Chapter 8

On Folk Psychology and Mental Representation

Peter Godfrey-Smith

Introduction

In the 1980s the problem of giving a naturalistic theory of mental content beckoned young philosophers like myself; this looked like a philosophical problem that was both fundamental and solvable. The aim was to give a theory of the physical or biological properties of the internal states of organisms that suffice to make these states representations of the world beyond them. Internal states that have these special physical or biological properties have semantic content, and they do so as a matter of objective fact. The folk-psychological concepts of belief and desire were seen as picking out, in a rough and imperfect way, both the kinds of inner states and the kinds of semantic properties that would figure in the more detailed naturalistic theory.

A theory of mental representation of this kind would not only become a centerpiece of cognitive science, but would be essential to epistemology and many other areas of philosophy as well. It would be important to all parts of philosophy that must use or assume a theory of thought. The guiding ideas for this project derived largely from the work of Jerry Fodor and Fred Dretske. Daniel Dennett and Stephen Stich looked on as skeptical but constructive critics (see Rey 1997; Sterelny 1990; Stich & Warfield 1994 for reviews).

Roughly twenty years on, how has the project fared? With some sadness and much caution, I suggest that things have not gone well for the Dretske-Fodor program. I doubt that we will ever see a satisfactory version of the kind of theory that

Representation in Mind
New Approaches to Mental Representation

ISBN: 0-08-044394-X

Fodor's *Psychosemantics* (1987) and Dretske's *Explaining Behavior* (1988) tried to develop. Despite this, I do think we have learned a lot from the development of this literature. Some good partial answers may have been given to important questions — but not the exact questions that Dretske and Fodor were trying to answer.

So I think it is time to start looking at different approaches to the network of questions surrounding belief and representation. This rethinking will involve looking again at some of the ideas of the nay-sayers of the 1980s, like Dennett and Stich, but looking further afield as well.

In this paper I will begin to sketch one alternative way forward. "Begin" and "sketch" are the right words; this paper will not give anything like a complete theory. It will not even tackle the central problem of saying how the ascription of content actually works. Instead, I will cautiously outline some ideas that might be pieces of a future theory.

Two Sets of Facts

We can start in familiar territory by recalling one of the fundamental disagreements that structured much of the discussion in the 1980s.

Dretske, Fodor and many others think that some organisms, including people, contain inner states and structures which represent the world, and do so as a matter of objective fact. Both ordinary people and theoreticians of various kinds engage in the interpretation of these organisms; we attribute beliefs, desires, and other less commonsensical intentional states to them. When we do this — both we the folk and we the theoreticians — we are trying to describe real features of how agents are wired and how they are connected to the world. Interpretation is based largely on behavioral evidence, but it is an attempt to describe what is really going on inside.

Other philosophers, including Dennett, think that this picture misconstrues the practice of interpretation. (Davidson's views overlap with Dennett's on some of these issues, but I will focus on Dennett here.) For Dennett (1978, 1981), to make an interpretation using folk psychological concepts is *not* to posit definite causally salient structures inside the head, which have special semantic connections to states of affairs in the environment. For an agent to have a particular belief is merely for the attribution of this belief to be compelling to an interpreter, where an interpreter has a characteristic viewpoint and a special set of goals. Interpretation is holistic, behaviorist and rationalizing.

From the point of view of someone like Dretske or Fodor, this "interpretation-based" view seems to have the tail wagging the dog. But for Dennett the dog exists only as a projection made by the tail, and to say anything different is to buy into the old view of the mind as a kind of "ghost inside the machine."

The dispute has proven hard to resolve. But one thing we can do is step back and say this: *however* the details go, we will in the end have to account for two sets of facts:

(1) Facts about the wiring and organization of behaviorally complex organisms, and the connections between their wiring and the world around them. (I will call these "wiring-and-connection facts.")
(2) Facts about our actual practices of interpretation and ascription of content. (I will call these "interpretation facts.")

Here I described these sets of facts *neutrally* with respect to the debates between people like Fodor and Dennett. Whether the interpretations made by people are *descriptions of* the wiring-and-connection facts or not, the world does contain these two sets of facts. Both are empirical phenomena, and in principle there could be complete empirical theories of each (see also Stich 1992).

