Journal for the Theory of Social Behaviour 29:4
0021–8308

FIC Descriptions and Interpretive Social Science: Should Philosophers Roll their Eyes?

TODD JONES

INTRODUCTION

There is a type of explanation one finds throughout the social sciences and the popular media that, as a philosopher, I find both fascinating and exasperating. I will call such account 'FIC descriptions.' FIC descriptions are Folk psychological, Interpretive and Collectivist descriptions of human thought and behavior. They attempt to explain what is going on in recent or not-so-recent history by interpreting the hidden *beliefs and desires* held by *groups* which determine what unfolds in a culture. Here's a statement, for example, that caught my eye in a recent *New York Times*.

> Among the elements that compose China's complicated psyche today, two stand out: its insular self-satisfied attitude, rooted in the ancient Chinese belief that the nation lay at the center of the world; and the deep sense of inferiority that came with the modern realization that China was actually far behind most of the developed world (Faison, 1997, p. 5).

Here's another example, one from the historian Geoffrey Gorer, attempting to explain American egalitarianism:

> Just as the colonists had rebelled against paternalistic Britain, and the second generation immigrants had escaped the patriarchal culture of Europe, so all American children were taught first to reject then surpass their fathers. The result was a permissive, freewheeling anti-authoritarian society, egalitarian in its human relationships if not in its distribution of wealth (Pells, 1997, p. 171).

I'm calling these sorts of accounts FIC descriptions partly to point out their family resemblance to an approach to social science that the anthropologist Clifford Geertz labeled 'thick description' that has become immensely influential in anthropology, history, and other social science disciplines. Ben Colby and his colleagues described the intellectual parentage of this style of social science as lying in

psychoanalysis, on the one hand, and in the sociology of knowledge, on the other. The primal documents are Freud's (1965 [1899]) The Interpretation of Dreams, with its emphasis on personally generated symbolism, and Durkheim's (1961 [1923]) The Elementary Forms of Religious Life, with its emphasis on collective representations and systems of classification and correspondence (Colby et al., 1981, p. 424).

The sorts of social explanations I'll be discussing here, which bring together a psychoanalytic-like approach emphasizing the deeper, darker, symbolic roots of social behavior, and a Durkheimian or even Hegelian emphasis on a collective zeitgeist, are found throughout the social sciences and popular media. The vast *pervasiveness* of these sorts of interpretivist and collectivist explanations is one of the main reasons that it's worth taking some time to examine the plausibility and sensibility of such statements. The other main reason they are worth looking at in detail is that *despite* the pervasiveness of these sorts of accounts among some social thinkers, there's also vast numbers of other thinkers, who are quite willing to *dismiss* these accounts out of hand as just plain silly – to roll their eyes derisively whenever someone starts to give this sort of description.

What I want to do in this paper is to try to sort out what's sensible and what's confused in these quasi-psychoanalytic neo-Hegelian accounts we so often see around us. Even if one isn't a fire and brimstone positivist, there are clearly some features of FIC descriptions that should give us pause. Accounts like these go considerably beyond our usual belief-ascribing practices in two ways. First, while beliefs and desires are prototypically mental states of individual human beings, such descriptions seem to ascribe beliefs, desires and other mental attributes to entire groups. Thus you have historian Simon Schama saying things like "it is not too much to say that *classical civilization* has always defined *itself* against the primeval woods" (1995, p. 82; emphasis mine). Or you have social critic Barbara Ehrenreich writing in the eighties that "All that is happening is that our collective values are shifting away from the liberal, unisex ideals of the seventies toward something more belligerent. The national wimp hunt, I have concluded, is an attempt to press men into line for the postdetente militarism of the eighties ..." (1985, p. 139).

In addition to their ascription of beliefs to entire groups, such ascriptions are also problematic in that they go well beyond the usual beliefs that we can safely assume that anyone with the normal cognitive equipment and a given cultural background will hold. If we want to say with Robert Jay Lifton that modern man can be characterized by his "suspicion of counterfeit nurturance" (1975, p. 326) we're going to need special access to peoples' "hidden" beliefs and desires. It's possible that historians, journalists, and anthropologists know how to uncover these hidden states better than ordinary folk can, but it's not obvious they do. For this reason we should be somewhat wary about uncritically accepting FIC description claims as easily as we tend to accept ascriptions like "Dan believes the car is close to the curb." On the other hand, such ascriptions are

extremely common and often at least seem to give us a great deal of insight into what's going on, so we shouldn't dismiss them out of hand either. What I'd like to do in this paper is look carefully at the metaphysics and epistemology of what I'm calling FIC descriptions and see if there's anything systematic to say about what's sensible and what's muddled about them.

COLLECTIVISM

Perhaps the simplest place to start is to consider whether we really might have good reasons to dismiss all such accounts as nonsense in one fell swoop. One way we might try to dismiss such claims is to say that all FIC descriptions claims are suspect because they ascribe beliefs to groups, when only *individual* organisms can really have beliefs and desires. Are there any arguments which could justify such a quick dismissal?

To decide if there are, one thing we need to acknowledge is that without more information collective-belief claims tend to be extremely ambiguous. Suppose the assistant dean says to the dean, "The philosophy department thinks they'd be better off hiring two new junior faculty than one senior faculty member." There's several things she might mean by this statement. She might mean that the majority of the individuals in the philosophy department believe this. She might mean that a controlling block have this belief, even though a numerical majority has a range of different opinions. It's possible that she's thinking merely that the department chair or the department's most distinguished professor thinks this, and to her those wants are synonymous with what "the department" thinks or wants.

Finally, someone making this claim might be talking as if the group as a whole is some sort of agent, a person-like entity in its own right. It's this unified holistic entity, and not the people in it that has that belief. The dean might say to the assistant dean, "the philosophy department *still* wants to add two faculty?" while not thinking of anyone in the philosophy department. He remembers they've been asking for two positions for years, with various chairs and varying personnel. Here the dean might well be thinking of "the department" as some sort of entity in its own right.

Now this ambiguity of our collective ascriptions can certainly be a source of misunderstandings, but it's also a source of utility for such descriptions. With so many potential meanings, one can't easily claim that a given group belief statement is categorically untrue or confused unless one's really ready to say that *each* of its possible parsings is untrue.

Can anything be said right away about which parsings tend to be systematically informative and important? One way that group belief ascriptions are routinely used in what I think is an undeniably informative way are when such descriptions are made as a species of statistical syllogisms aimed at telling you about

individuals in the group. A statistical syllogism is a way of making an inductive inference about individuals based on knowledge of the group that the individual is a member of. Statistical syllogisms follow the generic schema:

X percent of Fs are G
C is an F
So C is probably a G.

Sometimes what we are saying with group belief ascriptions is:

X percent of social group F believe G (usually a majority of a certain size).
A given member is a member of social group F.
So there's a good chance he or she believes G. (The higher the percentage the better the chance.)

No historians, sociologists, and journalists certainly ought to be more explicit about saying exactly what they have in mind when they make a group belief claim. But when what they have in mind are statistical syllogisms, then those claims can easily be quite legitimate and informative. When someone says "the French think that wine goes better with meals than beer," for example, that can be taken as a perfectly appropriate way to express the fact that any given French person you'll encounter is more likely than not to say that wine goes better with meals than beer.

So the fact that some group belief ascriptions should be read as legitimate statistical syllogisms is one reason that we shouldn't automatically dismiss FIC description claims. But clearly sometimes such claims are made when the speaker doesn't have any reason to believe that the majority of individuals believe X. What can be said about non-statistical syllogism collectivist claims? Let's begin looking at this by dividing group belief claims into claims about what I'll call aggregate groups and entity groups. An aggregate group belief claim is just a claim that some subset of the individuals in the group have a certain belief. It might be the majority of the individuals, as in statistical syllogism cases, or it might be just one prominent individual in the population having that belief. The English language is fairly tolerant of our using group belief phrases to cover cases where we mean merely that some individual group members have the belief. In what I'm calling an entity group, it might also be the case that any or all of the individuals making the group up have the belief ascribed to the group, but what's important is that, in addition, the whole group, as an entity, can in some sense be said to have that belief. The group as a whole somehow acts as a coherent, unified, individual agent, with beliefs and desires of its own. Sometimes when people make group belief ascriptions, the best way to make sense of what they are claiming is that they conceptualise the group as some sort of actor or agent. Thus when historian Robert Pells writes that with its Fullbright scholarships and America House performances, "Washington's principal target was not the young or the masses, but people in the chancelleries,

the media, the business and professional communities," he seems to be speaking about the aims of the American government as a whole, and not necessarily those of any particular people in it (1997, p. 87).

Can any sense be made of the notion of a group as an agent, or as the sort of thing that can have beliefs and desires? Let's start by seeing if perhaps these sorts of group belief claims can be dismissible outright as metaphysically muddled. To begin looking into this, we need to ask what it means for an entity to have beliefs and desires. There is, of course, a huge literature on what beliefs and desires are, with a large degree of disagreement on the fundamental conditions for being a belief or having beliefs. If one wanted to dismiss outright the claim that groups can be agents with beliefs, the easy way to do it is to select one of the theories with more stringent requirements for having beliefs (say Jerry Fodor's (1994)) and show that most or all groups must fail to have beliefs or desires, in that sense of belief. A more thorough and less contentious way to argue that groups couldn't have beliefs would be select one of the weaker, less stringent theories of what it means to have beliefs, and demonstrate that even using a weak construal with relatively few requirements, groups can't count as having beliefs. If one could show that groups fail to meet even our most lenient criteria for having beliefs, then this would go a long way toward showing that it's unlikely that any entity group belief claims are sensible.

