Mind & Language, ISSN: 0268-1064
Vol. 13. No. 4 December 1998, pp 499–512.

Rethinking Co-Cognition: A Reply to Heal

SHAUN NICHOLS AND STEPHEN STICH

1. Introduction

In cognitive science and philosophy of mind, there has been a wealth of fascinating work trying to tease out the cognitive mechanisms that are involved in understanding other minds or 'mindreading' (e.g. Baron-Cohen, 1995; Bartsch and Wellman, 1995; Fodor, 1995; Goldman, 1992; Gopnik, 1993; Harris, 1991; Leslie, 1991; Perner, 1991). This research has focused on evaluating the empirical evidence for various accounts of mindreading, predicting the results of future experiments, and carrying out experiments that might distinguish between the available theories. Our own previous work adopted this naturalistic approach (Stich and Nichols, 1992, 1995, 1997; Nichols et al., 1996; Nichols et al., 1995). In contrast to the naturalistic exploration of mindreading, Jane Heal has argued that simulation theorists have discovered an a priori truth about mindreading (Heal, 1994, 1995). In Heal's most recent paper (this issue), which is largely a response to an earlier paper of ours (Stich and Nichols, 1997), she maintains that we are committed to a view that conflicts with a simulationist thesis which is a priori true. In this paper we'll argue that this accusation is deeply muddled. Heal's putative a priori truth is so vague that it admits of many interpretations, and on some interpretations the claim does indeed conflict with our views. Unfortunately for Heal, on these interpretations the 'a priori truth' is simply false. On another, much weaker, reading, the claim is clearly true—but it is also completely uncontroversial. On that reading, it is quite preposterous to suggest that our view (or anyone else's) conflicts with the claim.

Here is how we propose to proceed. In the next section, we will review some of the notable agreements that have emerged in the ongoing dialogue between Heal and ourselves. In the subsequent section, we will present Heal's 'co-cognition' thesis and offer a friendly amendment to part of her

Address for correspondence: Shaun Nichols, Center for Cognitive Science, Psychology Building Addition, 152 Frelinghuysen Road, Busch Campus, Rutgers University, Piscataway, NJ 08854–8020, USA.
Email: nichols@cofc.edu, stephen-stich@email.msn.com.

story. Then, in section 4, we will present our critique of the co-cognition thesis.

2. *Points of Agreement*

Before we focus on Heal's vague and contentious co-cognition thesis, we'd like to put the current dispute in perspective by noting some points of agreement that have been achieved in our earlier debates with Heal (Heal, 1996, this issue; Stich and Nichols, 1997). First, in response to Heal (1996), we agreed that cognitive penetrability is not a useful tool for deciding among accounts of the mechanisms underlying mindreading, because mechanisms of the sort advocated by simulation theorists can be affected in various ways by what agents know or believe (Stich and Nichols, 1997, p. 315). Although the notion of cognitive penetrability now seems too coarse to assess simulationist hypotheses, we still maintain that the experimental results we've adduced pose a challenge for simulation theorists. They must give some account of how the sorts of mental mechanisms they posit can explain these results (see, for example, Nichols et al., 1996; Nichols et al., 1995). Heal (1996) agrees and also concedes that it's unlikely that simulation theorists can explain the results by appealing to the sort of cognitive penetration that might affect the simulation process.[1]

In Heal's current paper yet another point of agreement emerges. In Stich and Nichols (1997), we argued that the term 'simulation' needs to be retired. For 'the diversity among the theories, processes and mechanisms to which advocates of simulation theory have attached the label "simulation" is so great that the term itself has become quite useless. It picks out no natural or theoretically interesting category' (Stich and Nichols, 1997, p. 299). We were delighted to see that, in contrast with some other 'simulationists' (see, for example, Currie and Ravenscroft, 1997; Goldman, forthcoming), Heal is very much in agreement with us (Heal, this issue, p. 496). This is, we think, a very important point. Future discussion of 'simulation theory' will only be fruitful if the participants are sensitive to the distinctions among different simulation proposals. For, as we argued previously, some 'simulation' proposals are obviously right and others are quite controversial. Nothing but confusion can result from the expectation that they all stand or fall together. Productive debate can only result from a more sophisticated discussion that

