetermines the conclusions drawn. Perhaps in the face of pressure to recall accurately some story which they do not totally
remembering may well incorrectly substitute information from one's personal memory store either consciously or unconsciously. However, such a phenomenon is consistent with there being no scripts at all.

The second question regarding whether the existence of scripts constitutes vindication of cognitive state realism is most important. One problem with the script experimental paradigm is that the evidence regarding the roles that they are supposed to play in the causation of behaviour is grossly unclear. The script literature properly considered really has to do with complex concepts such as eating at a restaurant or visiting a doctor. The point in common is that these complex concepts happen to be activities which agents perform. There are many such concepts relevant to human action: marching, swimming and riding a bicycle. One may then think of scripts as the specification of a conceptual prototype. The trouble with this interpretation is that it gets us back to the problem of linguistic versus nonlinguistic knowledge representation. Scripts so construed might still be a form of linguistic representation, which would mean that it will fail to fill the types of causal roles required by cognitive state realism. To an extent, this problem is foreshadowed by Bower, Black and Turner when they claim that there is split between enacting and verbalising scripts in the collection of the script data (1979 p. 214).

Bower, Black and Turner also worry that the data collected about scripts through report and recall methods will give access to script knowledge that is consciously introspectable (p. 214). I'm not sure just what kind of knowledge unconscious nonintrospectable declarative knowledge would be like. Whatever it is like, it must be shown to get into the causing of behaviour in the way cognitive state realism requires. There is bound to be a great deal of unconscious and nonintrospectable knowledge about the concepts associated with human activities. Consider riding a bicycle. We list the requisite actions one must perform in order to successfully ride a bicycle. What one cannot do is to relay the information or declarative knowledge that would actually gain a listener the relevant skills. There's a distinction philosophers have long made between knowing how and knowing that. The information encoded in scripts would certainly seem to be knowledge that. Once one begins to go unconscious and nonintrospectable one would seem to be moving to knowledge how. Certainly knowledge how gets the kinds of causal roles required by cognitive state realism into our picture. However, we now have the wrong kind of representational information filling the causal role: it's skills encoded informationally and not the matters of fact about states of affairs in the world which cognitive state realism wants. The moral seems to be that you can have your scripts with all the propositional
information regarding human activities you like, and yet still not have that information playing a role in the causation of behaviour; some aspects associated with visiting a doctor might well just be like riding a bike.

At this stage there appear to be no arguments in support of cognitive state realism from script theory within cognitive psychology. At the moment, we are offered mere metaphor in place of theory.

Final Signpost

It is indeed an understatement to describe this chapter as a tentative start to determining what would have to be empirically the case in order for cognitive state realism to be judged true or false. To that extent the major component of this work is considered to be the theoretical foundations of the earlier chapters which allow one to pursue this empirical line of enquiry as a means to evaluating the thesis of intentional realism. There is a vast literature in the fields of neuroscience, neuropsychology, cognitive psychology and social psychology which may be surveyed in order to determine the correctness of the cognitive state realist thesis. The completion of that task awaits us in the future. Then, perhaps, the true topography of intentionality will be evident to all.

One trouble with playing the "only game in town" is that it can lead one to ignore underground swells of illegal games. One does not need to be a political or legal theorist to realise that what is considered legal often changes over time. I think cognitive legislators are going to pass some bills to legalise some other game in this town.