So imagine a future state of scientific knowledge in which we have highly detailed empirical theories of people. One thing this body of empirical knowledge will contain is a description of these two sets of facts. But as well as these two bodies of empirical knowledge, we will want a theory of how the two sets of facts are *connected* to each other. Here we find one of the roles for philosophy — to describe the coordination between the facts about interpretations and the facts about wirings-and-connections.

The same sort of thing is true in other parts of philosophy as well. We can think of the goal of philosophical theories of causation, and theories of knowledge, in the same kind of way. Philosophy should aim to describe the connections between facts about the *use* of difficult and controversial concepts, and facts about the parts of the world that the concepts are in some sense aimed at dealing with. Philosophy should link the empirical facts about human causation-ascriptions with the empirical facts about how the world runs. Philosophy should link the empirical facts about human knowledge-ascriptions with the facts about how beliefs are regulated and how beliefs enable organisms to get around the world. In describing these connections, it is natural and appropriate for philosophy to speculate about how the empirical stories on either side will turn out. We can speculate but also be guided by the empirical information as it comes in. Although this is not the only role for philosophy in these areas, it is one central role.

Let us return to the philosophy of mind. Imagine that we have complete descriptions of both the wiring-and-connection facts and the interpretation facts. How will familiar theories of mental representation in the literature look in relation to those descriptions?

The Dretske-Fodor program will look like an assertion of one way the two sets of facts might be tied together. For the Dretske-Fodor program, folk psychological interpretations are controlled by data which contain information about the wiring-and-connection facts. That is, the wiring-and-connection facts generate behavioral data which in turn generate interpretations. These interpretations function as attempts to accurately describe the wiring-and-connection facts. Some interpretations are false, of course, like any claims made about hidden things. We can be data-driven and also be wrong. But for people like Dretske and Fodor, interpretations are often fairly accurate, and this accuracy is a matter of accuracy about wiring-and-connection facts.

You might object at this point: how could interpretations given by people in earlier ages with no knowledge of the kinds of "wiring" we have in our heads possibly be trying to describe "wiring-and-connection facts?" This objection focuses too hard on my term "wiring." Any kind of inner structure that figures systematically in the causation of behavior will count for present purposes.

The Churchlands' program shows us another way in which the two sets of facts might be related. Folk psychological interpretations are intended to accurately describe wiring-and-connection facts, but these interpretations embody a false theory and fail in systematic ways (P. M. Churchland 1981; P. S. Churchland 1986).

Pl. check. We have changrd P. M. Churchland (1980) to P. M. Churchland (1981) according to the reference list.

Dennett's view shows us yet another option. On this alternative view, it is an error to think that the role of interpretations is to try to describe wiring-and-connection facts. Belief ascriptions are not attempts to pinpoint discrete, causally active, internal states with special semantic properties. Belief-attributions are not like gene-attributions. Instead, we should give a theory of the social role of interpretations that does not treat them as representationally aimed at the wiring-and-connection facts. We should think of belief ascriptions and other interpretations as part of a practice that has to do with various kinds of social coordination. Folk psychology is a "craft" (Dennett 1991). Inner wirings and physical connections between internal and external do exist, but it is a mistake to think that folk psychological interpretations function as attempts to make specific claims about wirings and connections.

From the point of view of Dennett's picture, the Dretske-Fodor view is mistaken because it does not take seriously the special properties of human interpretive practices; interpretation is not just an attempt to lay out the hidden structure of a complex machine. From the point of view of the Dretske-Fodor picture, Dennett's view seems to deny that understanding the mind must ultimately be understanding what is going on inside the skin. That is something that ordinary people know, and something psychology knows whenever it is not diverted by bad theory.