[1] Heal has urged that simulationists can explain these data by restricting the purview of their theories to what she calls 'rational' cognitive processes. And this is a claim that remains very much in dispute. We have argued that this strategy is utterly unconvincing since if rationality is assessed by any plausible intuitive standard it does not separate the cases in which people's predictions succeed from those in which their predictions fail. Of course, Heal may have some more technical, unintuitive notion of rationality in mind. But if so, she has not bothered to explain it, and until she does her proposal (interpreted in this way) is impossible to evaluate (see Stich and Nichols, 1997).

focuses on the specific and substantive proposals. Attempting to score points by asserting that one or another sort of mechanism or process really is a 'kind of simulation' is, we think, an utterly pointless exercise.

In 'Co-Cognition and Off-Line Simulation' (this issue),[2] Heal discusses another way in which the term 'simulation' has become useless and misleading. There are, she argues, fundamentally different ways of understanding the 'simulationist' idea, and the use of a single term leads to debates at cross-purposes. Here is how she puts the point:

> The central aim of this paper is to articulate and recommend the idea that the simulationist idea about what is involved in grasp and use of psychological concepts may be understood in two different ways, namely as an a priori claim about the relations of certain personal-level cognitive abilities or as an a posteriori hypothesis about the workings of sub-personal cognitive machinery. (p. 477)

Heal makes the distinction between these two claims explicit in her terminology. She writes, 'I shall use 'off-line simulation' to talk of the idea which is the focus of the a posteriori hypothesizing about what is implicated in thinking about others' thoughts; and I shall coin a new term, 'co-cognition', to talk of what the a priori claim says to be central to such thinking' (p. 478).

Here again, there is much that we can agree with. We agree that the a priori 'simulationist' claim is quite distinct from the various empirical claims made by simulationists and that nothing but confusion can result from failing to note the distinction. We would also note, for the record, that in all our previous work on this topic, our focus has been on the empirical debate. Our central concern has been to understand the cognitive mechanisms and processes that underlie mindreading. In this paper, however, our focus will be on Heal's a priori claim. To avoid confusion with other 'simulationist' theses, we will adopt Heal's terminology and refer to the claim as the (a priori) co-cognition thesis.

3. The Co-Cognition Thesis and a Friendly Amendment

Although much of the literature on mindreading attempts to provide a broad account of how we predict mental states, Heal sensibly focuses her attention on a narrower range of abilities. She confines her discussion to 'what is involved in our arriving at further psychological judgements about others, given information about some of their existing psychological states' (p. 479). The issue at hand, then, is how we move from information about another's thoughts to predictions about that person's further thoughts. In order to do this, Heal claims, we have to think about the same subject matter or 'co-

[2] All quotes are from this paper unless otherwise indicated.

cognize' with the target. Heal also offers an explanation for why this 'co-cognition' strategy is successful. In this section, we will sketch Heal's co-cognition thesis and her explanation for the success of co-cognition. We will then argue that her explanation for why the co-cognition strategy is successful is confused, but we will offer, as a friendly amendment, an alternative explanation for the success of the co-cognition strategy.

Heal maintains that in drawing inferences from thoughts to thoughts we rely on 'co-cognition', which is 'just a fancy name for the everyday notion of thinking about the same subject matter' (p. 483). Indeed, Heal claims that it is an a priori truth that co-cognition is *required* for this sort of mindreading. She offers the following thesis:

> (A) It is an a priori truth that thinking about others' thoughts requires us, in usual and central cases, to think about the states of affairs which are the subject matter of those thoughts, i.e. to co-cognize with the person whose thoughts we seek to grasp. (p. 484)

We think that this thesis is quite problematic, but, before we get down to our critique, we would like to consider her analysis of why the co-cognition strategy works.

