

RuCCS TR-5

# What *is* Folk Psychology?

**Stephen Stich**

Center for Cognitive Science  
Department of Philosophy  
Rutgers  
The State University of New Jersey  
stich@ruccs.rutgers.edu

**Ian Ravenscroft**

Department of Philosophy  
Research School of Social Science  
Australian National University  
Canberra, ACT 2601  
mraven@coombs.anu.edu.au

Technical Report #5, September 1993  
Rutgers University Center for Cognitive Science  
Rutgers University  
PO Box 1179  
Piscataway, NJ 08855

# What *is* Folk Psychology?

## I. Introduction

For the last two decades a doctrine called “eliminative materialism” (or sometimes just “eliminativism”) has been a major focus of discussion in the philosophy of mind. It is easy to understand why eliminativism has attracted so much attention, for it is hard to imagine a more radical and provocative doctrine. What eliminativism claims is that the intentional states and processes that are alluded to in our everyday descriptions and explanations of people’s mental lives and their actions are *myths*. Like the gods that Homer invoked to explain the outcome of battles, or the witches that Inquisitors invoked to explain local catastrophes, they *do not exist*. According to eliminativists, there are no such things as beliefs or desires or hopes or fears or thoughts. These putative states and processes are the badly misguided posits of a seriously mistaken theory, just like phlogiston and caloric fluid and the luminiferous ether.<sup>1</sup>

If eliminativism is right, then as Jerry Fodor has suggested, it might well be “the greatest intellectual catastrophe in the history of our species.” (1987, xii) To see why, we need only consider the consequences of the doctrine in various domains of intellectual activity. Let’s start with history: Did Lincoln sign the Emancipation Proclamation because he wanted to abolish slavery? Or was it because he thought it would be a strategically useful move, helping to weaken the Confederacy? If eliminativism is right, then neither of these explanations could possibly be correct, since there are no wants and there are no thoughts. Consider epistemology: From Descartes to the present, epistemologists have tried to construct a systematic theory that will tell us which of a person’s beliefs are justified, and which are not. But if eliminativism is right, there are no justified beliefs; there are no beliefs at all. Or consider anthropology: Some researchers have claimed that a variety of human emotions, like fear, surprise and disgust are cultural universals rooted in biology; others have urged that all emotions are “social constructions.” But if eliminativism is right, then there is something profoundly misguided about this dispute. For fear, surprise and disgust are intentional states, and eliminativism claims that there are no such things. Finally, consider psychology: If eliminativism is right, then much of what goes on in clinical psychology is bound to be useless. People’s problems can’t be remedied

---

<sup>1</sup> Another species of eliminativism claims that conscious states do not exist. In this paper, however, our focus will be on the version of eliminativism that takes intentional states as its target.

by removing irrational beliefs or making them aware of subconscious desires; there are no such things. And, obviously, if eliminativism is right, then as Fodor insists, many cognitive psychologists ought to “do [their] science in some other way....” Or at least, they “should stop spending the taxpayer’s money.” (1990, p. 202-3)

Although advocates of eliminativism are not always as clear or careful as one might wish, they are typically inclined to make four distinct claims that might be formulated as follows:

(1) ‘Belief,’ ‘desire’ and other familiar intentional state expressions are among the theoretical terms of a commonsense theory of the mind. This theory is often called ‘folk psychology’.

(2) Folk psychology is a seriously mistaken theory. Many of the claims it makes about the states and processes that give rise to behavior, and many of the presuppositions of those claims, are false.

(3) A mature science that explains how the mind/brain works and how it produces the behavior we observe will not refer to the commonsense intentional states and processes invoked by folk psychology. Beliefs, desires and the rest will not be part of the ontology of a mature scientific psychology.

(4) The intentional states of commonsense psychology do not exist.

It is clear that the first of these claims is a crucial presupposition of the second. After that, the putative relations among the claims gets a bit murky. It sometimes appears that both friends and foes of eliminativism assume that (2) can be used to establish (4). And, of course, if (4) is right then (3) comes pretty much for free. For if beliefs and desires don’t exist then surely a mature science has no business invoking them. In other places it seems that (2) is taken to support (3). If that’s the way the argument is supposed to go, then presumably (3) will be taken to support (4), though explicit arguments from one to the other are not easy to find.

Most of the literature debating the plausibility of eliminativism has focused on the second of these claims.<sup>2</sup> In this paper, however, our focus will be on the first. That premise of the eliminativist argument has already provoked a certain amount of debate, with some writers protesting that commonsense psychology

---

<sup>2</sup> See, for example, P. M. Churchland (1981); P. S. Churchland (1986); Fodor (1987); Horgan & Woodward (1985); Jackson & Pettit (1990); Kitcher (1984); Ramsey, Stich & Garon (1990); Sterelny (1990); Stich (1983); Von Eckardt (1993).

cannot be regarded as a causal or explanatory theory because its principles are partly normative or largely analytic. Others maintain that the basic function of folk psychology is not to predict and explain, but to warn, promise, congratulate and to do a thousand and one other jobs that are fundamental to the smooth workings of our interpersonal lives.<sup>3</sup> Eliminativists typically concede most of these points, but argue that it makes little difference. Whatever other uses it may have, they insist, folk psychology is still a causal and explanatory theory, and a seriously mistaken one.<sup>4</sup>

We propose to raise a rather different collection of concerns about the idea that ordinary intentional expressions are theoretical terms in a commonsense theory. Our central contention is that this idea can be unpacked (and, indeed, *has* been unpacked) in a variety of very different ways. Though many writers on both sides of the eliminativism debate take (1) to be unambiguous and unproblematic, there are actually *lots* of things that the label “folk psychology” might be (and *has* been) used to denote. Moreover, on *some* interpretations of (1) the remainder of the eliminativist argument will be in serious trouble. For on some readings, “folk psychology” is not the sort of thing that makes claims or expresses propositions. Thus it is not the sort of thing that *can* be either true or false. And obviously, on those readings the second step in the eliminativist argument couldn’t possibly be right. For if folk psychology makes no claims, it makes no false claims. Our goal in this paper is to pull apart these various readings of “folk psychology” and to get as clear as possible on which ones are and which are not compatible with the remainder of the eliminativist’s argument.