In the next section I will try to make some progress on these oppositions. Before leaving the methodological discussion I should note some idealizations

I have made. I distinguished two sets of facts, wiring-and-connection facts and interpretation facts, and imagined our gaining good theories of both of these. The relation between the structure of the mind and the structure of our interpretive practices might well be co-evolutionary (see Godfrey-Smith 2002). But beyond that, the relation between both of these sets of facts and our *theories* of these facts can be co-evolutionary. Highly theoretical ideas about the mind can filter down into everyday practices of interpretation. A recent example is found in Freud's ideas about unconscious desires, which have filtered quite far into everyday interpretation. Older examples might be furnished by religious ideas about the soul, and literary ideas about romantic love. Mental states, interpretations of mental states, and *theories* of both of these are all interlocked. The extent and importance of the interlocking depends on unresolved questions about modularity, plasticity and cultural change.

On Interpretation

After all the years of debate, how can we start to resolve the disagreements outlined in the previous section? The first point that should be clear is that we need more empirical knowledge about folk psychological interpretation. Fortunately, new empirical input is coming in. In particular, there is a growing body of work within psychology on the development of folk psychology in children (Davies & Stone 1995; Stone & Davies 1996). And work within cognitive science on the "wiring and connection" side continues as usual.

We can expect the oppositions outlined above to be transformed as this empirical work develops. In the meantime though, I would like to outline some simple ways in which we might find some of the disputes resolved.

Consider one very basic feature of the disagreements described earlier: is it the case that folk psychological interpretations make commitments or hypotheses about the internal structure of the person being interpreted?

Surely it is remarkable — and of some sociological interest — that there is still *so much disagreement* about this simple question. Dennett has argued for many years that folk psychology does not commit its users to anything about the internal causal structure of the people being interpreted. Dennett accepts a kind of "minimal logical behaviorism." Nothing about the insides of a person could affect the truth or falsity of an interpretation of that person in folk psychological terms.

So says Dennett. But to others this claim seems nothing less than outrageous. People who differ massively on other matters — Fodor, Stich, Armstrong, and the Churchlands — at least agree that folk psychology is committed to some claims

about what is going on inside people's heads. Surely beliefs and the like are, at least sometimes, internal mental causes of what people do.

Once we get past this fundamental opposition, the disagreements among those with a more "realist" interpretation of folk psychology are themselves substantial. Some, like Fodor and Stich, think that folk psychology involves rather strong claims about the insides of our heads, strong enough for eliminativism to at least be a relevant possibility. Others, like Jackson & Pettit (1990), think that folk psychology makes claims so weak that eliminativism is an almost impossible option.

Let us focus on the fundamental dispute about whether folk psychology does or does not involve commitments about internal structure. How should we deal with the alarming inability to agree on so basic an issue? Is one side so steeped in ideology that it cannot accept the blindingly obvious?

I suggest that some aspects of this position might be resolved with a simple assertion of false dichotomy. Folk psychology need not have a single role, with the respect to the question of whether interpretations are directed at describing inner structure. What we might need is some kind of pluralism on the issue. Folk psychological interpretation is a tool that lends itself to several different intellectual and practical tasks.

With some kind of pluralist view in hand, we can explain many of the strange battles of the past by noting that different philosophers have focused on different aspects of folk psychological practice. Compare the use of folk psychology in a law court, when the aim is to work out exactly what crime a person committed, and the use of folk psychology in freeway traffic, when the aim is to avoid collision and coordinate one's actions with surrounding drivers. These two uses of folk psychology *look* different and perhaps they *are* different. When one focuses on the law court case, it seems transparently clear that the apparatus of interpretation is being used to explore hypotheses about the inner causes of behavior. When one focuses on the freeway traffic case, the postulation of inner structure seems beside the point. All that is relevant in that case is the fast, accurate prediction of behavior. Philosophical theories devised under the influence of each of these two paradigm cases will look very different.

So we may need some kind of pluralist option about folk psychology — about the intended relation between wiring-and-connection facts and interpretation facts. But there is a variety of ways in which this pluralist view might be developed. Here I will sketch a couple of options, without taking a firm stand on any side.