Among current theories of belief ascription, which are the least stringent? Well, it's hard to get more lenient than Dan Dennett's theory. Dennett writes,

> To a first approximation the intentional strategy consists of treating the object whose behavior you want to predict as a rational agent with beliefs and desires and other mental states ... [A]ny object – or as I shall say any *system* whose behavior is well predicted by this strategy is in the fullest sense of the word a believer" (1987, p. 15).

For Dennett, for a system to be an intentional agent, it merely has to be one that can be seen as having goals it's trying to achieve (desires), some sorts of map of the outside world containing the constraints it will encounter trying to achieve these goals (beliefs), and some sort of tools for interacting with the external world in a way that attempts to achieve these desires, given these constraints. For Dennett, to be rational seems to be little more than for the beliefs, desires and actions to be consistent with each other. On having beliefs or representations Dennett writes,

> It's not what we attribute (or should attribute) beliefs and desires only to things in which we find internal representations, but rather that were we to discover some object for which the intentional strategy works, we endeavor to interpret some of its internal states or processes as internal representations. What makes some internal feature of a thing a representation could only be its role in regulating the behavior of an intentional system (1987, p. 32).

With this minimal criterion of what it means to be an intentional agent, what sort of entities count? In a number of places Dennett has discussed the beliefs

of clams, thermostats, mousetraps and coke machines. "Consider a soft-drink vending machine, designed and built in the United States and equipped with a standard device for accepting and rejecting U.S. quarters," says Dennett. "Let's call such a device a "two bitser." Normally, when the quarter is inserted into a two-bitser, it goes into a state, call it Q which "means" ... I perceive/accept a genuine U.S. quarter now. Such two bitsers are quite "clever" and "sophisticated." Dennett starts off with the terms "means," "clever," and "sophisticated" in scare quotes but then goes on to argue that the intentionality here is quite real.

On this very minimal view of belief, can a group as an entity, have belief? Certainly. All one has to do on this criterion is to set up a group of people to simulate the behavior of a coke machine. We could get a group of Kuwaities together and organize them in such a way so that one takes the quarter, one checks the shape, one checks the weight, one compares the picture of a George Washington with a prototype, one puts the quarter in the sack, one takes a coke from a cooler, etc. In this way, the group could be said, on Dennett's criteria, to go into the belief state "I'm accepting a genuine U.S. quarter now" even if *no* member of that group ever tokened such a thought.

Forming a collective agent

On Dennett's minimalist views of the requirements for intentional agenthood, it seems that, at bottom, what a system requires to have intentional states are some set of goals and some way of coordinating the goals with movements, with each other, and with various pieces of information about the environment. When an entity possesses these, almost everything it does can be seen as a rational way of trying to achieve those goals in a given world. There are various sorts of groups that would seem to be able to meet such requirements. Let's call groups to which we can clearly assign goals "programmed groups." A list of such programmed groups might include sports teams, corporations, or even nations. Among the main challenges a "programmed group" faces is coordinating the activities of its various members in ways that help ensure that its goals are met. There's a continuum of possible ways of achieving this coordination.

At one extreme end of this continuum, coordination can be accomplished by having a "dictator" who assigns group members requisite parts and who rewards and punishes their activities in a way that helps ensure certain goals are met. Let's call groups that work more or less like this "dictator groups." In a dictator group, while many of the group's beliefs and desires exist inside the head of the dictator, they could still be thought of as the group's since:

A. It's the group as a whole that acts to achieve the ends (no individual could achieve them).

B. Many of the group's sub-beliefs and desires, the "know-how" enabling the group to accomplish certain tasks, exist only as behavioral dispositions of coordinated *collections* of group members.
C. The members and the dictator himself could be replaced without our having a tendency to say that the beliefs and desires of the group changed.

In a dictator group, then, a powerful individual an take a varied set of believers and organize, coordinate, and transform them into a unitary agent that acts to achieve certain unitary ends.

Moving toward the middle of the continuum, one can imagine people coordinating their activities in such a way as to help achieve the goals of the group, not by having such activities directed and sanctioned by a dictator, but by each other. Let's call groups that coordinate activities in this way "collectively-controlled" groups. In collectively-controlled groups, it's agreement among the members themselves, rather than the policing of a dictator, that ensures that the requisite integration and cooperation take place. One important class of collectively-controlled groups are the sort of groups Margaret Gilbert describes in her book *On Social Facts*. A good example of the sort of group Gilbert has in mind is a pair of parents telling their teenager what time to be home from the dance. "Your mother and I believe you should be home by midnight," says the father. This is a worked-out compromise between father's belief that no curfew is necessary, and mother wanting her son home by 10:00. Here, the beliefs and desires talked abut aren't the beliefs and desires of *either* of the individuals – they're the negotiated beliefs and desires of the unit. The *actions* they'll take will be the coordinated rational actions of a unit trying to achieve certain agreed-upon goals (see Gilbert 1996). The sorts of groups she describes and the sorts of groups governed by other sorts of collectively-controlled forces are types of groups that can achieve the kind of integration and cooperation needed for the formation of a Dennettian rational agent, with beliefs and desires of its own.

In dictator groups and collectively-controlled groups, the beliefs and desires of some individual group members are coordinated together and integrated with actions to form the beliefs and desires of the group as a whole. The Kuwaiti coke machine case, however, shows that on Dennett's characterization of agenthood, none of the group members needs to have or to know anything about the beliefs and desires of the group as a whole. In that example, the group can be said to have the belief "I'm accepting a genuine U.S. quarter now" even though none of its members do. At the other extreme end of our continuum from the dictator group, then, are groups whose behavior has somehow come to be coordinated in a way requisite for the achievement of goals, even if none of the members have the slightest idea of what the beliefs and desires of the group as a whole are. Let's call groups that might be said to have certain beliefs and desires without any of its members necessarily having an awareness of them "cell-cluster groups." The "cell cluster" name is an allusion

to the fact that organisms can have certain goals, even though the various cells making up that organism have no idea what those goals are.[1]

The Kuwaiti coke machine case is an example of a very simple belief had by a cell-cluster group. But if we are willing to accept this minimal construal of what a belief is, then, in principle, there's no limit on how complicated the beliefs of the whole can be. How could "cell-cluster groups" have *complex* interesting beliefs that none of its members do? Bill Lycan (1988) has argued that a good way to conceptualize modern cognitive science is to assume something called "homuncular functionalism." Suppose you want to understand how a complex human behavior is enacted or how an expert system AI program works. A useful way to do this, according to Lycan, is to conceptually divide what's going on internally into a large set of subtasks, propose that there's a little man or a little machine in there in charge of doing each of those subtasks, and assume that an army of such little homunculi makes sure the task as a whole gets done. You end up understanding this complicated task by seeing how it's the net effect of lots of easier-to-understand tasks. Now, of course, you've got the same problem for understanding how each of the little men inside know how to do what they do. But you can explain the behavior of each of these little men in the same way. You propose there's an army of even smaller men doing each of the many subtasks in the homunculi's head, and the same for each of those homunculi, on down, until the entire task is understood in terms of a hierarchically organized set of units doing extremely simple tasks. At the top level description is a human trying to decide which route to take home. At the mid level are homunculi assessing various routes, and at bottom are neuron-like creatures pushing on and off buttons.

One can easily see how you could appropriate this "army of little men" model for ascribing beliefs and desires to groups. All you do is take the figurative homunculi literally. When you do this, you've explained how any entity that is minimally an agent can be simulated by a suitably organized army of people. If you've got a full blown human intentional agent described in this homuncular way – you could take this as a literal blueprint for creating a collective army that's a full blown intentional agent.

In the philosophical literature, there have been several famous thought experiments in which something like the cell-cluster groups instantiating homuncular functionalist "programs" have been described (usually to make different points from the one I'm making her). The most well known is Ned Block's story of convincing the government of China to use its people to simulate a human mind for an hour. Block first described how a set of homunculi could implement a Turing machine running an artificial intelligence program that simulates your mind:

> "Suppose the light representing input I17 goes on. One of the G-men [homunculi] has the following as his sole task: when the card reads 'G' and the I17 light goes on he presses output

> button 0191 and changes the state care to 'M'. This G-man is called upon to exercise his task only rarely. In spite of the low level of intelligence required of each little man, the system as a whole manages to simulate you because the functional-organization they have been trained to realize is yours."

Then Block asks us to imagine the people of China doing the same tasks and implementing the same program:

> We provide each of the billion people in China (I chose China because it has a billion inhabitants) with a specially designed two-way that connects them in the appropriate way to other persons and to the artificial body mentioned in the previous example. We replace each of the little men with a citizen of China plus his or her radio. Instead of a bulletin board, we arrange to have letters displayed on a series of satellites placed so that they can be seen from anywhere in China.
>
> The system of a billion people communicating with one another plus satellites plays the role of an external "brain" connected to the artificial body by radio (Block, 1993, p. 239).