Before getting on to that, however, it will be useful to consider another issue. The idea that ‘belief,’ ‘desire’ and other intentional locutions are terms embedded in a commonsense theory has become commonplace in the philosophy of mind. But though talk of a “folk theory” and its “posits” is all but ubiquitous in the recent literature, it is also rather puzzling. Ordinary folk certainly don’t take themselves to be invoking a theory when they use intentional terms to explain other people’s behavior. Still less do they think they are using a theory when they report their own beliefs and desires. So why is it that so many philosophers and cognitive scientists are convinced that our everyday predictions and explanations of behavior do involve some sort of theory? Why does this idea seem so plausible to many philosophers and psychologists, and so implausible to almost everyone else? One good way to approach these questions is to track down the history of the idea. That is what we propose to do in the two sections to follow. While we do not pretend to be serious historical scholars, we think it

---

<sup>3</sup> See, for example, Sharp (1987); Wilkes (1981, 1984, 1991).

<sup>4</sup> See Stich (1983), Ch. 10; P. M. Churchland (1989).

is pretty clear that the view set out in (1) has two major historical roots. One of them is to be found in the work of Wilfrid Sellars, the other in the dominant explanatory strategy of contemporary cognitive science.

## II. Folk Psychology's Sellarsian Roots

A major theme in Sellars' philosophy is a sustained attack on "the myth of the given" — the idea that some of our beliefs or claims have a privileged epistemic status because the facts that make them true are "given" to us by experience. One class of claims that has traditionally been accorded this special status are pronouncements about one's own "sense data" or the content of one's perceptual experience. On the traditional view, a person's sincere claim that she is now seeing a blue object might well turn out to be mistaken. But her sincere claim that she is now experiencing blue sense data (or that she is now having experiences "as if" she were seeing a blue object) could not turn out to be mistaken. Another class of claims that are immune from error, according to the traditional view, are claims about one's own apparent memories and beliefs. Stich can't be certain that he has in fact climbed Ayers Rock. But he can be certain that he now seems to remember climbing Ayers Rock. And while his belief itself might be false, his sincere claim that he believes he climbed Ayers Rock on his 42nd birthday can't be mistaken. Sellars was a trenchant critic of these claims to certainty, and of the foundationalist epistemology that typically went along with them. And though his assault on the traditional notion of sense data is not directly relevant to the eliminativist's skepticism about intentional states, his attack on the idea that our claims about our own beliefs and memories could not be mistaken most emphatically is. For, of course, if Stich's sincere claim that he believes he climbed Ayers Rock is enough to guarantee that he *does* believe it, then, since we make such sincere claims all the time, beliefs must exist, and eliminativism is a non-starter.

To counter the idea that our claims about our own beliefs and thoughts are underwritten by a special, introspective faculty that guarantees the truth of those claims, Sellars begins by "making a myth ... or, to give it an air of up-to-date respectability, by writing a piece of science fiction — anthropological science fiction." (1956, p. 309) For our purposes, Sellars' myth can be viewed as having three stages. The first of these is "a stage in pre-history in which humans are limited to what I shall call a Rylean language, a language of which the fundamental descriptive vocabulary speaks of public properties of public objects located in Space and enduring through Time." (309) At this stage in the myth, our "Rylean Ancestors" have no terms in their language for beliefs, thoughts or other "inner mental episodes." The second stage in the myth begins with the appearance, in this "Neo-Rylean culture" of "a genius — let us call him Jones." (314)

[I]n the attempt to account for the fact that his fellow men behave intelligently not only when their conduct is threaded on a string of overt verbal episodes — that is to say, as we would put it, when they ‘think out loud’ — but also when no detectable verbal output is present, Jones develops a *theory* according to which overt utterances are but the culmination of a process which begins with certain inner episodes. *And let us suppose that his model for these episodes which initiate the events which culminate in overt verbal behavior is that of overt verbal behavior itself. In other words, using the language of the model, the theory is to the effect that overt verbal behavior is the culmination of a process which begins with “inner speech.”* (317-318; emphasis is Sellars’)

At this stage of Sellars’ myth, the theory is only applied to other people. But in the third stage Jones and his compatriots learn to apply the theory to themselves. At first they apply it to themselves in much the same way that they apply it to others. They infer various theoretical claims by attending to their own behavior. A bit later, they discover a new way of applying the language of the theory to themselves. Here is how Sellars tells the tale:

[O]nce our fictitious ancestor, Jones, has developed the theory that overt verbal behavior is the expression of thoughts, and taught his compatriots to make use of the theory in interpreting each other’s behavior, it is but a short step to the use of this language in self-description. Thus, when Tom, watching Dick, has behavioral evidence which warrants the use of the sentence (in the language of the theory) “Dick is thinking ‘p’ ” ... Dick, using the same behavioral evidence, can say, in the language of the theory, “I am thinking ‘p’ ”... And it now turns out — need it have? — that Dick can be trained to give reasonably reliable self-descriptions, using the language of the theory, without having to observe his overt behavior. Jones brings this about, roughly, by applauding utterances by Dick of “I am thinking that p” when the behavioral evidence strongly supports the theoretical statement “Dick is thinking that p”; and by frowning on utterances of “I am thinking that p,” when the evidence does not support this theoretical statement. Our ancestors begin to speak of the privileged access each of us has to his own thoughts. *What began as a language with a purely theoretical use has gained a reporting role.*

(320; emphasis is Sellars’)

So, in Sellars’ myth, expressions of the form ‘I am thinking that p’ are theoretical expressions which have acquired “a reporting use in which one is not drawing inferences from behavioral evidence.” (321)

Now if, like Sellars, one is concerned to rebut the claim that our reports of our own thoughts are beyond challenge, the myth of Jones suggests how the argument might run. For suppose the myth were true. The inner episodes that Jones hypothesizes in stage two are supposed to be real events that are causally linked with behavioral episodes. Positing them to account for certain aspects of the observable behavior of people is, as Sellars stresses, on all fours with positing molecules to account for certain aspects of the observable behavior of gases. Thus, for mental states as for molecules, there will be some occasions on which the inference from the observed behavior to the theoretical claim may be mistaken. Occasionally, an anomalous event may cause the observed behavior in the absence of the hypothesized internal state. Similarly, when we have been trained to give “reasonably reliable self-descriptions, using the language of the theory, without having to observe [our own] overt behavior” it may occasionally happen that this process misfires, and that we describe ourselves as thinking that *p*, in the absence of the hypothesized internal state. Moreover, though Sellars himself did not stress the point, there is a more pervasive way in which our self-descriptions might turn out to be wrong. For it might turn out that Jones was less of a genius than we had thought — more of a Velikovsky, perhaps, than a Newton. His entire theory might turn out to be a bad idea. Other thinkers might discover better ways to explain the behavior that Jones’ theory was designed to explain — ways that don’t posit internal states modeled on observable verbal behavior. If that’s the way the myth unfolds, then it may not be just the occasional theoretical self-description that turns out to be false. They may *all* be false.