One thing we have to work out is whether folk psychology contains multiple *practices*, or a single practice with something like multiple *construals*. Deciding this issue requires that we work out where folk psychology itself stops, and application of and commentary on it begins. Dennett (1991) has argued that we should distinguish the "craft" from the "ideology" of folk psychology. The "craft"

is the unreflective use of the apparatus; the "ideology" is a set of ideas *about* the craft of folk psychology. The distinction I am making here is similar, but not the same. The "ideology" of folk psychology, for Dennett, is something that has no role (or very little role) in the ordinary social use of the apparatus. It is purely a piece of theory used to comment on the craft. The distinction I am making here is one between basic features of the craft, and more elaborate comments on and applications of the craft that *also have a role* within folk interpretation itself. So if we want to develop a pluralist option for folk psychology, one question we have to answer is whether there is a diversity of crafts, or a diversity of construals and applications, or both.

I will say something about each possibility. One simple way to resolve the old debates would be to claim that folk psychology is bifurcated all the way down to the most basic features of the practice. It might be argued that people switch between two different modes of interpretation, usually without realizing it. On a freeway, people apply a form of interpretation which is either behaviorist or very close to it. But when working out in a legal setting whether someone is guilty of murder or manslaughter, a different kind of interpretation is used. Then everything hinges on choosing between specific rival hypotheses about inner mental causes.

I pause to contrast this strategy with another recent proposal aimed at resolving some of these issues. Jackson & Pettit (1990) argue that folk psychology (as a whole) does make some commitments about internal structure, but these commitments are extremely minimal. So their reply to Dennett is that there is no real "instrumentalist" option for folk psychology because the right "realist" construal is so uncontentious. For Jackson and Pettit the *same* minimal claims are made when one uses folk psychology in the law court and on the highway. I suggest that their minimal realist construal is unlikely to make much sense of the most causally specific kinds of folk-psychological interpretation — the kind found in courts of law, soap operas, and so on. But more importantly, once we recognize the possibility of a pluralist view about folk psychology, there is no *need* to search for a single account that simultaneously makes sense of the striking obviousness of interpretations made in some contexts (on the highway), and for the apparent logical strength and specificity of interpretations made in other contexts (law courts). Different practices are at work in each case.

How likely is this first possibility to be right? One obvious problem is the fact that it seems strange for there to be two separate interpretive practices with so much in common. The same basic concepts are used in both kinds of interpretation (otherwise it would be much more obvious that there are two practices). It would be interesting to investigate whether some kinds of propositional attitude ascription only appear in one context or the other — whether some folk psychological

concepts are *only* useful in the service of interpretations aimed at uncovering mental causes. That would be evidence for this first pluralist option. More pressingly though, the psychologist Alison Gopnik tells me that there is no evidence in developmental psychology at all that seems to support this first option — especially, no evidence supporting the idea that children learn different practices of interpretation including a near-behaviorist practice for use when prediction is all that matters.

So I will now sketch a different, somewhat more complicated, version of the pluralist view. Suppose we think of folk psychology as a *model* of the mind, rather than (as it is often put) a *theory*. What is the difference? The difference that I have in mind has to do with the fact that a model, in science, can be interpreted as having many different kinds of relationship to the phenomena it is directed upon. In particular, there is a continuum of possible attitudes with respect to which features of a model are interpreted in a realistic way — as having real and distinct counterparts in the underlying structure of the world. Two scientists can use the same model while disagreeing about which features of a model are taken to have real counterparts, and a single scientist can use the same model while changing his or her mind about the right interpretation. A model is a conceptual structure which lends itself to a variety of different scientific uses. Sometimes predictive adequacy is central, but sometimes more than this is hoped for or required. Sometimes a model might be favored despite unimpressive predictive power, if it seems to get something about the basic causal structure right.

Folk psychology might be considered in something like the same way. The basic folk psychological model of the mind itself is normally acquired at a very young age, as the empirical literature has shown. What a child acquires by the time he or she is five years old might be *facility with a model*. This can be acquired without a grasp of some distinctions between different construals and applications of the model. Children might pick up different ways of using and construing the model as they grow older — they pick up different ways of highlighting some aspects of the model and downplaying others. A normal adult does not care about the inner structure of the driver alongside on the freeway; only behavioral patterns matter. The folk psychological model is then *used as* an input-output device. But in other social contexts the model will be used to try to get a grip on the exact nature of mental states which act as intervening variables in the causation of behavior. Behavior now functions only as evidence for hypotheses about inner causes. The aim is to elucidate the fine structure of mental processes ("did she realize the likely effect of those words?", "did he really fear for his life at that point?"). In some social contexts, the details of the underlying mental processes, as far as we can

discern them, do make a difference to how we treat a person or problem, while in other contexts they do not.