To explain the success of co-cognition, Heal appeals to the claim that we are 'rational' in the sense that our thoughts track the truth. She links truth to rationality as follows: 'The idea that connections in thought follow connections between states of affairs is the idea that we are rational' (p. 486). Heal then claims that the success of co-cognition presupposes some such assumption of rationality. She writes:

> Given the assumption of such very minimal rationality [i.e., truth tracking], we can show why reliance on co-cognition is a sensible way to proceed in trying to grasp where another's reflections may lead. The other thinks that p_1—p_n and is wondering whether q. I would like to know what will she conclude. Her thoughts (I assume) will follow the connections between things. (p. 487)

We think Heal's attempt to explain the success of co-cognition by linking truth, rationality and co-cognition is muddled and obscures what is really going on. First, let's focus on the connection between rationality and truth tracking. Here the dispute may be largely terminological. It certainly seems that if 'rational' is used in an ordinary, non-technical sense, then we can be highly rational even though the connections in our thought *do not* follow connections between states of affairs. To see the point, we need only think about Descartes's evil genie who misleads his victims into an utterly mistaken view of the world. The more rational the victim is, the less likely it is that her thoughts will follow the connections between states of affairs. To be fair, however, Heal is careful to say that 'The idea that connections in thought follow connections in states of affairs is the idea that we are

rational—*in some sense of that term*' (p. 486, emphasis added). And, while we are dubious that there is really any *common* sense of the term that fits Heal's description, we also think it is silly to debate such matters.

Next, and much more important, let's focus on the link between truth tracking and co-cognition. As far as we can see, the idea that our thoughts are true—that they track reality—is really quite irrelevant to Heal's basic account of co-cognition. To see the point, it is useful to think about the case of two people, Pam and Sam, who share what is in fact (and what we ourselves take to be) an utterly benighted, mistaken, false theory about some particular subject matter. To be concrete, let us assume that they share some odd and deeply mistaken body of religious belief and theory. Now, as far as we can see, the fact that their theory is profoundly mistaken and that 'connections in [their] thought' *do not* 'follow connections between states of affairs' will not in the least interfere with Pam's ability to reason from the fact that Sam believes $p_1 - p_n$ to the conclusion that Sam will also believe q, where $p_1 - p_n$ and q are religious claims couched in the language of the thoroughly false theory that they share. Moreover, this example does not constitute a challenge to Heal's thesis A, according to which 'thinking about others' thoughts requires us, in usual and central cases, to think about the states of affairs which are the subject matter of those thoughts' (p. 484). For it is certainly possible, perhaps even plausible, that when Pam thinks about Sam's thoughts she thinks about the subject matter of those thoughts—the gods, mysterious powers and strange rituals that they both believe in.

Now, no doubt, Pam and Sam both take themselves to be rational in Heal's sense. They think that 'connections in [their] thought tend to mirror connections between states of affairs' (p. 486). But they are wrong, at least when it comes to their religious thoughts. And the fact that they are wrong—that they are not Heal-Rational when it comes to religion—poses no problem at all for their ability to 'mindread' one another, i.e. to infer from thoughts to thoughts. The point we are trying to bring out here is that Heal's emphasis on what we've called 'truth tracking', the presupposition that connections in thought tend to mirror connections between states of affairs, is really doing no work in her account. So long as Pam and Sam are co-cognizers, their ability to draw inferences (generally quite good inferences) from thoughts to thoughts will proceed just fine.

In suggesting that successful co-cognition doesn't require truth tracking, we mean only to offer a friendly amendment to Heal's characterization of co-cognition. This amendment doesn't pose a real threat to Heal's co-cognition thesis. If there is to be a real challenge to Heal's proposed a priori truth, it will come from considering cases in which people attempting to 'read' each other's minds are *not* co-cognizers. But before turning to such a case, we want to note in passing the way in which the points we've just made echo a discussion that took place in the literature twenty years ago, back in the days of what Dennett's *Lexicon* refers to as the Davidsonic Boom (Dennett, 1987a). In his early papers on interpretation, Davidson proposed a version of the Principle of Charity which required that we assume that

most of the beliefs of the person we are interpreting be *true* (e.g. Davidson, 1967). But in later papers he often added an important qualification. The Principle doesn't require that the target's beliefs be *true* (full stop), but that they be *true by our lights*—i.e. that by and large they believe what we believe (e.g. Davidson, 1973). There are, of course, important differences between Davidson and Heal, both in their views and in their projects. What they share is the idea that the process of attributing mental states to others presupposes that the target and the attributer *share* important aspects of their mental life.