Before we take these possibilities seriously, however, there is a pair of problems that must be confronted. The first of these focuses on the status of the myth itself. The possibilities set out in the previous paragraph were preceded by the *supposition* that the myth is true. But surely that’s just silly. There was no historical Jones or anyone much like him, and there is no reason at all to suppose that talk about thoughts and other inner mental events was introduced as an explicit theoretical posit. Presumably what Sellars would say here is that the myth is a bit like Wittgenstein’s ladder which we kick away after we have climbed it. The reason for telling the myth about Jones is to loosen the grip of another myth, this one the Cartesian “myth” in which introspection gives rise to infallible knowledge about our own mental states. If the Sellarsian myth were true, then we would talk *just as we now do* about inner mental states. But this talk would be both theoretical and fallible. Once we appreciate the point, the myth is irrelevant. It doesn’t much matter what the actual history of our language is. What matters is that people could talk about inner mental states just the way we do, and their sincere self reports could be mistaken.

A second problem in assessing the significance of Sellars’ myth focuses not on the truth of the myth but on the nature of the theory that the myth describes. As Sellars tells the story, Jones actually set out a theory and taught it to his compatriots. But nothing much like that seems to go on in our current practice.

We don't explicitly teach our children a theory that enables them to apply mental terms to other people. Indeed, unlike Jones and his friends, we are not even able to *state* the theory, let alone teach it. If you ask your neighbor to set out the principles of the theory of the mind that she has taught her children, she won't have the foggiest idea what you're talking about. An essential step in Sellars' argument is the claim that, if the myth were true, we would talk just as we now do about mental states. But isn't this rather implausible? Surely if each of us had been taught an explicit theory of the mind, and if we invoke this theory in applying mental state terms to others, then both our developmental history and our adult linguistic practice would be rather different from what they actually are. At a minimum, there would be more awareness and discussion of the theory, and more explicit appeal to its principles than we find in current linguistic practice. Rather than tackling this problem head on, we think the best strategy, at this point, is to break off our discussion of Sellars, and attend to the other major source of the idea that mental states are the posits of a folk theory. As we noted earlier, that source is to be found in the dominant explanatory strategy of contemporary cognitive science. As our discussion of that strategy proceeds, we will find an obvious way in which a neo-Sellarsian might respond to the objection that none of *us* can even state Jones' theory, let alone teach it.

### III. Cognitive Science and the Appeal to Tacit Theory

From its earliest days, a central concern of cognitive science has been the explanation of various cognitive or behavioral capacities. The capacity of speakers to recognize and make various judgements about sentences of their language was among the first to be studied in detail. Other pioneering studies attempted to explain the capacity to solve various sorts of problems including problems in logic, chess problems, criptarithmic problems and a variety of others. During the last three decades, many additional capacities have been explored. In all of this work, the most common explanatory strategy is to suppose that people's performance in executing these various capacities is guided by an internalized "knowledge base" or "knowledge structure." Typically a knowledge structure will be a set of principles or rules that constitute a recipe or program enabling people to carry out the activity in question by exploiting more basic capacities in a systematic way. Those more basic capacities can themselves be explained by reiterating the strategy at a lower level. They are decomposed into still more basic ones which are systematically marshalled under the guidance of another recipe or program.

The first philosophical account of this approach to psychological explanation that we know of was provided in Jerry Fodor's paper, "The Appeal to Tacit Knowledge in Psychological Explanation." (1968) And, though the picture has become a fairly familiar one, we think it is worth quoting Fodor's vivid exposition at some length.

Here is the way we tie our shoes:

There is a little man who lives in one's head. The little man keeps a library. When one acts upon the intention to tie one's shoes, the little man fetches down a volume entitled *Tying One's Shoes*. The volume says such things as: "Take the left free end of the shoelace in the left hand. Cross the left free end of the shoelace over the right free end of the shoelace..., etc.

When the little man reads the instruction 'take the left free end of the shoelace in the left hand', he pushes a button on a control panel. The button is marked 'take the left free end of a shoelace in the left hand.' When depressed, it activates a series of wheels, cogs, levers, and hydraulic mechanisms. As a causal consequence of the functioning of these mechanisms, one's left hand comes to seize the appropriate end of the shoelace. Similarly, *mutatis mutandis*, for the rest of the instructions.

The instructions end with the word 'end'. When the little man reads the word 'end', he returns the book of instructions to his library.

That is the way we tie our shoes. (63-4)

Now, as Fodor goes on to note, there are some obvious things wrong with this story. First, of course, the whimsical bit about the cogs and wheels will have to be replaced by a more biological story. Second, and more seriously,

some of the behaviors ... involved in shoe tying are of considerable complexity.... A serious theory of the behavioral integrations involved in tying one's shoes must explicate this complexity.... Prima facie ... grasping a shoelace should be considered complex behavior, because doing so involves motions that also play a role in other actions.

We might thus consider expanding the population in one's head to include subordinate little men who superintend the execution of the "elementary" behaviors involved in complex sequences like grasping a shoelace. When the little man reads 'take the left free end of the shoelace in the left hand', we imagine him ringing up the shop foreman in charge of grasping shoelaces. The shop foreman goes about supervising that activity in a way that is, in essence, a microcosm of supervising tying one's shoe. Indeed the shop foreman might be imagined to superintend a detail of wage slaves, whose functions include: searching inputs for traces of shoelace, flexing and contracting fingers on the left hand, etc. (64-5)

A bit later, Fodor explains how this process ultimately comes to an end:

We refine a psychological theory by replacing global little men by less global little men, each of whom has fewer unanalyzed behaviors to perform than did his predecessors...