Thinking about agents in this way, we see how it would be conceptually possible for a *country* to, say, play chess with another country and for the country to count as having the belief that it ought to get its queen out early, just as a chess-playing computer could be said to have this belief in some minimal sense. Indeed, given resources and time, a suitably organized country could do anything a digital computer could do.

The interesting thing about cell-cluster groups instantiating a program in a homuncular functionalist way is that such a group not only satisfies the minimal rational agent construal of intentionality accepted by Dennett. It conforms to the much more stringent *functionalist* view of what beliefs are, accepted by the *majority* of cognitive scientists. On this view, an entity doesn't merely have to be rational, it has to have a certain style of functional organization to have the belief that p. But there's no in-principle reason why a group couldn't instantiate that functional organization. If one agrees with the functionalist view of belief, then one ought to agree that, in principle, groups can have even quite complicated mental states.[2]

What's the moral to be drawn here? The moral is that we really can't, in a quick and painless way, dismiss the group-as-agent view as nonsense. Group belief seems to be a serious coherent metaphysical possibility. If we want to say groups can't have beliefs, we need to argue that a comparatively more stringent theory of what beliefs are is the correct one.[3] Indeed, to deny the metaphysical possibility of such group beliefs, we'd have to argue for a much stronger view of belief than most cognitive scientists are willing to accept. We might want to do this, but unless one does this, we can't just dismiss group belief ideas as silliness.

Before moving on to talk about interpretivism, let me briefly end this section with a note of caution. I've argued that entity group beliefs are a coherent metaphysical possibility. But I can see numerous potential *epistemological* problems concerning what one can say about the thinking and behavior of groups like

these. When dealing with a person, you know beforehand that you've got something like a rational agent. But you don't know beforehand whether a given group is really a true rational programmed group. Without knowing this you aren't going to be able to say much about it. Furthermore, with individual humans, even if we know very little about them, we still tend to know something about systematic ways in which perfect rationality can't be counted on. We also know something about the additional functional organization of their minds, aside from rationality. We even usually know something about other default beliefs and desires it's safe to attribute to them in certain circumstances. For programmed groups we know none of these other things about them without a great deal of additional study about exactly how these groups function. The question, then, is whether, given we have to do detailed studies about these groups anyway, rather than just using folk psychology to understand them, we really gain anything by popping up a level and discussing the group as an agent? These are issues I'll have to leave for another occasion.

INTERPRETIVISM

If some of us tend to roll our eyes at FIC descriptions, then, it shouldn't be because we know we can easily dismiss collective belief as a metaphysical possibility. Let's now look at the other feature that people find problematic about FIC descriptions – their interpretive nature. What is that many people find objectionable about this?

By interpretive social science, I mean social science with a family resemblance to psychoanalysis. In interpretive social science, actors (individuals or groups) are said to think and do what they do because of deeper beliefs, desires and other states that are usually hidden. The task of the interpretive social scientist is to uncover those hidden beliefs. Anthropologist Emiko Ohnuki-Tierney, for example, tries to get us to understand Japanese monkey performances by showing us how, at their end, the audience "realizes that it was they who were the untamed nature to be culturalized by the monkey" (1984, p. 304). Similarly, Andrew Feldman seeks to explicate the famous Rodney King police brutality case by showing us how, for Sergeant Stacey Koon,

> The successful confinement of King – the symmetry of a body lying at attention with the face in the dirt – and the acquisition of linguistic reciprocity marked the neutering of the animalized body and its internalization of the will of the state. A "gorilla in the mist" a black "bear" that was insistent on rising on its haunches was turned by violence into a speaking subject (1994, p. 410).

If there's a single central objection to interpretive social science, it's the suspicion that interpreters are ascribing interesting beliefs to people and groups

without sufficient evidence that those beliefs are really there. In what follows I'll examine, as before, whether such ascriptions tend to deserve the eyeball-rolling they are sometimes met with by tough-minded scholars, or whether there is something sensible about them. At the outset, let me note that interpretivists are often unclear about just who they are ascribing the interpretation *to*. Interpretive group ascriptions (e.g., "through the E.U., Germany again wishes to control Europe") certainly have all the ambiguities that ordinary group ascriptions do. Are such beliefs being ascribed to certain people *in* Germany or to Germany as a whole? For ease of exposition, I'll initially assume that he beliefs in question are being ascribed to individuals within such groups. Later, I'll discuss how what I've said applies to ascribing beliefs to groups as collective entities.

Interpretation with minimal theories of mind

Informally, the beliefs, desires, and other mental states we can put in the vague category of "hidden" states that interest interpretive social scientists, are simply those that tend *not* to be ascribable using whatever methods we typically use to ascribe everyday beliefs. Most of us endowed with functioning perceptual equipment and a basic knowledge of our cultural practices could easily infer that during the beating, Stacey Koon believed Rodney King was trying to get to his feet. But it's clear we have to make some sort of special effort in order to learn that King was seen by Koon as a gorilla/bear that needed to be neutered and made to signal a return to the internalization of the will of the state by speaking. What methods do interpretivists use to uncover these sorts of hidden beliefs?

Interpretivists themselves tend to be very unclear about how they know they've really located the hidden beliefs of an agent. Geertz himself once wrote, "You either grasp an interpretation or you do not, see the point of it or you do not, accept it or you do not." And "This raises some serious problems of verification, all right – or if 'verification' is too strong a word for so soft a science ... of how you can tell a better account from a worse one. But that is precisely the virtue of it" (1973, p. 16). Despite this (rather appalling) glibness, I believe that there actually are systematic methods interpretivists tend to use. To begin with, there are only a limited number of methods that they *could* use. The task of describing any unobservable states of minds in others is just one instance of the very common and general problem of trying to uncover information about entities we can't directly observe. There have long been several methods for justifying claims about the unobservable. One of the main ways is to try to infer what the unseen entity must be like by deriving information about it from our current theories of what the world is like under certain conditions, and knowledge of those conditions. There are two general methods of gathering evidence and

using theories to make this sort of inference. One method might be termed "the environmental strategy." Here one starts by observing external conditions that are thought to cause certain unseen states of affairs to result, according to certain general regular patterns or laws. In the case of belief ascription, the environmental strategy begins with the idea that certain beliefs *result from* exposure to certain perceptual/environmental situations. Showing that a person was exposed to a certain natural or social environment is taken as evidence of her having certain resulting beliefs. The other method, which might be termed "the behavioral strategy," starts with the assumption that only certain sorts of things can *cause* certain resulting actions, according to our theories. So when those resulting behaviors are observed, that's taken to be good evidence that those purported hidden causes are in fact there. In the cases we are discussing, observing certain resulting behaviors (including verbal utterances) is taken to be evidence that certain internal beliefs must be there, causing such behaviors to occur.

Whether they ever explicitly discuss it or not, interpreters, like most scholars examining unseen entities, are almost certainly using a combination of environmental and behavioral strategies to reach the conclusions about beliefs that they do. What's less clear is which of the needed *theories* of *inner* mental processing they tend to use in combination with these. Anthropologist Brad Shore once wrote of his participation in a cognitive science conference, 'My assignment at the conference was to try to characterize the implicit theory of mind that anthropologists employed in their cultural analyses. In such heady company, it soon became clear to me that most of our work in symbolic anthropology proceeded innocent of any well-formed theory of mind whatsoever' (1996, p. vii). This lack of explicit theory is unsurprising, given most interpretivists' romantic resistance to utilizing any well-researched scientific theories of the mind or brain. Still, almost any use of the behavioral or environmental strategies requires that one use *some* theory of how the mind works.[4] My own experience with interpretivists and their works suggest that many make use of what one might call "minimalist" theories of mind. Whatever theories of mind they actually use, let me say what's problematic about trying to infer hidden mental states using certain minimalist theories of mind.

One technique for trying to ascribe beliefs to people using only a very minimal theory of mind, is to 1) focus on behavioral output, and 2) assume only that the mind is a sort of "rationality machine", which coordinates beliefs and desires with each other to enable to agent to reach her goals. This is the sort of theory and method Dennettians advocate for ascribing beliefs to people. Using what's commonly called "inference to the best explanation," or abductive reasoning, an interpretivist using this technique might collect a large sample of behaviors, and then postulate that beliefs-desire set S is the one that a rational mind is most likely to have utilized in generating the observed behaviors. We can see an example of someone attempting to make this sort of abductive inference in social critic Warren Farrell's book *The Myth of Male Power*. In this work, Farrell (1993) takes a look at a large range of observations such as these:

1. the military's not giving combat assignments to women
2. 24 of the 25 professions rated as most hazardous are virtually all male
3. the more hazardous the job, the higher percentage of men it has
4. men are twice as likely to be the victims of violent crimes and three times more likely to be murder victims
5. the suicide rate among men in their early 20s is six times higher than that of women the same age, and the suicide rate of men over 85 is 1350 times higher
6. breast cancer receives 600 percent more funding than prostate cancer even though death rates from each are equal.