4. A Critique of the Co-Cognition Thesis

In this section, we want to explore the co-cognition thesis more critically. In Heal's exposition of the thesis, proposition (A) plays a central role so it might be useful to quote it a second time:

> (A) It is an a priori truth that thinking about others' thoughts requires us, in usual and central cases, to think about the states of affairs which are the subject matter of those thoughts, i.e. to co-cognize with the person whose thoughts we seek to grasp.

We maintain that, in light of what Heal says about co-cognition, (A) is crucially vague along three quite distinct dimensions.

(1) It's not clear what 'requires' is supposed to mean.
(2) It's not clear *how much* one needs to know (or think) about the subject matter in order to co-cognize.
(3) It's not clear which cases are the 'usual and central ones'.

We will consider three different ways of resolving the vagueness and making the claim more precise. On two of these interpretations, we will argue that the claim is just plain false. Thinking about others' thoughts does not require co-cognition. On the third interpretation we offer, the co-cognition thesis is true. However, on this reading it is a claim that we have never disagreed with, nor, as best we can tell, has anyone else.

4.1 Interpretation 1

First let's assume that when Heal says 'thinking about others' thoughts *requires us* to co-cognize' she means that you *can't* think about others' thoughts unless you think about the same subject matter. Now, the question is, when does one satisfy the condition of thinking about the same subject matter? Well, it depends on how you individuate subject matters. Heal seems to address this issue by introducing a distinction between subject matters that are 'independent' of each other and subject matters that aren't.

To get a sharper idea of when two people count as thinking about the same subject matter, let's look at Heal's characterization of independent subject matters. She writes,

> two subject matters are independent of each other when the principles of classification, knowledge and conceptual skills relevant to dealing with the one are of little help in dealing with the other and vice versa and when, in consequence, a person may have rich and adequate grasp of one subject matter but an exceedingly minimal grasp on the other and vice versa. For an example, consider as subject matters vegetables on the one hand and stocks and shares on the other. A person brought up in the country might have a great deal of information about vegetables, their varieties, patterns of growth, required nutrients, soil types, climate etc. while knowing very little of stocks and shares. Conversely, a person brought up entirely in a built environment might have excellent grasp of stocks and shares, their varieties, legal complexities, profitability and so forth, while being extremely ignorant of the vegetable world. (p. 481)

Since the subject matters of vegetables and stocks are independent, the country boy who knows a great deal about vegetables but is extremely ignorant of stocks cannot co-cognize with the city boy's thoughts about stocks. More specifically, although the country boy believes (among other things) that potatoes are tubers, he cannot co-cognize with the city boy's thought that if there is a sell-off in blue chip stocks then bond prices will rise. For the country boy doesn't know about the subject matter of stocks, independent as it is from the subject matter of vegetables. Now, as Heal notes, even vegetables and stocks are not entirely independent subject matters. However, a rich grasp of one is 'independent of all but minimal grasp' of the other (p. 481). Thus, we suggest as interpretation 1, that according to Heal's thesis (A) it is impossible to think about another's thoughts unless one has some non-minimal grasp on the subject matter of the other's thoughts.

Is this true? To address the question, let's elaborate a case in which the mindreader and the target are patently *not* co-cognizers in the current sense. The case of Pam and Jack will do nicely. Pam, recall, is deeply involved in an exotic New Age religion. She has lots of beliefs which she expresses using terms like 'karma', 'nsisim', 'kong', 'bekong', and 'Clear'. Jack, on the other hand, is a devout atheist who thinks that all religion is nonsense and that New Age religions are nonsense on stilts. But he is, near enough, completely ignorant about Pam's religion. He couldn't even begin to explain the difference between a kong and a nsisim. He certainly has no more than the most minimal grasp of the subject matter. At any rate, his grasp on Pam's religion is certainly no greater than the country boy's grasp of stocks and shares. When it comes to religion, then, Jack and Pam are not co-cognizers. How well could Jack do at thinking about Pam's thoughts?