A completed psychological theory must provide systems of instructions to account for the forms of behavior available to an organism, and it must do so in a way that makes reference to no unanalyzed psychological processes. One way of clarifying the latter requirement is the following. Assume that there exists a class of *elementary* instructions which the nervous system is specifically wired to execute. Each elementary instruction specifies an *elementary operation*, and an elementary operation is one which the normal nervous system can perform but of which it cannot perform a proper part. Intuitively speaking, the elementary operations are those which have no theoretically relevant internal structure. (65-6)

There are three additional points that need to be made before asking how this explanatory strategy links up with our concerns about folk psychology. The first is that the strategy does not require that bosses have any conscious access to the information their underlings are using. In some cases a person may be able to tell us a great deal about the principles that guide his activity; in other cases he may be able to report some of the principles but not others; and in still other cases he may not have a clue about how he accomplishes the task. In those cases where the person can't recount or even recognize the principles he (or one of his sub-systems) is using, it is often said that the principles are "tacit" or "unconscious." The second point is that it is often natural enough to describe the principles being used as a "theory" of the task domain or of how to accomplish the task in question. So, putting this point together with the previous one, it will sometimes be the case that the principles specified in an explanation of the sort Fodor envisions will constitute a "tacit or unconscious theory" of the domain or task in question. Here, of course, we have an obvious way to address the problem that we left unresolved in our discussion of Sellars' myth. If people regularly exploit *tacit* theories of the sort that Fodor describes, then we should not expect them to be aware of the principles of the theory or to appeal to those principles in discussion.

The third point is a bit more subtle. In Fodor's account, the little man inside the head has a single book specifying a set of rules for accomplishing the task at hand. But we might also imagine that in some instances the little man has *two* books for a given ability. One of the books contains declarative sentences rather than rules. These might, for example, be a set of axioms for some branch of mathematics or science. Or they might be a set of principles detailing generalizations and more idiosyncratic facts in some other domain. Now, of

course, axioms or generalizations or statements of fact cannot, by themselves, tell us how to do anything. That's the job of the second book, which is much the same as the book imagined in Fodor's shoe tying example. It provides rules for using the information in the first book to accomplish some task. So, for example, if the first book contains statements of the laws in some branch of physics, the second book might contain rules which specify how to use these laws to solve physics problems. Or perhaps the first book contains an axiomatic specification of all the sentences in a given language, and the second book contains a set of rules for using this specification efficiently in determining whether or not a given sentence is in the language. If one thinks of theories as the sorts of things that make claims, and thus as the sorts of things that can be true or false, then one might be inclined to say that only the books containing declarative sentences count as "theories". The books that contain programs or recipes can do a good job or a bad job at accomplishing their task. But since they don't make any claims, they don't count as theories. We don't think that anything much turns on this terminological issue. What is important is the point about truth. If the little man accomplishes his task using only a recipe or a program, we may, if we wish, choose to describe that program as a theory. But it makes no obvious sense to say that the "theory" he is exploiting is either true or false.

#### IV. Interpreting "Folk Psychology"

The central goal of this paper, it will be recalled, is to explore various possible interpretations of the assumption that beliefs, desires and other commonsense mental states are posits of a folk theory of the mind. We are now in a position to tackle that project head-on.

Cognitive science, as we have just seen, typically attempts to explain cognitive and behavioral capacities by positing an internalized and often tacit theory. Among the various capacities that normal adults display, there is a fascinating cluster in which talk of mental states looms large. It is hardly surprising that many people have been tempted by the idea that this cluster of capacities might be explained by appeal to a tacit theory. Before considering the various ways in which such an explanation might work, we would do well to assemble at least a partial list of the "folk psychological" capacities or abilities that need explaining.

i) Perhaps the most obvious of these is the one that was center stage in Sellars' myth of Jones. We use terms like 'believe,' 'think,' 'want' and 'desire' to *describe* ourselves and each other. We say things like the following all the time:

Columbus believed that the earth was round.  
Henry VIII wanted a male heir.  
Anna thinks there is someone knocking at the door.

And, while we occasionally dispute such claims, in the overwhelming majority of cases they are readily accepted and strike us as completely unproblematic. This widespread agreement is a manifestation of a widely shared capacity to describe - or as philosophers sometimes like to say, to *interpret* - people using intentional idioms.<sup>5</sup>

ii) We not only describe people using these folk psychological idioms, we also use the descriptions to construct *explanations* of people's behavior. We say things like:

Henry VIII divorced Catherine of Aragon because he wanted a male heir.  
and  
Anna is looking through the peep-hole because she thinks that there is someone knocking at the door.

And here too, in the vast majority of cases these explanations are widely accepted and strike us as entirely unproblematic. Within our culture a capacity to construct this sort of explanation seems to be universal among normal adults.

iii) In addition to offering explanations of behavior, we are also quite adept at producing *predictions* of how people will behave. Sometimes these predictions are embedded in a discourse that also mentions the beliefs, desires and other mental states of the person whose behavior is being predicted. But on other occasions we predict someone's behavior without saying anything about her mental states. In one respect, our ability to produce predictions is rather similar to the previous two abilities on our list. For in this case, too, there is remarkably widespread inter-personal agreement. Asked to predict what the motorist will do as she approaches the red light, almost everyone says that she will stop. But there is another respect in which the ability to predict is much more impressive than the ability to offer descriptions and explanations. For in the case of predictions, there is often an independent and readily available check on their accuracy. And, as many philosophers have noted, it is a striking fact that in the vast majority of cases our predictions turn out to be *correct*. Though we are certainly not infallible, it is very often the case that people do what we predict they are going to do.

---

<sup>5</sup> It is often claimed that at least some of those intentional idioms, and the capacity to apply them, vary markedly from culture to culture. (See, for example, Hacking (1982), Needham (1972) and the essays collected in Harre (1986).) That sort of cultural relativism, if it turns out to be correct, is entirely compatible with the various accounts of folk psychology to be set out in this section. If people in different cultures use different intentional categories, and if their use of these categories is guided by a tacit theory, then the tacit theories will also differ from culture to culture.

iv) The ability to produce predictions of people's behavior is one which manifests itself in our *linguistic* behavior. But we also have a capacity to *anticipate* other people's behavior without saying anything at all. In watching a baseball game, our eyes immediately jump to the hot tempered manager when the umpire throws his star player out of the game. We anticipate that he will come storming out of the dugout. Seeing the furious manager approaching, the umpire may anticipate what is to come and beat a hasty, though silent, retreat. Now it might be thought that these cases are just like the ones in which people make verbal predictions, except that they don't actually utter the prediction. They just say it silently to themselves. But there is also a *prima facie* similarity between our ability to anticipate people's behavior and the ability that animals have to anticipate the behavior of other organisms. The family cat is very good at anticipating which way mice will run, and at anticipating which way the neighbor's dog will run. In both cases he reacts appropriately, though we rather doubt that he is saying anything to himself as he does it.

v) The final entry on our list is the only one that overtly involves what might be thought of as *principles* or *generalizations* of a folk psychological theory. There is a wide range of generalizations about the interactions among stimuli, mental states and behavior that people in our culture occasionally utter, and are generally quite willing to endorse when asked. To give you a feel for the sort of ability we have in mind consider whether or not you agree with the following claims:

(v-i) When a normal person is looking at a traffic light which changes from red to green she usually comes to believe that it has changed from red to green.