In the course of examining hundreds of observations like these, Farrell, rightly or wrongly, concludes that the beliefs underlying such a pattern must be that women are actually perceived as the valuable gender (especially in evolutionary terms) that needs to be protected and preserved at all cost, while men (a dime a dozen in evolutionary terms) are thought of as essentially disposable. Farrell clearly comes to this conclusion by trying to infer what sort of rational thinking could have generated these observable situations.

As many philosophers of science have documented, abductive inference is one of the central strategies used throughout successful sciences. While interpretivist social scientists may not have direct evidence for the belief states they postulate, they seem to be trying to use abductive inference to inform us about the hidden structure of the mind in the same way Perrin used abductive arguments to inform us about the hidden structure of the atom. In interpretive social science, as in other areas of inquiry, one proceeds by collecting lots of evidence, proposing that a certain unobserved causal structure best fits the observations, and inferring that such a structure must be present. Observing further behavior that would be predicted by such an inner structure, is thought to help confirm that one has hit upon the correct view of the inner structure.

But this strategy tends to be problematic when used to try to uncover *beliefs* – even garden variety ones. It is a point of elementary logic that merely showing that one can confirm a prediction entailed by a hypothesis isn't enough to show that that hypothesis is true. If there are plenty of viable alternative hypotheses that could generate the observed prediction, then observing that prediction doesn't give you *any* evidence that the hypothesis in question, rather than its equally well-predicting rivals, is true (see Laudan (1996) for a good articulation of this point). If different beliefs and desires could have led to the same behavior, then observing that behavior provides no evidence for the existence of any particular beliefs or desires. One of the root difficulties of belief ascription is that, unlike the sparse fundamental building blocks of some other sciences, there exists not merely a few dozen or even a few thousand different possible beliefs and desires – but an infinite number of them. We must begin, then, by selecting from an unlimited number of potential belief posits. The beliefs we can reasonably ascribe using a behavioral strategy are, of course, only those that could possibly

cause the behavior we observe. This, however, is a fairly weak restriction. We can think of beliefs as something like maps used for getting around the world. A central problem is that many different sorts of maps could usefully lead us to the same destination. Any given behavior is, thus, consistent with positing numerous different core beliefs and desires. When Malinowski's (1922) Trobrianders initially pointed to an outrigger canoe, for example, and said "Kewo'u," He initially had no firm way of telling whether they were thinking "there's a boat," "there's a group of undetached boat parts," or "there's a stage in a boat's existence."[5]

And there are still further difficulties. Beliefs do not cause behavior by themselves, but do so in conjunction with desires and, often, with other beliefs. A selection of vanilla over chocolate may be based on the belief that vanilla is tastier and a desire for the tastiest ice cream. It may also, however, stem from the belief that chocolate is tastier but also more fattening, and a desire to lose weight. Selecting vanilla could also stem from a superstitious belief that chocolate should never be eaten on Wednesdays, and a desire not to offend the gods. The vanilla-choosing behavior alone won't tell you which of these beliefs and desires are behind it. (This realization marked the downfall of the philosophical behaviorist view that a belief statement was merely a statement about a disposition to behave (see Stich, 1983; Churchland, 1988).) If one's task is to find a belief that, along with a string of auxiliary beliefs and desires, would lead to the production of a given behavior – with no prior restrictions on the number and type of such strings – then the task is analogous to one of guessing a number which, added to some string of positive or negative numbers, yields the sum of 5. If one makes the appropriate adjustment in the strings of added numbers, literally *every* number can qualify. Similarly, with the right adjustments in auxiliary beliefs and desires, it is logically possible for *any* belief to cause *any* behavior. Increasing the numbers of behaviors one observes can help rule out some possible belief-desire sets by showing that some of the predictions of that set are incompatible with the further observed behaviors. But even large sets of behaviors can be shown to be compatible with any given belief-desire set, so long as one is willing to postulate the existence of enough additional (perhaps very odd) beliefs and desires that make all those behaviors rational. Behavioral observations and assumptions of rationality, alone, then, could not justifiably enable one to say what sort of hidden beliefs lay behind them. What one needs are some further constraining theories that say which of the belief-desire sets that *could* generate the behavior we observe are likely to be the ones that *do*.

A different minimalist strategy that interpretivists could use is to approach the problem from the other direction. One could look at the sorts of environmental inputs people in a particular culture tend to have, and use theories of mental state formation processes to infer what sorts of mental states they are led to have by these environmental inputs. A family of minimalist theories of mind that has been around at least since Hume's time is that the contents of mind,

at any given time, arise due to certain features of environment or of another internal state "calling to mind" other features *associated* with these. For "ordinary" beliefs (see footnote 6), the clusters of features brought to mind at a given time are the ones that are most commonly seen to be clustered together. The sight of a bat, thus, easily brings to mind thought of flight and shrieks. It might also, however, bring up thoughts of caves, Batman, or Dracula. I suggest that many of the mental states that seem especially "hidden" are thought to be so because they are the results of more peripheral associations that are far too numerous for us to make confident guesses about, merely by seeing what our compatriots see. Besides being numerous, each person's associations are highly idiosyncratic, due to their varied personal histories. If we don't know the details of these life histories (and even if we did know them), there are innumerable associations one might be making, looking at any given thing. The mental states people find themselves in through following chains of peripheral association mechanisms are therefore "hidden" from third parties in a way that ordinary world-mapping feature-cluster beliefs are not. Such thoughts might also be hidden from the conscious view of the cognizer herself. The question, then, is how could interpretivists uncover what these hidden mental states are. Observation of what's constantly saliently associated with what might tell us, for example, that the Swat of Pakistan believe that leather workers won't discipline their patron's children. But how could we infer, as Lindholm (1981) does, that these leather workers are seen as symbolic male mothers?

One thing interpretivists often do is look extensively at the cultural environments of the people studied, to see what kind of concrete or abstract features tend to be paired together in that culture. In the analysis of the Japanese monkey performance mentioned above, Ohnuki-Tierney comes to the conclusion that,

> At the end of the performance, [the audience] realizes that it was they who were the untamed nature to be culturalized by the monkey. Put another way, the monkey and the outcast are the small eyes in yin and yang. For this reason, I think, even amidst the laughter at the monkey performance the audience is reminded, albeit vaguely, of their darker side, as represented by the monkey and the outcast trainer (Ohnuki-Tierney, 1984, pp. 301–304).

She tries to convince us that this is so by describing in great detail the various ways in which monkeys are depicted in Japanese history and folk tales. Such exposure is supposed to show us the sorts of symbolic associations Japanese people might have with monkeys that we would not initially be able to see.

While it is certainly helpful to be made aware of possible associations that we wouldn't have thought of by learning about other people's history and culture, the amount we can rely on such a strategy for accurate hidden belief ascription is acutely limited. After all, even if we knew *all* of a person's environmental inputs, this, by itself, would tell us nothing about where that person's train of thought tended to go at any given time. Any seen feature has millions of other

features possibly associated with it (more on this below). To know which ones are more likely than others requires additional knowledge of the internal mental mechanisms by which certain items are frequently called to mind by other items.

Along with looking at common cultural pairings, interpretivists could also begin looking at what other features might be mentally associated with other features (and more abstract categories) by looking at which items are connected by resemblance and less salient contiguity. When confronted with a candidate for something that has a deeper symbolic meaning, interpretivist social scientists of all persuasions, much like psychoanalysts and literary critics, suggest that what's going on is that the observed surface features are mentally associated with some other features, in native minds, because of a vague resemblance or contiguity. Thus Wilson interprets the meaning of eating a banana in a ritual performed by the bride in Nyakyusa culture as a symbol of the sex she will have with her husband (1959). Shells and coral have been said to be symbolically associated with the ocean by Levi-Strauss because of spatiotemporal contiguity (1963). Interpretivists also posit that something can symbolize something else, not merely by being mentally associated with it in some way, but just by being associated with something else that is. Hence, something can be symbolically associated with something else through elaborate chains of association. Such a convoluted chain can be seen in Sapir's discussion of the symbolic association of lepers and hyenas among the Kujamaat Diola. Lepers, on Sapir's account, are thought to be burned by a magic fire associated with iron working forges. Leprosy is associated with the forge because the way that leprosy acts on the body is seen to be isomorphic to the way that forge fire works on iron. The forge is thought to send leprosy when someone attacks something that the forge, a source of spiritual power, is thought to protect (primarily cattle or children). If a cow is killed through witchcraft, it is thought likely to have been done by the person in the form of a were-hyena. "Hence," writes Sapir "if you had leprosy, you were caught stealing something protected by the forge; and if you were stealing, you might have been stealing in the guise of a hyena" (1983, p. 533). Thus the symbolic connection between lepers and hyenas.