At first blush it looks like the answer is: not very well at all. And this, of course, is just what Heal's a priori principle would lead us to expect. Suppose, for example, that Sam tells Jack something about Pam's current beliefs:

Pam believes

> p1: that Sam's karma has been prigilated to his nsisim
> p2: that Sam's kong has been debinating him from a hostile bekong

and

> p3: that the bekong is trying to undermine Sam's Clear.

And then Sam raises a question for Jack:

> Q: Does Pam also believe that Sam will lose his Clear?

Not surprisingly, Jack hasn't a clue. He has no idea whether Pam believes that Sam will lose his Clear. And this, note, is just what Heal's a priori principle would predict. Jack *can't* think about Pam's thoughts about religion because he can't think about the subject matter of those thoughts. When it comes to religion, Jack and Pam can't co-cognize.

But we are inclined to think that there is a better explanation for Jack's inability to answer Sam's question. The problem is not that Jack can't co-cognize with Pam, but rather that Sam hasn't told Jack enough about Pam's beliefs. What Jack needs to answer Sam's question is not the ability to co-cognize, but simply *more* information about Pam's beliefs. So, for example, suppose that after Sam poses his question about Pam's beliefs, Tam tells Jack about some more of Pam's beliefs:

Pam believes

> p4: that when someone's karma has been prigilated to their nsisim they are no longer karmafied
> p5: that kongs don't debinate those who are not karmafied
> p6: that people in Sam's Klingon can only be debinated by their kong

and

> p7: that unless a person is debinated hostile bekongs will succeed in doing what they are trying to do.

Now, with this further information, Jack might well infer an answer to Q:

> p8: Pam believes that Sam will lose his Clear.

And he might well be right.

If all of this is correct, then it seems to pose a problem for our first interpretation of Heal's a priori principle. For after Tam tells Jack more about Pam's beliefs, Jack *can* think about Pam's thoughts in just the way that is central in Heal's discussion. He can reason from the fact that Pam believes $p_1 - p_n$ to the conclusion that she will also believe q. But despite this, Jack and Pam still can't co-cognize. Jack not only *doesn't* think that when someone's karma has been prigilated to their nsisim they are no longer karmafied. He *can't* think it. As far as he as concerned the sentence

> When someone's karma has been prigilated to their nsisim they are no longer karmafied

is utter nonsense. The reason that Jack could not answer Sam's question about Pam's thoughts prior to getting additional input from Tam was *not* that Jack and Pam could not co-cognize, but rather that Jack just didn't know *enough* about Pam's thoughts. So the lesson to be learned from this example is not that mindreading requires co-cognizing, but that successful mindreading requires knowing *a lot* about the target's thoughts, even if you can't think them yourself.

4.2 Interpretation 2

In our example of Jack and Pam, the belief attributions do include some terms that we (and Jack) know the meaning of. For instance, $p_1 - p_7$ contain logical connectives. Thus, one way to respond to our case is by saying that there is still *some* co-cognition involved in the case—via the logical connectives. Indeed, this *is* the way Heal responded when we initially presented the case to her, and her final draft reflects this. She now writes that 'At one limit N may be able to co-cognize with M only in the very skeletal sense of grasping the logical form of M's thought' (p. 483). Heal thus might maintain that Jack *does* co-cognize with Pam, in virtue of thinking about the same logical form as Pam. This suggests a second interpretation of the co-cognition thesis. On this interpretation, the *slightest grasp* of the subject matter, including grasp of logical form, suffices for co-cognition. Let us continue to suppose, on this interpretation, that (A) claims one *can't* think about others' thoughts without co-cognition. So, on interpretation 2, thinking about another's thoughts is impossible without co-cognition, but co-cognition turns out to be a lot easier than Heal's example of stocks and vegetables might lead one to expect; it occurs whenever we grasp anything at all about the subject matter of the other's thoughts.