(v-ii) If a person believes that all scorpions are poisonous, and if she comes to believe that Henry's pet is a scorpion, then she will typically come to believe that Henry's pet is poisonous.

(v-iii) If a person sitting at a bar wants to order a beer, and if she has no stronger desire to do something that is incompatible with ordering a beer, then typically she will order a beer.

We trust you agreed with all of them. In so doing you were manifesting the widely shared ability to recognize folk psychological generalizations.

That's the end of our list of capacities. It is, as we noted earlier, only a partial list. Normal adults in our society have lots of other abilities in which talk of mental states plays a central role. In the absence of a theory about the mechanisms underlying these abilities, there is no obvious or natural way of drawing a boundary and saying exactly which capacities do and do not count as "folk psychological capacities." That vagueness will make for problems as we

proceed. But for the moment we propose to ignore it and focus on the five capacities we have listed.

Let's begin by assuming that something like Fodor's story provides the correct explanation for those abilities, and let's consider some of the ways in which this story might be developed. To start us off, let's ask how the explicit generalizations mentioned in (v) are related to the underlying knowledge structure - the book that the little man in the head consults. Perhaps the simplest answer is that *these very generalizations* are encoded in the underlying knowledge structure. Indeed, to tell a really simple story we might suppose that the underlying knowledge structure consists of *nothing but* these explicit generalizations. If this simple story were right, then all the principles we use in employing our various folk psychological abilities would be readily recognizable. But, for two very different reasons, we think this is an unlikely option. First, there are just too many generalizations. People who find (v-i)-(v-iii) to be intuitively obvious will find an all but infinite class of similar generalizations to be equally obvious. And it seems absurd to suppose that all of those generalizations are internally represented. A natural suggestion here is that what we have internally represented is a set of more abstract generalizations - we might think of them as "axioms" of folk psychology - which in conjunction with other internalized information entail (v-i)-(v-iii) and all the other more concrete generalizations that we recognize as intuitively obvious. The second problem with our simple story is just the opposite of the first. If we restrict ourselves to the generalizations that people are prepared to endorse, then in all likelihood there are too few of them to do the job required. A serious and fully explicit account of how we accomplish the feats recounted in (i)-(v) will almost certainly require more rules and principles than people can state or recognize as intuitively obvious. It is to be expected that in this area, as in every other area that cognitive scientists have explored, there is a great deal of information being used that is simply not accessible to introspection. If this is right, then at least part of what is written in the little man's book will be "tacit" or "unconscious". Moreover some of the information that is not available to introspection may not even be statable in terms that the agent can understand. Linguistic ability provides a valuable analogy here. If Chomsky and his followers are even close to being right about the "tacit knowledge" that subserves a speaker's ability to judge whether or not a given sentence is grammatical, then most people would require a crash course in generative linguistics before they could begin to understand an explicit statement of what they "tacitly know".

We can now begin to pull apart some of the very different ways in which the label "folk psychology" might be used. In a series of influential papers, David Lewis drew attention to what he called the "platitudes" of commonsense psychology. (Lewis, 1970, 1972) These are generalizations that are "common knowledge" amongst ordinary folk. Almost everyone assents to them, and almost everyone knows that almost everyone else assents to them. These platitudes are, near enough, the intuitively obvious generalizations discussed in

the previous paragraph. On Lewis' view these platitudes constitute an implicit definition of the terms of commonsense psychology. But suppose that the speculations in the last paragraph are correct, and that platitudes like (v-i)—(v-iii) are the consciously accessible consequences of a substantially richer set of mostly tacit or unconscious psychological rules and generalizations that people in our culture share. Suppose, further, that these tacit rules and generalizations also play a central role in explaining folk psychological capacities like (i) - (iv).<sup>6</sup> If these suppositions are correct, then we might well be tempted to reserve the term "folk psychology" for the *underlying internally represented* rules and generalizations. Moreover, a neo-Lewisian might well propose that it is these internal generalizations that fix the meaning or reference of the terms of commonsense psychology.

There is, however, nothing mandatory about this terminology. We might equally well elect to use the term "folk psychology" in a way that is more akin to Lewis' usage — as a label for the collection of folk psychological "platitudes" that people in our culture readily recognize and assent to.<sup>7</sup> Or, since the collection of "platitudes" is likely to be large and ungainly, we might reserve the label "folk psychology" for a set of more abstract generalizations — a "theory" if you will— that systematizes the platitudes in a perspicuous way and that (perhaps in conjunction with some other commonly known information) entails them. That systematization might well invoke terms and concepts that are quite unfamiliar to ordinary folk, in the same way that an attempt to systematize our linguistic intuitions probably would. What makes the proposals set out in this paragraph fundamentally different from the various possibilities considered in the previous paragraph is that on these readings, *"folk psychology" is not claimed to be an internally represented knowledge structure or body of information; it is not part of the mechanism that subserves the abilities recounted in (i) - (v).* On these readings, folk psychology "ain't in the head." To mark this distinction we propose to call these

---

<sup>6</sup> As we noted earlier, there is no obvious way of deciding which capacities to count as "folk psychological capacities." Thus the current supposition is itself a bit vague. It might turn out that the rules and generalizations subserving (i) - (v) are a tightly integrated set, perhaps even a "module" in something like the sense set out in Fodor (1983). If that is the case, then we would do well to let these integrated rules and generalizations determine which capacities to count as "folk psychological" — the folk psychological capacities are the capacities that these rules and generalizations explain. But it might also turn out that the rules and generalizations we use are not modular or even tightly integrated — that different subsets of rules and generalizations explain different capacities. If that's how things work, then the only way to eliminate the vagueness is by stipulation. Folk psychological capacities will not be a psychologically natural kind.

<sup>7</sup> Here, again, the proposal is a bit vague, since there is no obvious well motivated way to determine which platitudes count as "folk psychological." Nor is it clear how widely accepted a claim must be in order to count as a "platitude".

accounts of folk psychology *external* accounts. Accounts on which folk psychology *is* part of the mechanism subserving (i) - (v) we call *internal*.