The basic problem with using association to find symbolic meaning is that, as Anderson and Bower (1973) demonstrated long ago (and players of the "Kevin Bacon game" have discovered recently), virtually anything can be associated with anything else in the right circumstances. Bananas can serve as phallic symbols, but they could also possibly serve as symbols of the tropics, of monkeys, of banana bicycle seats, of the Velvet Underground, or of Bob Dole. Showing an interesting *possible* set of associated remindings by itself does nothing to establish that such a set of remindings is *actually* present among the people studied. It's possible, as Feldman suggests, that Officer Koon saw Rodney King as a wild bear that needed to be culturalized by submitting to state authority. So is the idea that King was seen as a symbol of a black revolutionary movement, one that threatened the American government and way of life. But perhaps

Koon saw King as a symbolic snake, and believed that it is proper for snakes to be lying on the ground. Maybe, in trying to put King down, Koon remembered a tree that he chopped and chopped at but couldn't fell as a child, and King became the symbol of Koon's continually straining and failing. With only resemblance and contiguity as constraints, there is simply no telling what mental states someone has found himself in at a given time.

Problems with less-minimal theories of mind

Environmental and behavioral strategies, combined with only minimalist theories of mind, then, seem inherently capable of coming up with little more than hidden mental states a person *possibly could be* in. The remedy, however, seems clear. Interpretive belief ascribers need to combine these strategies with more substantive, less minimal theories of mind. To make plausible the idea that it is one set rather than another of the potentially infinite symbolic associations that is actually invoked in certain circumstances, for example, one needs to make use of constraining theories about what symbols tend to be invoked when.[6] A Freudian theory, for example, will put some constraints on what a symbol is a symbol of in light of its "sex drive/hydraulic" posits about which types of thought are the ones most commonly brought to mind. There are lots of problems, however, that keep the constraining theories that interpretivists *currently* tend to rely on from being much help. First, the theories of our symbolic mental life that interpretivists tend to rely on (Freudian , Jungian, Levi-Straussian) are extremely dubious at best.[7] More importantly, however, even if one of these theories were well established, they don't tend to provide enough constraints to keep dozens and dozens of different thoughts and associations as counting as potential states of mind at a given time, even within the constraints of the theory. *All* of the different possible interpretations of King's attempt to rise to his feet that I suggested for Koon above, for example, are compatible with *all* of the theories of the unconscious I just mentioned. If we want additional constraints on possible ascriptions that less minimal theories of mind might give, we clearly need to look toward different theories of mind than the ones currently most favored by interpretivists.

A telling example

If interpretivists are using the minimalist or inadequate theories of mind they seem to be, then they cannot give us reasons to believe that the beliefs and desires that they suggest as underlying the behaviors we observe are any more likely to be there than a huge number of competing ones. To really vividly see this, all one has to do is sit down one afternoon with a copy of any interpretive

analysis and some knowledge of the observations the analysis tries to account for, and see how many different belief and desire ascription sets you can come up with for the same observations. I did this one day with Geertz's essay "Deep Play, Notes on the Balinese Cockfight" (see XXX, 1998). I chose that one because this essay is widely regarded as a model case of interpretive analysis by a gifted ascriber.

In "Deep Play," Geertz tries to make sense of the strange practice in Bali whereby the normally calm regal "hyperpolite" Balinese get together and scream and shout and spend huge sums of money betting on two roosters tearing each other to shreds with tied-on spurs. What's going on in people's minds here? Well, first, Geertz looks at various features involved in the cockfight and uses the basic crude resemblance and contiguity association method to suggest how all this is seen by the Balinese. Fighting cocks, he suggests, are read to be phallic symbols because of their vague resemblances to penises, their tendencies to be held by men between their thighs and stroked, their tendencies to be cared for and fussed over exclusively by men, etc. Balinese national identity is shown to be symbolized by the cock by discussing a famous legend in which the island of Bali was separated from Java by a Javanese prince trying to escape from a heroic cockfighter. The cockfight is shown to symbolize social solidarity by discussing how groups of kinsmen pool their money to make large bets in favor of an underdog cock, etc. Geertz then posits a coherent set of beliefs, desires and symbolic associations that accounts for this large range of elements of the cockfight and Balinese behavior in general. When all is said and done, the central picture that seems to emerge for him is this: The cockfight is a safe arena where the Balinese can view the raw naked essence of who they really are. In most parts of their overtly observable lives, the Balinese are hyper-polite, and deferential. Hierarchical status is never something that is overtly sought after. Elegance and grace are highly prized and anything that hints of animality is found repulsive. Children's teeth are even filed at puberty so they won't look like animal fangs. But deep down the Balinese know that conflict, chaos and instability are everywhere. This is all transparent in the cockfight where one's carefully nurtured cock (along with a considerable sum of money) can be suddenly lost in a violent attack of kicking and pecking. People are also not really as deferential and impervious as they seem. The grace and elegance are a facade. Deep down the Balinese are as filled with aggression and anger as people anywhere, perhaps more so, as there is less chance for overt outlet. In the fury of the cockfight, such anger and aggression can be given full expression. "In identifying with his cock," writes Geertz, "the Balinese man is identifying not just with his ideal self, or even his penis, but also, and at the same time with what he most fears, hates, and ambivalence being what it is, is fascinated by – 'The Powers of Darkness'" (1973, p. 420). Most importantly, in the transparent world of the cockfight, the Balinese make clear that status hierarchies, among men and among kin groups, are truly serious business, something that

blood will be shed for, something that people are willing to put substantial amounts of money on the line for. Ironically, it's only in a "game" that the Balinese really allow themselves to say that status, my group's superiority (something that ordinarily can not be overtly striven for), is for real.

However insightful Geertz's analysis seems, it remains the case that there are numerous other possible interpretations of the Balinese cockfight. On one occasion, I sat down and wrote four rival interpretations of the cockfight. Each one appears to be consistent with all of the observations and factual information Geertz relates, and coherently fits together various associated meanings in the standard ways used by interpretivists' and the analysts and literary critics they model themselves on to come up with "readings." Let me present two of them:

1. The Balinese cockfight is a political protest. As Geertz notes, cockfights have been outlawed by the political authorities centered in the neighboring island of Java. Javanese Indonesian authorities worry about what tourists or various heads of state will think about their country if they know that such barbaric violent pastimes are taking place. The central meaning of each cockfight staging for the Balinese can thus be seen as a political protest against Javanese authority. Nearly all the policemen (at the time of Geertz's work) in Bali were Javanese. So it is Javanese authorities who will be directly provoked and angered by this display. Moreover, it has always been the cockfight that symbolized Balinese autonomy. As mentioned before, Bali was mythologically separated so a Javanese prince would be spared the wrath of a cockfighting hero. "Even the island itself is perceived from its shape as a small, proud cock," writes Geertz, "poised, neck extended, back taut, tail raised, in eternal challenge to the large feckless shapeless Java" (1973, p. 418).

It is significant that the myth mentioned above is a Hindu myth. Bali remains a Hindu-Buddhist outpost in one of the world's most populous Moslem countries. Where Javanese authorities want Indonesia to be an austere artless Islamic country, the pageantry of dressed-up cocks fighting serves to remind Balinese people of their proud shadow puppet tradition where battles between Hindu gods and monsters are depicted. There are also myths where fighting cocks turn into rescuing garuda birds, central to Hindu mythology. The cockfight itself is seen as a reminder of this rich Hindu heritage that Javanese authorities want to sweep out of Indonesian life. The main story told by the cockfight is one of Balinese joining arms and flying a forbidden flag – sending Java the message that they prefer their colorful barbarian pleasures to bleak visions of modern statehood.

2. The Balinese cockfight is thinly disguised homoeroticism. Cocks are clearly phallic symbols to the Balinese. This is evident from the first time that one hears sabung (cock) as the standard slang for penis. The degree to which cocks and cockfights are symbolic of a very deep male bonding can be seen everywhere. Wherever you see a group of men gathered, sitting in a circle, many of them will be holding a cock between his thighs, stroking it. This is the symbolic equivalent of the "circle jerks" once common among American adolescents.

What's more, according to Geertz, "Now and then, to get the feel for another bird, a man will fiddle this way with someone else's cock for a while, but usually by moving around to squat in place behind it ..." (1973, p. 419). Such symbolic homoerotic behavior surrounds the cockfight everywhere. Geertz reports that red peppers are often stuffed up the cocks' anuses "to give them spirit." When a cock is losing a match, his owner "blows in its mouth, putting the whole chicken head into his own mouth and sucking and blowing, fluffs it ..." (p. 423). Cruising-like behavior is displayed when, in between every match, men flutter into the center ring, looking for partners to bet with (in a process even Geertz refers to as "wooing"). Women are carefully excluded from any of this, as cockfighting is an all male affair. Indeed, women are not even allowed to look at the spikes that men affix to the legs of their cocks before battle. During the cockfight, Balinese can retreat from their hyperpolite feminine world and become symbolic Greco-Roman warriors, passionately interested in battle and the eroticized camaraderie of their fellow men.

Given the outlandish interpretations commonly seen in anthropological journals today, these interpretations are fairly tame. And there is nothing stopping one from looking at the sort of chaining associations that interpretivists often do to uncover numerous other meanings in the cockfight. The cockfight, for example, could be thought of as a depiction of the anxieties one feels in trying to bring up children. The animal cocks can be thought of as calling to mind still unsocialized animal-like children. As the winning cock's owner is paid his gambling winnings in silver rupia, the silver reminds him of shiny dagger blades or the shiny mysterious ocean over which many unknown countries lie, harboring hidden unknown dangers for one's children. Having worked through all these dangers symbolically in numerous cockfights, however, Balinese men are able to face domestic life, hardened to its burdens, and able to face its challenges, etc ...