So on this view there is a unitary, low-level "craft" of folk psychological interpretation. This consists in the ability to apply a model. But a model itself does not determine its proper interpretation; it can be construed either as a set of hypotheses about hidden causes, or as a purely predictive device. It can also function as a mixture of these — some core features of the model might be taken to have real counterparts, while most details of the model are not interpreted this way. My suggestion is that the ability to use the model in a variety of different ways is itself part of the set of skills that a folk-psychological interpreter comes eventually to acquire. The distinction between different construals or applications of the folk-psychological model is not something peculiar to philosophical discussion, but is part of the tool-kit ordinary people use to negotiate different contexts in social life.

I am unsure how best to further develop these ideas, and how best to connect them with empirical work on the structure, evolution and ontogeny of the folk psychological model of the mind. If the second pluralist proposal is correct, we should be able to empirically distinguish two different aspects of the acquisition of folk psychology — the acquisition of facility with the model, and the acquisition of alternative construals and applications. In correspondence, Alison Gopnik suggested to me that the ease with which young children handle the concept of *pretending* might tell against this view. Children do have a grasp of the distinction between literal and non-literal applications of all kinds of frameworks. And their take on mental states is that these states are real. If this is right, then the acquisition of folk psychology takes place in a way that is accompanied by a simple realist construal of the model. Any pluralism must then result from later revision.

Clearly these are subtle empirical issues, and my discussion here has the nature of speculation and sketching. I do think the possibility of a pluralist view of the commitments of folk psychology is important though. It suggests that there is no need to "thread the needle"; no need to give a single account of the commitments made by folk psychology that does simultaneous justice to the role of folk psychology in freeway traffic, soap operas, law courts, historical reconstruction of the causes of World War I, and devising marketing plans for new products. This needle-threading project may well have been, to use an Australian expression, a mug's game.

It is possible that the pluralist option might help with some other questions that philosophers battle endlessly about in this area. An example is the problem of animal belief.

On the Representational Concepts Used Within Cognitive Science

Suppose the ideas sketched in the previous sections are right. What would this tell us about the role of folk psychological concepts in cognitive science, and the issue of eliminativism? In this section I will discuss that issue, plus another linked set of issues concerning the naturalistic relations between thought and the world that have figured in philosophical attempts to give a reductive analysis of semantic properties.

If folk psychology gives us a model that can be applied to mental processes, then folk psychology can be associated with either a strong *or* a weak set of hypotheses about the mind. We can look for a detailed mapping between the structure of the model and the structure of the mind, or a minimal mapping. Folk psychology itself does not resolve the issue, does tell us which are *the* issues that are crucial for assessing whether beliefs (for example) are real or not. If the folk psychological model was both predictively unimpressive and turned out to structurally very unlike real mental processing, then the answer would be clear. Eliminativism would be vindicated. If folk psychology was predictively impeccable and the mind turned out to work the way Jerry Fodor has claimed it does — with a set of core sentence-like structures representing the world and being manipulated computationally during thought — then eliminativism would be clearly false. But the more likely outcomes lie in the middle. Folk psychology is clearly very predictively powerful in many of its domains of normal use. As a picture of mental processing, we do not know what will come of folk psychology yet. But let us look at cases where the two extreme options are false, and some pieces of folk psychology remain useful.