The distinction between internal and external accounts of folk psychology is at least roughly parallel to a distinction between two different ways of interpreting the sorts of generative grammars produced by linguists. Linguistic intuitions are a principle source of data for generative grammar. These intuitions are spontaneous judgements about the grammaticality and other grammatical features of sentences presented to a speaker. And it is generally agreed that a correct grammar is a theory which entails most of those judgements. On the analogy we are urging, linguistic intuitions are analogous to people's spontaneous judgements about the correctness or incorrectness of proposed folk psychological platitudes. Some theorists (e.g. Stich, 1972 and perhaps Soames 1984) claim that capturing the intuitions (along with "simplicity") is all that is expected of a grammar; a correct grammar, on this view, is nothing more than a useful systematization or axiomatization of linguistic intuitions. Other theorists (e.g. Chomsky & Katz, 1974; Fodor, 1981a) have higher aspirations for grammar. On their view, a grammar should not only capture (or entail) most linguistic intuitions, it should also be part of the mechanism that is causally responsible for the production of those intuitions, and for a variety of other linguistic capacities. On this view, people are assumed to have an internally represented body of linguistic rules or principles that is brought into play in producing linguistic intuitions and in processing and producing speech. A correct grammar is one that specifies the internally represented rules. Understood in this way, grammar is analogous to folk psychology, construed *internally*. On the other view, grammar is analogous to folk psychology construed *externally*.

It has often been noted that when grammars are construed externally they may be seriously underdetermined both by the data of linguistic intuition and by considerations of simplicity. There may be a number of quite different, though comparably simple, ways to construct a theory that entails most of a speaker's linguistic intuitions. So on an external view of grammar, there may be several quite different grammars of a language or a dialect, all of which are equally correct. Much the same is true for external accounts of folk psychology. For even if we find some principled way of saying which "platitudes" or folk psychological intuitions a theory must systematize, there may be very different, though comparably simple, ways in which this can be done. This point is particularly important if, as we speculated earlier, a good systematization of folk psychological intuitions may invoke terms or concepts that are unfamiliar to the people whose intuitions are being systematized. For these concepts might well be viewed as among the deeper "theoretical" notions of folk psychology. They are thus prime candidates for the eliminativist critique. But if there is no unique external systematization of folk psychology, then *the eliminativist who adopts an external account of folk psychology will have no determinate target*.

Let us return, now, to internal accounts of folk psychology. For the remainder of this section we propose to explore various ways in which attempts to construct internal accounts might turn out, and to consider the implications of these various possibilities for the eliminativist's argument.

At the end of Section III, we noted that for any given capacity, the little man in the head may have *two* books rather than one. The information guiding the exercise of various cognitive capacities may divide into two parts, one consisting of declarative sentences or some similar notation for encoding propositions about the relevant domain, and the other consisting of rules or procedures which specify what to do with these declarative sentences — how to use them to accomplish the task at hand. Applying this distinction to the case of commonsense psychology, we might conjecture that the knowledge structure underlying the skills in (i)-(v) divides into two parts. One of these is a set of (putative) laws and/or generalizations and/or facts about the ways in which the states of commonsense psychology interact with each other, with environmental stimuli, with behavior and with other relevant aspects of an agent's situation. The other part is a program — a set of rules for extracting predictions, explanations, etc. from these laws, generalizations and facts. If this is how the system works, it suggests two quite distinct ways in which the term “folk psychology” might be employed. It might be used as a label for the entire knowledge structure (the program plus the propositional information), or it might be reserved just for the part that contains the propositional information. On that latter usage, it makes perfectly good sense to ask whether folk psychology is true or false. On the former usage, however, only part of folk psychology is the sort of thing that can be true or false.

There is, however, another possibility to reckon with here, one which is much less congenial to the eliminativist's argument. It might turn out that the system subserving folk psychological skills contains only one book, not two, because the system is all rules and no propositions. If that's how the system works, and if we adopt the internal reading according to which “folk psychology” is used as a label for this system, then it will make no clear sense to say that folk psychology is either true or false. So if, as is entirely possible, folk psychology (construed internally) is all rules and no propositions, then the second step of the eliminativist's argument cannot possibly be correct. The upshot of all of this is that eliminativists who adopt an internal view of folk psychology are committed to an empirical speculation — the speculation that folk psychology *isn't* all rules and no propositions — and this speculation might well turn out to be mistaken.<sup>8</sup>

---

<sup>8</sup> A caveat: Even if folk psychology (construed internally) is all rules and no propositions, it may be the case that the rules of this folk psychological program, or the concepts they invoke, presuppose various claims that could turn out to be false. The notion of presupposition being relied on here could do with considerable clarification. But assuming that adequate clarification can be provided, the presuppositions of folk psychology might be a suitable target for the eliminativist's critique.

There is also quite a different way in which it might turn out that folk psychology, construed internally, is not the sort of thing that can possibly be either true or false. Thusfar we have been following Fodor in assuming that the “knowledge structure” underlying our folk psychological abilities is encoded in something akin to a language. But during the last decade there has been a growing tendency in cognitive science to propose theories in which the information subserving various abilities is not encoded in anything like linguistic form. Perhaps the most widely discussed theories of this type are those that propose to account for certain cognitive capacities by positing one or another sort of connectionist device. Quite a different idea for the non-linguistic encoding information is the hypothesis that mental models of various sorts underlie our cognitive capacities. (Johnson-Laird, 1983) Both the connectionist approach and the mental models approach are very controversial, of course. Moreover, to draw the conclusions we want to draw from them, we have to proceed rather carefully. It is perfectly possible to view certain sorts of connectionist models and certain sorts of mental model theories as storing information that can be straightforwardly captured by a set of propositions or a set of sentence. (McGinn, 1989, Ch. 3) Indeed, in some cases mental models and connectionist networks can actually be used to encode declarative sentences, or something very much like them. Thus it is not the case that connectionist or mental model approaches are inevitably in conflict with propositional or even linguistic accounts of information storage. However, in other connectionist and mental model theories, there may be no unique and well motivated way to map or translate the information stored into a set of propositions or declarative sentences. If a theory of this sort should prove to provide the best account of folk psychological skills like those sketched in (i)-(v), then we might well use the label “folk psychology” for the connectionist network or mental model posited by the theory. But since *ex hypothesis* there is no well motivated mapping from the network or model to a set of declarative sentences or propositions, it would make no obvious sense to say that folk psychology is either true or false. So in this case, too, eliminativists who adopt an internal view of folk psychology are buying into a controversial empirical assumption. They are assuming that folk psychological skills are *not* subserved by connectionist networks or mental models of the sort lately considered. Without this assumption, the eliminativist’s argument couldn’t get started.