The problem with making hidden belief ascriptions in the way that interpretivists do, is that they've really got nothing preventing one from coming up with numerous sets of ascriptions that account for everything observed as well as rival ascriptions do. If all I have to do is come up with some set of hidden beliefs and desires that would account for a variety of observed Balinese behavior, constrained only by a rule which says to describe things as having a deeper meaning based on other things they are associated with, then I can come up with as many interpretations as someone will pay me to make. Neither the behavioral evidence nor the mentalistic theories interpretivists generally use, rule enough of them out.

Better ascriptions through better theories

In section 3.2, I mentioned that there was an obvious remedy for the tremendous under-determination problems stemming in trying to ascribe beliefs using the environmental and behavioral strategies: One must combine these strategies

with more constraining, far less minimal models of mind than many scholars use. Above I mentioned that constraining theories like Freudian or Levi-Straussian are still far too liberal. But there's no reason to think that, ultimately, better theories of mind couldn't give us what we are looking for.

And what might such more constraining theories look like? Two possibilities immediately suggest themselves. For any "black box" entity that one can't observe directly, one way to come up with a model of its workings, is to start with a theory of the component parts of which it is constructed, and use theories of how such parts link together in space to create more complex constructions (as in atomic theory or chemistry). Alternatively, for any black box mechanism at a given time, one might start with theories about sets of earlier initial conditions, then use dynamic law theories to talk about how forces from within and without will change this set to produce a resulting construction (as in predicting how a satellite will stabilize into an orbit).[8] If one is lucky enough to have well supported theories of components, combination rules, initial conditions, and modification rules, one has important additional ways to build models of black box mechanisms, besides relying on observations of inputs and outputs. If the black box we are talking about here is the mind or brain, we have at least the skeletons of models constructed along these lines based on decades of research by teams of researchers from neurology, psychology, and computer science. However incomplete these models are, they are far more advanced than the non-experiment-based speculations about the mind made by late nineteenth and mid twentieth century sages like Freud or Levi-Strauss. The "composite-construction" strategy is pursued at a very low level by neurologists and connectionist modelers looking at how neurons and neuron clusters are linked together to create specific machines that tend to bring certain sorts of things to mind in certain situations (see, for example, Squire, 1987; Morris, 1994; McClelland and Rumelhart, 1986; Clark, 1993). At a higher level, numerous cognitive modelers have investigated how various items of knowledge fit together to form mental models and theories (for example, Bower and Glass, 1976; Palmer, 1977; Johnson-Laird, 1993). Versions of the initial-conditions-and-changes strategy can be looked at over both high and low levels, and over large and small time scales. At smaller time scales various researchers have studied how initial mental states are changed by new perceptual information (Marr and Nishihara, 1978; Beiderman, 1987), operant conditioning (Reynolds, 1968) and learning to relax or tighten conditions under which certain behavioral or informational schemas are invoked (Holyoake and Koh, 1987). Over larger time scales, biologists and evolutionary psychologists have been studying how the process of natural selection has shaped and developed probable earlier mental structures (e.g. Barkow, Cosmides and Tooby, 1993). Information from these realms promises to tell us much about which sorts of beliefs and desires people are predisposed to have innately, and which they are likely to have at a given place and time. If we want to understand what Geertz and others call "the

native point of view," I believe it is by borrowing methods from cognitive science, not psychoanalysis or literary theory that we're most likely to succeed. In general, if we want to know which sorts of associated items are brought to mind in certain situations, a model like John Anderson's ACT* system, supported by decades of clinical research and computer modeling should surely be preferred to a Freudian's armchair speculations about which associations people are making when. (For example, was Freud really in a position to know that a man's reciting a line from a poem mentioning flowing water was an indication that he was unconsciously worried about his lover's missed period? (see Grunbaum, 1984).) Interpretive belief ascribers need to cure themselves of what look like allergies to all research they see as "scientistic" if they don't want to be overwhelmed by a plethora of possible plausible ascriptions.

Interpretive ascriptions for groups

We have been discussing the problems and potential solutions surrounding claims about things like "China's belief in its own inferiority" when such claims are construed as being about the mental states of individuals. What can be said about such claims read as ascribing hidden beliefs to groups as entities? Actually, many of the same things can be said. If groups can have beliefs and desires, as they apparently can on minimalist theories of intentionality, there is no reason they can't have bizarre beliefs with a vague resemblance to the sort of mental states that psychoanalysts are interested in uncovering in individuals. But we are likely to have many of the same problems in uncovering those beliefs that we've discussed for individuals, and a few more, as well. A group may function like a giant computer that produces certain behavioral outputs. But then knowing about the output behaviors of groups won't, by itself, tell you any more about the internal states that produced them, than knowing a computer's output would. Similarly, knowing a group's environmental "inputs" still leaves you with a huge number of potential world-maps and associations that could be formed on the basis of these. In addition, with groups, we can't even rely on the minimalist theories of how minds work that can help us with ascriptions to individuals. With human beings, we automatically know not only that we are dealing with rational agents, but also some of the ways in which agents typically do and don't fall short of ideally rational behavior. For a group, we have no idea of this without a good deal of further study. For fellow humans, we also know, from personal experience, that they tend to be "associative engines." And we have at least some ideas about what sorts of things are going to be associated with what through observed resemblance and contiguity. For any given group we want to ascribe beliefs to, on the other hand, we don't automatically even know *if* it works by being any kind of "associative engine" computer. And we certainly don't know which sorts of things are going to be associated with what (e.g.,

through resemblance) for such a machine that likely works very differently from the way our own minds work.

But the abstract logic of how we *could* successfully ascribe beliefs to groups should be the same as it is for individuals. If particular groups really have beliefs, we should be able to uncover them using a knowledge of inputs, outputs, and theories of how internal states are organized. And we should be able to develop models of the internal states of groups using the same abstract composite-construction and initial-conditions-plus-changes types of theories used for developing the constraining internal-state models for individuals. Where models of individual beliefs should be constrained by theories of the sorts of maps of the world our neuron clusters create, models of group beliefs should be similarly constrained by theories of how the individuals and subgroups that make up the groups cluster together, communicate, and integrate the information they have in order to produce behavior (see for example, Burt, 1982; Hutchins, 1995; Satz and Ferejohn, 1994; and Denzau and North, 1995). Where models of individual beliefs should be constrained by theories of which innate beliefs and desires are likely to be there after a long shaping by evolutionary forces, we could also constrain theories of group beliefs and desires by theories of how various group goal-seeking and world-mapping strategies would likely have been shaped by the forces of natural selection. Nelson and Winter (1982) have been developing such principles in order to explain how corporations evolve over time. At a more abstract level, Axelrod and Hamilton (1981) and numerous others have been developing theories about which sorts of individual behaviors will tend to become stably fixed in groups over time, due to the dynamic unfolding, via the constraints of the other actors, of certain overall net effects for the group. There's no reason to think that these sorts of theories couldn't serve just as well at helping us choose between various behavior-generating models of group inner states as their analogues could in helping us find the best models of individual internal mental states.

Closing a Dennettian loophole

I've just described some of the means by which one might successfully ascribe beliefs to groups, by doing more than just looking at the group's output behavior. But the reader who remembers my account of how groups could have beliefs may well wonder whether I am here proposing a solution for which there is no known problem. The *apparent* problem that the above sections aim to solve stems from the gigantic number of hidden beliefs and desires that could underlie behavior, given the information interpretivists use. Given this number, an interpretivist ascribing a particular belief-desire set to an individual or groups tends to be rather arbitrarily selecting a set as being the one that *does* underlie behavior, from numerous ones that *could*. Our previous discussion of theories of intentionality, however, point to what might be seen as a potential loophole

through which belief ascriptions might be able to escape the above "too many potential ascriptions" charge.

On standard "realist" assumptions, when a scientist is using abductive reasoning, she is looking for an actual hidden state responsible for generating observed behavior. On most theories of intentionality, including homuncular functionalist ones, we are trying to uncover the actual beliefs and desires that make people act as they do. Finding states that *theoretically could* generate certain behaviors is not enough. If we assume a minimalist Dennettian theory of intentionality, however, we are not trying to uncover an actual real belief state for individuals or groups. Dennett's theory of intentionality is an instrumentalist one. Dennett's theory, sometimes referred to as "holistic behaviorism," maintains that a belief is a state we usefully ascribe to an entity to understand its behavior as part of a rational way of getting its goals. Dennett has explicitly stated (1995) that sometimes it is possible for different ascriptions of belief to make rational sense of something's behavior. If one is a Dennettian, the fact that one's method of belief ascription can attribute lots of different possible beliefs underlying a given behavior set isn't automatically the embarrassment it would be for those using realist theories of intentionality. This is true both for talking about the beliefs of groups as collective entity, and for ascribing beliefs to the individuals in groups. In principle, the various different hidden beliefs about the cockfight one might come up with could all be "correct" attributions on a Dennettian theory.