Even this much weaker version of the co-cognition thesis can be seen to be wrong if we simply modify our above example. For even the minimal traces of co-cognition can be eliminated if we elaborate the case along the lines of Searle's Chinese Room example (Searle, 1980)—albeit with a very different goal in mind. Consider the case of Cam, an adherent of a distinct

sect of Pam's New Age religion. Sam tells Jack that in that the language of that sect '*' (pronounced *blip*) and '%' (pronounced *blap*) are connectives of some sort, and when members of the sect believe something that they express with 'p * q' for more than three days (or during a full moon) they also come to believe something they would express with '%q'. Sam also tells Jack that for the last several weeks, Cam has had a belief which he expresses with the sentence:

p9: karma kong prigilate * kong Clear debinate.

From this information, Jack might well infer that Cam has a belief which he would express with the sentence:

p10: %kong Clear debinate.

In this case, Jack has absolutely no idea what any of the beliefs mean. He can't even co-cognize the connectives, which are utterly foreign to him. Nonetheless, as in the previous case with Pam, it seems that Jack might figure out quite a lot about Cam's further beliefs from the sort of information that Sam provides. Hence, contra the second interpretation of Heal's co-cognition thesis, it's possible to think about another's thoughts without even the slightest bit of co-cognizing. Mindreading does not require co-cognition.

4.3 Interpretation 3

In both of the previous interpretations, we assumed that on Heal's view, it is *impossible* to think about another's thoughts without co-cognition. The cases we've presented show that co-cognition, even in the attenuated version of interpretation 2, is not *required* for thinking about another's thoughts. If the relevant facts about the target's beliefs, and their patterns of causal interaction, are available to us, we can think about their thoughts in just the way that Heal focuses on—we can arrive 'at further psychological judgements about others, given information about some of their existing psychological states' (p. 479)—even if the subject matter of their thoughts is entirely unintelligible to us. But perhaps Heal's thesis is not intended to claim that it's *impossible* to think about another's thoughts without co-cognition. When Heal claims that thinking about another's thoughts 'requires us, in the usual and central cases' to use co-cognition (p. 484), perhaps she only means that we would be quite at sea if we could not use co-cognition in trying to mind-read. This suggests a third interpretation of the co-cognition thesis: in thinking about others' thoughts, we typically use co-cognition and if we could not rely on it, the ordinary mindreading that we employ in our daily interactions with one another would be severely disrupted. This interpretation seems to fit with Heal's characterization of the a priori. 'An a priori claim', she writes, 'is one we rely on unhesitatingly in making inferences; in cases where it seems threatened our automatic assumption is that the threat is

illusory and we seek ways of explaining it away; if challenged we are thoroughly at a loss to describe realistically or in any detail how we would carry on intellectually if we could not rely on it' (p. 480). Hence, perhaps Heal's proposed a priori claim is just that co-cognition plays a crucial role in everyday mindreading.

If this interpretation of the co-cognition thesis is what Heal has in mind, then we are in complete agreement with her. It's obvious that we routinely use co-cognition in everyday life, and that typically we'd be lost without co-cognition. Indeed this is so obvious that we wonder who is supposed to disagree with it. According to Heal, the viewed is opposed by those who hold the 'strong theory-theory'—the view that 'thoughts about X form a separate subject matter which is independent of the subject matter X' (p. 485).