It’s time to sum up. A central theme in this section has been that there are various quite different ways in which we might choose to use the term “folk psychology.” A first choice turns on the distinction between external and internal readings. External accounts either collect or systematize the intuitively recognizable generalizations of commonsense psychology, while internal accounts focus on the cognitive mechanism that underlies our ability to have those intuitions, to predict behavior, etc.. If we opt for an external reading of “folk psychology,” then folk psychology is clearly the sort of thing that makes claims, and some of those claims might turn out to be false. So on external readings, the eliminativist is guaranteed a target. But since there may be many quite different,

comparably simple ways to systematize the intuitively recognized generalizations, that target may be far from unique. Also, the target may not be all that interesting, since on external accounts, the principles of folk psychology may have little or nothing to do with the impressive range of commonsense psychological skills that people possess. They may have nothing to do with the processes by which we actually produce intentional descriptions, predictions and explanations of behavior.

Internal accounts use the label “folk psychology” for the knowledge structures that actually underlie skills like those recounted in (i) - (v). So on internal accounts folk psychology plays a major role in the explanation of our ability to predict and explain each other’s behavior. But on internal construals of folk psychology, the eliminativist’s argument *may* turn out to be incoherent. For it is entirely possible that the knowledge structure underlying our commonsense psychological skills consists *entirely* of instructions, or it may be a connectionist device or mental model that does not map comfortably on to a set of sentences or propositions. The eliminativist who adopts an internal reading of “folk psychology” must make the risky bet that none of these options will turn out to be correct. For if one of them is correct, then premise (2) in the eliminativist’s argument can’t be right, since folk psychology is not the sort of thing that can be either true or false.

We’re afraid that all of this has turned out to be rather complicated. Perhaps Figure 1 will be of some help in seeing how the various possible interpretations of “folk psychology” that we have been pulling apart are related to each other. The options that are compatible with step (2) in the eliminativist argument are on the left; those that are not are on the right. There is, however, one increasingly important view on folk psychology that does not fit comfortably anywhere in Figure 1. To conclude this paper we propose to take a brief look at that view and its implications for eliminativism.

---

### Figure 1 About Here

—*but not in this, the electronic version.*

*To obtain, send stamped self-addressed envelope to RuCCS.*

---

## V. Eliminativism and Simulation Theory

As we noted in Section III, the most common explanatory strategy in cognitive science posits internalized bodies of information, or “knowledge structures,” that are brought into play when people exercise cognitive capacities. Language processing, various sorts of problem solving, visual recognition and a

host of other abilities have been studied from within this explanatory paradigm. To many people it seems inevitable that an explanation of folk psychological capacities like those recounted in (i) - (v) will also posit some sort of internally represented information store. But in recent years a number of philosophers and psychologists, most prominently Robert Gordon (1986, 1992), Alvin Goldman (1989, 1992), and Paul Harris (1992), have suggested a way in which some of the capacities on our list might be explained without using anything much like the internalized knowledge structures that are omnipresent in contemporary cognitive science. Their central idea is that we might predict other people's behavior by using our own decision making system to simulate someone else's decision making. To make this strategy work we must first (consciously or unconsciously) imagine ourselves in the situation of the "target" — the person whose behavior we are trying to predict. We imagine what his beliefs are and what his desires are (to the extent that these are relevantly different from our own). We then feed these imagined (or "pretend") beliefs and desires into our own decision making system and allow it to decide what to do. But rather than acting on that decision, we use it as a basis for predicting what the target person will do. If we have done a good job in imagining the target's beliefs and desires, and if his decision making system works in much the same way that ours does, then our decision will be the same as his, and our prediction will be correct.<sup>9</sup>

Some critics of this proposal have suggested that it could not possibly work, or that it must *covertly* appeal to internalized rules or knowledge structures. We think both of these contentions are mistaken. The simulation theory provides a real alternative to the prevailing explanatory strategy in cognitive science for explaining our capacity to predict and explain other people's behavior. We are far from being advocates of the simulation theory, however, because we don't think it does a very good job of explaining the experimental data on the acquisition and deployment of folk psychological skills. The details of this critique make for a long story, which one of us has tried to tell elsewhere. (Stich & Nichols, 1992, forthcoming)

Our present concern is not to renew the debate over the correctness of simulation theory, but rather to ask what happens to the eliminativist argument if it should turn out (contrary to our expectations) that simulation theory is correct. A number of authors on both sides of the debate have maintained that if simulation theory is right, eliminativism will be undermined. Here is how Stich & Nichols (1992) argue for this conclusion:

---

<sup>9</sup> Advocates of simulation theory have also proposed ways in which this process might be used to generate intentional descriptions and explanations. See Stich & Nichols (1992), Sec. 3.

*Eliminativists* maintain that there really are no ... [intentional states]. Beliefs and desires are like phlogiston, caloric and witches; they are the mistaken posits of a radically false theory. The theory in question is “folk psychology” — the collection of psychological principles and generalizations which, according to eliminativists (and most of their opponents) underlies our everyday explanations of behavior. The central premise in the eliminativist’s argument is that neuroscience (or connectionism or cognitive science) is on the verge of demonstrating persuasively that folk psychology is false. But if Gordon and Goldman are right, they will have pulled the rug out from under the eliminativists. For if what underlies our ordinary explanatory practice is not a theory at all, then obviously it cannot be a radically false theory. There is a certain delightful irony in the Gordon/Goldman attack on eliminativism. Indeed, one might almost view it as attempting a bit of philosophical jujitsu. The eliminativists claim that there are no such things as beliefs and desires because the folk psychology that posits them is a radically false theory. Gordon and Goldman claim that the theory which posits a tacitly known folk psychology is *itself* radically false, since there are much better ways of explaining people’s abilities to interpret and predict behavior. Thus, if Gordon and Goldman are right, *there is no such thing as folk psychology!*<sup>10</sup>

Now it’s our contention that Stich & Nichols (and Gordon and Goldman) were being much too hasty in concluding that simulation theory puts eliminativism out of business.<sup>11</sup> And, in light of the distinctions we drew in the previous section, the reason should be obvious. Simulation theory suggests an account of the mechanisms underlying our capacity to predict and explain people’s behavior, and that explanation makes no appeal to an internalized theory or knowledge structure. So if simulation theory is correct, then *on the internal reading* there is no such thing as folk psychology. But simulation theorists do not deny the obvious fact that people have intuitions about folk psychological platitudes; nor do they deny that it might be possible to systematize those intuitions by constructing a theory that entails them. That theory would be a folk psychology *on the external reading*, and it might well turn out to be a seriously

---

<sup>10</sup> For similar arguments, see Gordon (1986), p. 170; Goldman (1989), p. 182.