For some, doubtlessly, the possibility of multiple different "correct" ascriptions is a reason to hold a more stringent theory of what a belief is than Dennett's. I think that even if one is an unreconstructed Dennettian instrumentalist, however, there are still several reasons one could find current interpretive practices too lax. First, even an instrumentalist would want to make ascriptions which generate *the most* true and the fewest untrue behavioral observations. Seeing which ones do that requires *comparing different* belief ascriptions and seeing which do better by this standard. One could look at different belief ascriptions which generate the same set of behavioral observations, and ask of each one which *other* behaviors should be expected from this ascription. Do we, in fact, see these behaviors in those circumstances? And what other heretofore undiscussed features do we regularly see in that culture? Might these be more consistent with one rather than another ascription set? None of this can be assessed, however, if interpretivists don't proposed and specifically examine the merits of the different interpretive attributions that I've argued are inherently legion in this sort of activity. Currently, however, most interpretivists seem to be content to come up with ascriptions that seem to account for their observations rather than specifically comparing and contrasting them with other models that also seem to do the job. Explicitly engaging in systematic comparisons could certainly make the practice of ascribing interpretations better and more stringent than it currently is.

The practice of comparing different interpretations is also currently hampered by the fact that many interpretive belief ascribers tend to be satisfied with

ascribing internal state models that are more vague and incomplete than is desirable. As I just mentioned, even an instrumentalist theory is interested in coming up with models which generate true behavioral consequences, and eliminating models which generate false ones. The problem with many current interpretive ascriptions is that they are not postulating belief-desire set models that are detailed enough to infer specific observational consequences from them. In the description of Japanese monkey performances mentioned previously, for example, Ohnuki-Tierney writes that, "the monkey and the outcast are the small eyes in yin and yang. For this reason, I think, even amidst the laughter at the monkey performance the audience is reminded, albeit vaguely, of their darker side, as represented by the monkey and the outcast trainer" (1984, p. 301–304). What does this tell us about what either the performers or the audience will do at a monkey performance? Indeed, what does it even tell us they will think? What does being vaguely reminded of your dark side entail? What features are present in their conscious or unconscious minds at this time? This description gives us almost no information whatsoever about what to expect about the people described this way. Specifying less vague belief-desire models, then, would also improve interpretive practice, even on instrumentalist theories.

Finally, even if one is not a realist about beliefs, and is more interested in finding predictive models than in uncovering "real" internal states, there are lots of reasons that some models can be considered superior to others that make the same predictions. Larry Laudan describes how historically, scientists, for good reasons, have tended to prefer models that make surprising predictions or are not *ad hoc* (1996). More importantly, however, it's clear that if we want our knowledge to be as unified and as systematic as possible we should prefer models that not only enable us to infer predictions that turn out to be empirically adequate, but ones that are consistent with all our other theories about the sorts of entities that generate this behavior. If theory/model A and theory/model B both yield the same observational consequences, but B is inferable from a well-supported theory C, and A is not, then B should be the preferred model, even if one is an instrumentalist. In the present context, imagine Bob is doing all he can to make sure his wife never comes in contact with his handsome, randy friend Ted. One model of what is going on might attribute to Bob the desire that his wife sleep with Ted, and the belief that the more he keeps them apart, the more mystery and allure his wife will have for Ted, making it more likely that they'll sleep together. Alternatively, one could attribute to Bob the desire to prevent his wife from sleeping with Ted, and the belief that if they met, they would find each other very attractive. Both models equally well account for Bob's observed behavior. The second model should be preferred to the first, however, as this second model is the more likely consequence of a well-supported theory about what inner states are like. This theory holds certain beliefs and desires are more likely to be the ones present, given human evolutionary history. Here again, equal observational consequences in a given realm need not mean

equally preferable theories for an instrumentalist. The same thing can be said for ascribing beliefs to groups. Ascription A, for example, might claim that a large airline company slashed its prices because it hoped to drive a smaller rival airline out of business and regain its monopoly. Ascription B might claim the corporation is cutting prices because it felt guilty about past price gouging. Here, ascription A is to be preferred because it is more consistent with general theories about corporations usually taking whatever actions they do in order to maximize profits (and perhaps more consistent also with theories saying this company is not the sort of collective entity that has any analog of "guilt" feelings). The sort of constraining theories discussed above, then, could help instrumentalists, as well as realists, select better ascriptions of hidden beliefs.

WHAT'S SENSIBLE IN FIC DESCRIPTIONS

Let me put the two different parts of this essay together, now, and summarize what I think is sensible in FIC descriptions and what's confused. Suppose someone claims that an all-male philosophy department fears hiring women, as this would be taken as a sort of symbolic self-castration. Is this automatically eye-rolling nonsense?

One of the main points of this paper is that such claims should not automatically be dismissed as muddled nonsense. There are a number of ways in which it makes sense to talk about groups having beliefs. And both groups and individuals can be sensibly ascribed difficult to access "hidden" beliefs. First, such a claim might be intended to mean that the *majority* of members of the philosophy department believe this. While this ascription might be wrong, there's nothing nonsensical about the collectivist part of this claim. What we'd have there is a straightforward statistical syllogism. Saying "the department believes X," here is saying that the odds are high that any given individual member does. We can then use everything we know about human psychology to give us predictions and explanations about what those individuals will think or do in given situations. Secondly, if one has a minimalist instrumentalist theory of what a belief is – and even if one holds a much more stringent functionalist theory of belief (as most cognitive scientists do) – then there is no reason, in principle, to think that *groups as entities* couldn't have beliefs and desires, just as individuals do. Where there may well be difficulties, in practice, is determining whether a given group actually functions like a rational agent, and whether it functions sufficiently like us to be able to use anything that much resembles "folk" or scientific psychology to say what the group will do. I will not explore this issue here. I tend to be skeptical that most groups have the requisite amount of internal organization to be rational enough for a folk psychology based understanding to be applicable. I do believe, however, that the matter is worth investigating. Certain dictator groups and certain small collectively controlled groups might

indeed posses the requisite rational coordination. And again, the *possibility* of many such groups existing cannot be ruled out.[9]

What about the somewhat strange content of ascriptions like the above? Can we ever be justified in ascribing such "hidden" beliefs and motives to groups and individuals, or is this an example of muddled undisciplined speculation? As before, there is nothing wrong with such ascriptions, in principle. From personal experience, we all know individuals can enter into mental states where what they see recalls features (or categories) far afield from what's perceived, and that these recallings can have powerful effects on behaviors. And if a group can come to be arranged so it behaves like an integrated organism or computer, there is no reason it couldn't also have hard-to-uncover internal states that effect its behavior.

The problem with such ascriptions is that in current *practice*, most theorists interested in making these sorts of ascriptions do not seem to be interested in learning the sorts of theories that could give us our best models of individuals' inner states. The behavioral and environmental strategies combined with the minimalist or antiquated theories of mind interpretivists currently seem to favor are capable only of identifying innumerable inner states that *could* generate behavior. The choice of a particular set of these as the one that *does* generate behavior (or which provides the best model of states which do so) tends to be an idiosyncratic or arbitrary one. Current interpretive social science ascriptions of hidden beliefs to individuals or groups should be treated as inspired or uninspired guesswork. In practice, if not in principle, such ascriptions tend to be muddled results of inadequate evidence.

I believe, however, that we have theories available to us that could, if utilized, help provide us with quite useful accurate pictures of individual's "hidden" beliefs and desires. Ascriptions to individuals in groups that were made on the basis of these theories have the potential to be highly trustworthy. I would encourage theorists interested in making hidden belief ascriptions to become more cognizant of developments in neuroscience, psychology, etc., and to make use of the well-researched models of which kind of mental contents tend to come to mind when. Ascriptions of mental states made to individuals on this basis could be quite sensible.

And hidden beliefs ascribed to groups? If groups can have beliefs, there's little reason to doubt they can have controlling internal states that are "hidden" in the sense that they are beyond what our usual folk psychological ascription strategies can uncover. What's less plausible is that groups are likely to really have hidden beliefs with *anything remotely like* the sorts of contents that interpretivists currently tend to ascribe. Groups are composed of internal parts (people) quite radically different from those that form the brains of individual humans (neurons). These parts are arranged according to quite different principles. The evolutionary pressures and learning histories of groups is likely to be quite different from that of humans. A group might well come to be organized in such away that certain rough principles of folk psychology are true of it (e.g., it "*does* what it *thinks* will

bring about what it *wants*" (Bennett, 1990)). But the burden of proof is surely on anyone who thinks that a given group can be organized such that it does anything like feel that it is "untamed nature to be culturalized by the monkey." Groups that "think" this way would have to have internal information processing structures that *mimic a human unconscious* structured along Freudian or Levi-Straussian lines. The idea that any group could come to have that degree of complex organisation, realizing just these sorts of structures, is so implausible that one should regard any such ascription to a group with even more skepticism than one should for individual interpretive ascriptions. Such ascriptions should, then, be regarded as muddled implausible speculations.