Now we don't deny that a strong theory-theorist might reject the co-cognition thesis even on the third interpretation. But who is supposed to *hold* the strong theory-theory? The only explicit target Heal cites is one of our papers (Stitch and Nichols, 1995). She suggests that since Figure 1 in Stich and Nichols (1995) has separate boxes for folk psychology and folk physics, we are committed to strong theory-theory. Heal writes, 'If one were alive to the possibility that in thinking about another's thoughts about physical objects a person would have to call upon his or her knowledge of physical objects, it would be distinctly misleading to configure the diagram as Stich and Nichols do' (note 4). We are, we confess, simply flabbergasted by this reading of our view. For surely a much more sensible interpretation of Figure 1 is simply that there is a body of information about psychology that is distinct from a body of information about physics. That doesn't entail (or even suggest) that all folk psychological predictions and attributions derive solely from the 'folk psychology box'.[3] Moreover, even if our diagram admits of several interpretations, the text that surrounds it includes a number of quite explicit passages that are obviously incompatible with the strong theory-theory. In one of them we agreed with Paul Harris (1995) that the prediction of other people's grammaticality judgements plausibly exploits the predictor's own tacit grammatical knowledge (Stich and Nichols, 1995, p. 93), and obviously that knowledge is not stored in the folk psychology box. We then went on to give further examples along these lines:

> Suppose . . . that we were asked to predict what one of our Rutgers colleagues would say when asked: 'Who is the President of Rutgers University?' Or suppose we were asked to predict what our wives

[3] It's even possible that some of the principles in the folk psychology data-base make explicit that one should exploit other mechanisms. For instance, the folk psychological theory might contain a heuristic that says that I should assume that other people believe what I believe (at least in certain domains). Or perhaps there's a heuristic that says that I should assume that other people believe what is true. On either of these stories, the heuristic in the folk psychology box would call for using the folk physics box to assist in mindreading.

> would say when asked: 'Who was the third President of the United States?' In both cases, we suspect, we would proceed by first answering the question for ourselves—recalling who *we* think is the President of Rutgers or who we think was the third President of the US. Then, since we assume that our colleagues (in the first case) and our wives (in the second case) believe the same things we do on questions like this, we would predict that they would say the same thing we would (Stich and Nichols, 1995, p. 93).

Obviously our knowledge of Presidents is outside the folk psychology box, but we explicitly claimed that we can and do use this knowledge to make attributions to others. Since that paper, we have continued to be quite explicit that on our view people's understanding of other minds is not independent of their own beliefs and inferences. For instance, in our previous response to Heal, we wrote, 'Suppose Stich knows that Nichols has paper and pencil handy and is about to add 123 + 456 + 789. If Stich wants to predict what Nichols will believe the answer is, the obvious strategy is for Stich to reach for a pencil and do the addition himself' (Stich and Nichols, 1997, pp. 300–301). We even claimed that it's difficult to elaborate any 'theory-theory' story about mindreading that doesn't involve tapping into one's own beliefs about various subject matters. 'Indeed', we wrote, 'the hypothesis that people use their own beliefs as the default value in assigning beliefs to a target is just about the only serious proposal we know of for how this sort of tacit-theory-driven account of belief prediction could be elaborated' (p. 309). If we are the most likely suspects for being strong theory-theorists, then we doubt that there are any strong theory-theorists out there.[4] The strong theory-theorist is a straw man, a figment of Heal's

---

[4] As far as we can discover, no one has *ever* endorsed the singularly implausible doctrine that Heal calls 'the strong theory-theory'. On the contrary, several prominent theory-theorists have been fairly explicit about rejecting the view. For instance, Hartry Field (who is arguably one of the first theory-theorists) maintains that logic plays a role in drawing inferences from thoughts to thoughts. He suggests that when we know about the target's explicit beliefs, we attribute beliefs that are 'obvious consequences' of those beliefs. He writes that

> This kind of account would make clear why someone who believes that either Russell was hairless or snow is white would almost certainly also believe that if Russell was not hairless then snow is white: almost any stock of core beliefs from which the former was an obvious consequence would also be a stock of core beliefs from which the latter was an obvious consequence. (Field, 1981, pp. 83–4)

Daniel Dennett, who defends an account of mental state attribution quite different from the one that Field advocates, also has made claims that are clearly incompatible with the strong theory-theory. He maintains that we attribute beliefs according to the following principles: 'A system's beliefs are those it *ought to have*, given its perceptual capacities, its epistemic needs, and its biography. Thus, in general, its beliefs are both true and relevant to its life, and when false beliefs are attributed, special stories must be told to explain how the error resulted' (Dennett, 1987b, p. 49). So, on Dennett's view

imagination. So if we interpret Heal's co-cognition thesis in the third way, the a priori strand of simulationism turns out to be a banal truth that no one has ever questioned.