In science, a model can be used or rejected as a whole, but it can also be mined, piecemeal, for elements that can be used in new models. This mining is what we often find in the case of folk psychology. When a concept like *belief* appears in a discussion in cognitive science, what has often happened is that the cognitive scientist is taking the folk psychological picture of mental processing and "stripping it down" to yield a few core structural elements that might have scientific value. For example, a core feature of the folk psychological idea of a belief is the *contrast* between beliefs on the one hand and wants and desires on the other. A theory about the mind that is in many ways at variance with folk psychology might retain a basic contrast between "how things are" registrations of the world and "what I want" registrations of the world. It might also retain the idea of rational or well-adapted cognitive processes as involving a systematic interaction between the two.

Imagine some theory of this kind — a theory recognizing the core structural contrast between belief and desire, but embedding this contrast within a detailed

picture that departs in many ways from folk psychology. (This might be a radically connectionist theory, a theory borrowing from dynamical systems theory in the style of Tim Van Gelder, or a model of distributed cognition in the style of Rodney Brooks.) And suppose the theory is successful and developed in detail. Some cognitive scientists will want to retain folk-psychological terms like "belief" for the states that are posited and described in this process; others may want to avoid these terms. To retain the term "belief" is to stress the continuities between the scientific model and the folk psychological model, with respect to the basic interaction between "how things are" and "what I want" states. To drop the term is to stress the discontinuities between the scientific and the folk psychological picture. But there is no fact of the matter about whether the psychological states that appear in such a psychology "really are beliefs" — whether they are the same states posited by folk psychology but more accurately described (as envisaged in Lycan 1988). There is no fact of the matter because folk psychology itself does not commit to a sufficiently definite specification of what beliefs are supposed to be like. However, it will be appropriate to conclude in such a case that there is *some* non-trivial coordination between the folk-psychological model and the scientific theory.

Here is an example to illustrate these general points. Ramsey *et al.* (1991) claim that propositional attitudes like beliefs are conceived by folk psychology as "functionally discrete, semantically evaluable states" that play a causal role in the production of behavior and other propositional attitudes. The key idea here is "functional discreteness" — beliefs are seen as individually revisable and deployable. They can be added and lost individually, in virtue of reasonably localized alterations to the cognitive system. Ramsey *et al.* claim that connectionist models of a certain kind do not treat inner representations as having these features. So if connectionist models of this kind are accurate models of cognition, they support eliminativism about belief. This is a good example of people taking the folk psychological model of the mind, and insisting on a *very* realistic construal of the structure of the model. The structural match between the folk psychological model and the real nature of cognition has to be very good, or folk psychology has been undermined. But folk psychology itself does not contain any commitment to these "discreteness" properties. When mental processing is modeled by folk psychology, beliefs appear in the model in a way that *suggests* this kind of discreteness, but *only* if the model is interpreted in such a way that nearly everything in it is supposed to have a real-world counterpart. And this is not the only, or the most natural, interpretation of the model; there is no one right way to construe the model. So it is a mistake to see folk psychology as definitively committed to the falsity of the class of models Ramsey *et al.* discuss. But it is true that if these models were accurate, that would significantly *reduce* the number

Pl. check. We have changrd Ramsey et al. (1990) to Ramsey et al. (1991) according to the reference list.

of elements of the folk psychological model that can be taken to have real-world counterparts.

So far I have mostly discussed folk psychology's handling of inner structure. There is another side to these issues as well, and this has to do with connections between internal structures and the world. To finish the paper I will make some suggestions about this second topic that follow up some of the same themes as the earlier discussion.

During the 1980s we saw many attempts to show a coordination between some set of naturalistic relations between internal and external states, and the one hand, and folk psychological semantic relations, on the other. The aim was to use this coordination to *reduce* semantic properties to more basic, physical properties. I doubt that such a theory can ever work (again, I say this with caution). I will not hazard a guess about which of the existing alternatives to the reductive approach is most close to being right. Instead I will discuss a related topic. Although the reductive project failed, the philosophers pursuing it did succeed in describing *some* interesting kinds of natural connection between thought and the world. What should we make of these naturalistic connections? And what should we make of the use of representational concepts in cognitive science that have the same naturalistic orientation?

To approach the issue, we can begin by looking again a passage from a key early discussion of these issues. Here is the passage Dretske used to begin his book *Knowledge and the Flow of Information* (1981):

> In the beginning there was information. The word came later. The transition was achieved by the development of organisms with the capacity for selectively exploiting this information in order to survive and perpetuate their kind (p. vii).