<sup>11</sup> This constitutes a change in view for one of us (SS), prompted, in large measure, by the arguments urged by the other (IR). The argument set out in this paragraph was first proposed by Ravenscroft. At about the same time, much the same argument was proposed by Gerard O’Brien (forthcoming) and by Ian Hunt. We’ve also heard a similar argument from Kenneth Taylor. Stich wishes to thank all of these philosophers for convincing him that his previous view was mistaken.

mistaken theory. So the right conclusion to draw from the assumption that simulation theory is correct is not the jazzy claim that eliminativism has been undermined, but only the much more modest claim that eliminativists must opt for an external account of folk psychology.

Stephen Stich  
*Rutgers University*

Ian Ravenscroft  
*Australian National University*

---

## REFERENCES

- Chomsky, N. & Katz, J. (1974). "What The Linguist is Talking About," *Journal of Philosophy*, 71.
- Churchland, P.M. 1981. "Eliminative Materialism and the Propositional Attitudes," *Journal of Philosophy*, 78.
- Churchland, P.M. 1989. "Folk Psychology and the Explanation of Human Behavior," in Churchland, *A Neurocomputational Perspective*, Cambridge, MA: MIT Press.
- Churchland, P.S. 1986. *Neurophilosophy*, Cambridge, MA: Bradford Books / MIT Press.
- Fodor, J. (1968). "The Appeal to Tacit Knowledge in Psychological Explanation," *Journal of Philosophy*, 65. Reprinted in Fodor (1981). Page references are to Fodor (1981b).
- Fodor, J. (1981a). "Some Notes on What Linguistics is About," in N. Block, ed., *Readings in Philosophy of Psychology*, Cambridge, MA: Harvard University Press. 197-207.
- Fodor, J. (1981b), *Representations*, Cambridge, MA: Bradford Books/MIT Press.
- Fodor, J. (1983). *The Modularity of Mind*, Cambridge, MA: Bradford Books / MIT Press.

- Fodor, J. (1987). *Psychosemantics*, Cambridge, MA: Bradford Books/MIT Press.
- Fodor, J. (1990). "Roundtable Discussion," in P. Hanson, ed., *Information, Language and Cognition*, Vancouver: University of British Columbia Press.
- Goldman, A. (1989). "Interpretation Psychologized," *Mind and Language*, 4.
- Goldman, A. (1992). "In Defense of the Simulation Theory," *Mind and Language*, 7.
- Gordon, R. (1986). "Folk Psychology as Simulation," *Mind and Language*, 1, 158-171.
- Gordon, R. (1992). Replies to Stich and Nichols and Perner and Howes, *Mind and Language*, 7.
- Hacking, I. (1982). "Wittgenstein the Psychologist," *New York Review of Books*, 29, 5.
- Harre, R. (1986). *The Social Construction of Emotions*, Oxford: Basil Blackwell.
- Harris, P. (1992). "From Simulation to Folk Psychology: The Case for Development," *Mind and Language*, 7.
- Horgan, T. & Woodward, J. (1985). "Folk Psychology is Here to Stay," *Philosophical Review*, 94.
- Jackson, F. & Pettit, P. (1990). "In Defense of Folk Psychology," *Philosophical Studies*, 59.
- Johnson-Laird, P. 1983. *Mental Models: Towards a Cognitive Science of Language, Inference and Consciousness*, Cambridge, MA: Harvard University Press.
- Kitcher, P. (1984). "In Defense of Intentional Psychology," *Journal of Philosophy*, 81.
- Lewis, D. (1970). "How to Define Theoretical Terms," *Journal of Philosophy*, 67.
- Lewis, D. (1972). "Psychophysical and Theoretical Identifications," *Australasian Journal of Philosophy*, 50.
- McGinn, C. (1989). *Mental Content*, Oxford: Basil Blackwell.
- Needham, R. (1972). *Belief, Language and Experience*, Chicago: University of Chicago Press.

- O'Brien, G. (forthcoming). "A Conflation of Folk Psychologies," in K. Neander & I. Ravenscroft, eds., *Prospects for Intentionality*, Working Papers in Philosophy, 2, Canberra: Research School for Social Science, Australian National University.
- Ramsey, W., Stich, S. and Garon, J. (1990). "Connectionism, Eliminativism and the Future of Folk Psychology," *Philosophical Perspectives*, 4.
- Sellars, W. (1956). "Empiricism and the Philosophy of Mind," in H. Feigl & M. Scriven, eds., *Minnesota Studies in the Philosophy of Science*, 1, Minneapolis: University of Minnesota Press.
- Sharp, R. (1987). "The Very Idea of a Folk Psychology," *Inquiry*, 30.
- Soames, S. (1984). "Linguistics and Psychology," *Linguistics and Philosophy*, 7.
- Sterelny, K. (1990). *The Representational Theory of Mind*, Oxford: Basil Blackwell.
- Stich, S. (1972). "Grammar, Psychology and Indeterminacy," *Journal of Philosophy*, 69.
- Stich, S. (1983). *From Folk Psychology to Cognitive Science*, Cambridge, MA: Bradford Books / MIT Press.
- Stich, S. & Nichols, S. (1992). "Folk Psychology: Simulation or Tacit Theory?" *Mind and Language*, 7.
- Stich, S. & Nichols, S. (forthcoming). "Second Thoughts on Simulation," in M. Davies & T. Stone, eds., *Mental Simulation: Philosophical and Psychological Essays*, Oxford: Basil Blackwell.
- Von Eckardt, B. (1993). *What Is Cognitive Science?* Cambridge, MA: Bradford Books / MIT Press.
- Wilkes, K. (1981). "Functionalism, Psychology and the Philosophy of Mind," *Philosophical Topics*, 12.
- Wilkes, K. (1984). "Pragmatics in Science and Theory in Common Sense," *Inquiry*, 27.
- Wilkes, K. (1991). "The Relationship Between Scientific Psychology and Common Sense Psychology," *Synthese*, 89.