Department of Philosophy
The College of Charleston

Department of Philosophy
Rutgers University

References

Baron-Cohen, S. 1995: *Mindblindness: An Essay on Autism and Theory of Mind*. Cambridge, MA: MIT Press.
Bartsch, K. and Wellman, H. 1995: *Children Talk about the Mind*. Oxford University Press.
Currie, G. and Ravenscroft, I. 1997: Mental Simulation and Motor Imagery. *Philosophy of Science*, 64, 161–80.
Davidson, D. 1967: Truth and Meaning. *Synthese*, 17, 304–23.
Davidson, D. 1973: Radical Interpretation. *Dialectica*, 27, 313–28.
Davies, M. and Stone, T. (eds) 1995: *Mental Simulation*. Oxford: Basil Blackwell.
Dennett, D. (ed.) 1987a: *The Philosophical Lexicon*, 8th ed. Newark, DE: The American Philosophical Association.
Dennett, D. 1987b: Three Kinds of Intentional Psychology. In *The Intentional Stance*. Cambridge, MA: MIT Press.
Field, H. 1981: Mental Representation. In N. Block (ed.), *Readings in the Philosophy of Psychology*, vol. 2, 78–114.
Fodor, J. 1995: A Theory of the Child's Theory of Mind. In Davies and Stone (eds) 1995.
Goldman, A. 1992: Empathy, Mind, and Morals. *Proceedings and Addresses of the American Philosophical Association*, 66, no. 3, 17–41.
Goldman, A. forthcoming. The Mentalizing Folk. In D. Sperber (ed.), *Metarepresentation*, Oxford University Press.
Gopnik, A. 1993: How We Know Our Own Minds: The Illusion of First-Person Knowledge of Intentionality. *Behavioral and Brain Sciences*, 16, 1–14.
Harris, P. 1991: The Work of the Imagination. In A. Whiten (ed.), *Natural Theories of Mind*. Oxford: Blackwell.
Harris, P. 1995: From Simulation to Folk Psychology: The Case for Development. In M. Davies and T. Stone (eds), *Folk Psychology*. Oxford: Blackwell, 207–31.
Heal, J. 1994: Simulation vs. Theory Theory: What Is at Issue? In C. Peacocke (ed.), *Objectivity, Simulation and the Unity of Consciousness*. Oxford University Press, 129–44.
Heal, J. 1995: How to Think about Thinking. In Davies and Stone (eds) 1995, 33–52.

---

we must make broad use of our own knowledge of subject matters in attributing beliefs to others.

Heal, J. 1996: Simulation and Cognitive Penetrability. *Mind and Language*, 11, 44–67.

Leslie, A. 1991: The Theory of Mind Impairment in Autism: Evidence for a Modular Mechanism of Development? In A. Whiten (ed.), *Natural Theories of Mind*. Oxford: Blackwell.

Nichols, S., Stich, S. and Leslie, A. 1995: Choice Effects and the Ineffectiveness of Simulation: Response to Kuhberger et al. *Mind and Language*, 10, 437–45.

Nichols, S., Stich, S., Leslie, A. and Klein, D. 1996: Varieties of Off-Line Simulation. In P. Carruthers and P. Smith (eds) *Theories of Theories of Mind*. Cambridge University Press, 39–74.

Perner, J. 1991: *Understanding the Representational Mind*. Cambridge, MA: MIT Press.

Searle, J. 1980: Minds, Brains and Computers. *Behavioral and Brain Sciences*, 3, 417–57.

Stich, S. and Nichols, S. 1992: Folk Psychology: Simulation or Tacit Theory? *Mind and Language*, vol. 7, no. 1, 35–71.

Stich, S. and Nichols, S. 1995: Second Thoughts on Simulation. In Davies and Stone (eds) 1995, 87–108.

Stich, S. and Nichols, S. 1997: Cognitive Penetrability, Rationality, and Restricted Simulation. *Mind and Language*, 12, 297–326.