Properly understood, this is a good summary of some fundamental features of the physical involvement that organisms have with their environments. Information, in the sense analyzed by Dretske, is a *resource* that can be utilized by organisms, like the carbohydrate molecules in fruits and seeds. The evolution of sensory and cognitive mechanisms is the evolution of ways of making use of this resource. But the relationship between information in Dretske's naturalistic sense and the semantic properties posited in folk psychological interpretation is more complicated than Dretske and others envisaged. Informational properties in the Dretske sense cannot be used to give a "reduction" of folk-theoretic semantic involvement between internal states and external conditions. What Dretske has done instead is isolate a real and important kind of naturalistic relation between organisms and environments, a relation which has a *partial*

similarity to the kinds of semantic relations that figure in folk psychological interpretation.

In that initial statement I focused on Dretske-style informational links, which are in some ways fundamental to a theory of organism/environment relations. But similar points apply to a variety of other naturalistic relations described in the recent literature. In particular, something similar is true of the properties involving biological functions and mapping relations that Millikan has described with such care (1984). These are real relations between internal and external, which probably have a significant role in cognitive science. But they can have this role without being the material for a reductive theory of folk-psychological semantic properties. (I should add that although I would make these claims about both Dretske's informational relations and Millikan's teleo-functional relations, I do not think this is true of Fodor's relations of asymmetric dependence (1987), as these are not fully naturalistic relationships at all.)

I suggest that when a cognitive scientist works on mental representation, what we often find is a special kind of meeting between two conceptual frameworks and mindsets. The people doing cognitive science are people, who bring with them ordinary habits of folk psychological interpretation. But they are also scientists, and science brings with it special criteria for what to look for and describe. For example, any connections between brain states and the world which figure in a scientific theory should be describable in physicalistically acceptable terms.

What results is a special kind of interpretive practice, born of the meeting of folk interpretive habits and the special features of science. We should think of the "representational" concepts used in cognitive science as *amalgams*, or *hybrids*, born of the interaction between the ordinary interpretive habits that cognitive scientists have just in virtue of being people, and the scientific aims of describing precise, naturalistic and empirically studiable relations between organisms and environments. The representational concepts used in cognitive science are products of marriages between folk semantic concepts and a family of naturalistic concepts of physical specificity — concepts of connection and directedness that are based on causal, nomic and functional concepts.

The folk interpretive practices and the scientific concepts of specificity meet . . . and what results is a hybrid description tailored to the demands of some part of cognitive science. What results is a description of "what the frog's eye tells the frog's brain," or a description of how the visual system "infers" shape from shading. We see a similar kind of hybrid in some everyday descriptions of computational devices — when we say "the email program found that it was not connected to a server, so it told the modem to make a connection." (I hope this reminds people of the now-neglected program of "homuncular functionalism," defended by Dennett in one form (1978) and by Lycan in another (1981). The

intermediate levels of description envisaged by homuncular functionalists tend to borrow both from folk-psychological interpretive practices on the one hand and from physical forms of description on the other. They are hybrids too.)

Cognitive scientists will sometimes express their ideas as if they are using their scientific concepts to give a reduction or other unitary explanation of folk psychological semantic relations. I think this is a harmless mistake. In support of my claim that this is a mistake, I point to the fact that different cognitive scientists tend to have in mind radically different views about *which* naturalistic concepts of specificity are most fundamentally connected to representation and meaning. Some think the most important naturalistic relation here is covariation, of a special kind; others think it's resemblance, of a special and abstract kind. Others might think it's a concept of teleo-functional specificity, deriving from the biological concept of function. No one is right. These relations should not be conceived as rival attempts at the reduction of a folk concept of representation or meaning. Cognitive scientists forge different kinds of hybrid semantic concepts in different circumstances — in response to different theoretical needs, and different ways in which scientific concepts of specificity and folk habits of interpretation interact with each other. It might turn out that some one of these relations is more *scientifically* fundamental than the others — more fundamental to the project of explaining how intelligent systems work. But that, again, does not make this scientific relation into something that yields a reductive explanation of folk-psychological semantic properties.

Conclusion

As I said at the outset, this paper is a collection of pieces that might be part of an alternative picture of folk psychology and mental representation. One theme seen in the two main discussions (third and fourth sections) is a kind of pluralism. Old disputes about the role and status of folk psychology might be resolved with the aid of a pluralist view of how folk psychological interpretation operates. There are several different ways in which this idea might be developed. And we might think of much of the work of Dretske, Millikan and others as describing a range of naturalistic relations between internal and external states, relations which may well have importance within cognitive science even though they cannot be used to give a reductive analysis of the content of folk psychological representational states. I do not want to give the impression that "pluralist" alternatives are always the way to resolve problems, however. Pluralism is not always an advance (especially when there is only one thing, as Mark Twain might have said). But one way or another, we do need to explore some new approaches to folk psychology and mental representation.

Acknowledgments

Thanks to Hugh Clapin for organizing the Sydney conference, and to all those who participated in discussion of these topics. This paper is partly an attempt to reassess the force of some of Daniel Dennett's and Stephen Stich's arguments (especially those in Dennett 1981). This paper has been also influenced by discussions with many people over the years, but I should give special acknowledgment to the role that a coffee with Huw Price at *Badde Manors* and a beer with Kim Sterelny, both about five years ago, had in getting me to think along different lines about these problems.

References

Churchland, P. M. (1981). Eliminative materialism and the propositional attitudes. *Journal of Philosophy, 78*, 67–90.

Churchland, P. S. (1986). *Neurophilosophy: Toward a unified science of the mind/brain*. Cambridge, MA: MIT Press.

Davies, M., & Stone, T. (Eds) (1995). *Folk psychology: The theory of mind debate*. Oxford: Blackwell.

Dennett, D. C. (1978). *Brainstorms. Philosophical essays on mind and psychology*. Cambridge, MA: MIT Press.

Dennett, D. (1981). Three kinds of intentional psychology. Reprinted in: D. C. Dennett (Ed.), *The intentional stance*. Cambridge, MA: MIT Press.

Dennett, D. C. (1991). Two contrasts: Folk craft vs. folk science, and belief vs. opinion. In: J. Greenwood (Ed.), *The future of folk psychology: Intentionality and cognitive science*. Cambridge: Cambridge University Press.

Dretske, F. (1981). *Knowledge and the flow of information*. Cambridge, MA: MIT Press.

Dretske, F. (1988). *Explaining behavior*. Cambridge, MA: MIT Press.

Fodor, J. A. (1987). *Psychosemantics*. Cambridge, MA: MIT Press.

Godfrey-Smith, P. (2002). On the evolution of representational and interpretive capacities. *Monist, 85*(1), 50–69.

Jackson, F., & Pettit, P. (1990). In defence of folk psychology. *Philosophical Studies, 59*, 31–54.

Lycan, W. G. (1981). Form, function and feel. *Journal of Philosophy, 78*, 24–49.

Lycan, W. G. (1988). *Judgment and justification*. Cambridge: Cambridge University Press.

Millikan, R. G. (1984). *Language, thought, and other biological categories*. Cambridge, MA: MIT Press.

Ramsey, W., Stich, S., & Garron, J. (1991). Connectionism, eliminativism and the future of folk psychology. In: J. Greenwood (Ed.), *The future of folk psychology: Intentionality and cognitive science*. Cambridge: Cambridge University Press.

Rey, G. (1997). *Contemporary philosophy of mind: A contentiously classical approach*. Oxford: Blackwell.

Sterelny, K. (1990). *The representational theory of mind: An introduction*. Oxford: Blackwell.

Stich, S. P. (1992). What is a theory of mental representation? *Mind*, *101*, reprinted in: Stich & Warfield (1994).

Stich, S. P., & Warfield, T. A. (Eds) (1994). *Mental representation: A reader*. Oxford: Blackwell.

Stone, T., & Davies, M. (Eds) (1996). *Mental simulation: Evaluations and applications*. Oxford: Blackwell.