If groups can have beliefs, however, they'll have some kind of internal structuring that processes information and coordinates behavior. In principle, if a group of people could come to be organised to compute complex functions, it *could* also come to be organized to have the sorts of unrecoverable-through-folk-psychology, chains-of association-like structures that interpretive social scientists tend to be interested in. Such groups are unlikely to be internally organized along quasi-Freudian or some such lines, but it is worth investigating whether there *are* any other structures (or useful models of them) centering around hidden beliefs that actually are responsible for the behavior of entity groups. Any hidden beliefs possessed by such entities should not be expected to be uncovered using even the best theories of how "hidden" beliefs function in individuals. These sorts of collective "organisms" are too different from us. But if such hidden beliefs exist in groups, I suggest the best way to uncover them would be to use the group-level analogues of these "inner organization" theories used in cognitive science. These theories could be developed, as I describe above, by using the composite-construction and initial-conditions-plus-changes methods for inferring the structures of unseen "black box" mechanisms. In principle, then, groups could have these sorts of beliefs, and there are ways such beliefs could be uncoverable. There is nothing muddled about attempting to ascribe such beliefs. But unlike for individualist hidden beliefs, ascribing such beliefs in practice currently remains a long way off.

Let's summarize this summary and conclude. In principle, it's not muddled nonsense to speak of groups having beliefs. In principle, it's also not muddled nonsense to claim that either the individuals in groups, or groups as entities, have the sorts of unusual quasi-Freudian beliefs that interpretive social scientists are interested in uncovering. If tough-minded scholars automatically roll their eyes in disdain anytime such ascriptions are made, interpretive social scientists may well justifiably consider these scholars narrow minded. So considered, it is easy for interpretivists to ignore their potential good advice.

And FIC description interpretivists are in great need of good advice, While good FIC descriptions remain an in-principle possibility, in practice it's not clear whether or which entity groups can really usefully be said to have beliefs – much less the interesting hidden beliefs interpretivists seek. In practice, the methods typically

used by interpretivists to uncover hidden beliefs are inadequate for doing so. If interpretivists really want to look at collective beliefs and uncover hidden beliefs, they need to be utilizing information held by the "scientistic" scholars they eschew.

I suggest, then, that a better course to steer is to try to bring FIC description thinkers back into the fold of careful scholarly activity, by considering these ascriptions as something worthy of serious study rather than disdain. By treating such claims as serious social science, and working together, both interpretivists and tough-minded scholars would be in a better position to separate confused muddled ascriptions from claims that could really tell us something interesting about social behavior. For a long time, the social sciences have found themselves torn between C.P. Snow's scientistic and humanistic cultures. Having tough-minding scholars take humanistically-orientated FIC descriptions seriously, and having interpretivists open mindedly look at how their ascriptions might be improved with up-to-date scientific ideas about the mind might help Snow's breach come a little closer to being mended.

Todd Jones
University of Nevada Las Vegas
Department of Philosophy
4505 Maryland Parkway
Box 455028
Las Vegas, Nevada 89154-5028
USA
tjones@nevada.edu

Acknowledgements. I'd like to thank John Greenwood, David Biesecker, Brian McGlaughlin, Ron Mallon, Mike Bishop, Virginia Held, Josh Jorgenssen, Jackie Maloy, Brian Loar, and an anonymous referee for many helpful comments on earlier drafts of this paper.

NOTES

[1] Here, note that I'm discussing these three types of entity-group beliefs as a quick way of illustrating how such groups can exist. I am not attempting to give a sophisticated taxonomy of types. I do want to mention, however, how I think a sophisticated taxonomy would go. First, while the above discussion has individual people as the atomic unit out of which more complex collectives are built, it's important to recognize that a group can also be a building block unit out of which more complex groups can be built. I also think there are four ways that units (either individuals or groups) can be connected with each other in ways that create goal-seeking believing entities:
A. Units can independently, without interacting, jointly act in ways that tend to robustly bring about a particular end state in a range of environments. (This can be described as "having a goal.")

B. Units can come to interact in ways that produce a larger goal-seeking entity – without the other units being the primary shaper's of each unit's activity.
C. Units can interact in ways they do, because some units reward and punish the activities of other units (with hierarchical dictator groups being the clearest case).
D. Units can interact in ways they do because some units seek to take advantage of the behavioral activities of other units.
All of these ways of connecting could certainly be at work simultaneously in various groups of units. For a full articulation of this taxonomy see Jones (forthcoming).

[2] Indeed, if the functionalist/computational view of cognition is correct, and a suitably arranged group can compute any function a Turing machine can, then the limits of what a group can believe, desire, etc., are the same as the limits on computer cognition. Many theorists don't see how various types of thinking could be accomplished computationally. But other theorists (especially those who advocate what philosopher John Searle calls the "strong AI position") have not been shy about attempting to give functionalist/computational stories of all kinds of thinking behavior, from qualia, to self awareness, to emotions (see, for example, Churchland, 1988). If such functionalists stories are correct, then there is no reason to believe that suitably arranged groups could not have the full gamut of intentional states, from *de dicto* and *de re* beliefs to self-deception.

[3] For the record, note that my own view is that Dennett often misrepresents his own theory. For reasons described later in the paper, one can't possibly ascribe a belief, knowing nothing more than that the behavior is a rational agent (any behavior set can be the rational result of any belief if you make the surrounding beliefs and goals extensive and bizarre enough). One might, however, be able to ascribe beliefs by using little more than rationality assumptions if one combines this with some minimal proscriptions on certain types of belief-desire sets, and assumes the presence of certain others.

[4] Simulation theories of belief ascription (described in footnote 6) may be the exception.

[5] This example is just a recasting of philosopher Willard Quine's (1960) famous example. The idea that behavioral evidence, by itself, can never really confirm a particular belief ascription is a central idea in his celebrated claim that whether an alien's cry of "Gavagai" in the presence of a rabbit really means "rabbit" or "undetached rabbit parts" is, in principle, indeterminate.

The problem of underdetermination of theory by data is, of course, a problem all sciences have to face. A number of scholars, however (e.g., Quine, 1970; Rosenburg, 1985; Jones, 2000), have argued that this problem is much more serious for belief ascription.

[6] A somewhat less minimal way of ascribing beliefs to people is to use ourselves as models. Numerous theorists have suggested that in everyday (non-interpretive) belief ascription, when we want to know what someone else believes, we tend to look at what we would believe, were we in that other person's situation. We could infer what others' beliefs and desire are like by making analogies with our own beliefs in several ways. One proposal found in the belief ascription literature is that we attribute beliefs to others by performing a kind of *simulation*. Ascribing beliefs in this way requires very little prior *knowledge* about how people's minds work, or about the various primary and surrounding beliefs and desires they hold. All one has to do to see what they believe is to physically put oneself in their position – or imagine oneself in the other's position – and then check to see what beliefs and desires pop into one's own mind. If others' minds indeed work like ours do, and the simulation is a realistic one, this provides a pretty good indication that these thoughts are what appears in their minds in such situations (see Gordon, 1986; Goldman, 1993; Davies and Stone, 1995).

Other philosophers and psychologists have proposed that in ascribing beliefs to others we do make use of vague theories about how minds work (sometimes termed "folk

psychology"), and about the types of beliefs and desires people tend to have. When using such theories to ascribe beliefs to others, however, instead of using our observations of others to try to justify the positing of auxiliary beliefs one by one, we tend to make a blanket default assumption that the relevant surrounding interacting beliefs held by others in a given situation are similar to those that we have – unless there are specific reasons to believe otherwise. Which beliefs others will form on the basis of certain environmental exposure and certain previously held beliefs, then, are generally assumed to be the same ones that we would form in these circumstances (see Stich and Nichols, 1997).

Whatever the details of how we go about actually ascribing beliefs to others, it's clear that for some beliefs, if others really are like us, we can indeed use ourselves as models, both of what prior beliefs exist and which beliefs get formed in certain circumstances. But it's clear that however successful we are at ascribing ordinary beliefs to others using methods of using ourselves as models, we can't be confident about ascribing the sorts of states *interpretivists are interested in* in this way. The sorts of more peripheral associations people may be thinking of when they see something, are far too numerous for us to make confident guesses about, merely by seeing what our compatriots see. Imagining we were in Bill's shoes, we might easily infer that Bill thinks his garden hose is next to the rake. This method would not tell us that Bill is being reminded of the time his father bought him a pet snake, probably hoping the snake would bite and kill Bill.

[7] To begin with, the theories of the unconscious that interpretivists typically rely on are simply not well-established theories that we can have much confidence in. There are literally dozens of competing theories about what sorts of unconscious beliefs and desires lie behind behavior, and how they come to be there. Within the psychoanalytic tradition alone there are Freudian theories, Jungian, Adlerian, Horneyite, Sullivanian, Frommian, Reichian, and Eriksonian theories. There are also the Structuralist and Structural Marxist theories of Lacan, Levi-Strauss, Douglass, and Friedman. This lack of consensus about the nature of the unconscious ought to make one pause before using one of these aging theories as the basis for ascribing particular unconscious beliefs to people. In the face of such division, we should at least try to insure that the particular theory we do end up using is one for which we have a great deal of evidence. Unfortunately, most proponents of these theories eschew making systematic attempts to provide evidence for them (see Harris, 1979; Grunbaum, 1984). Worse, when attempts by independent researchers have been made to test the two most prominent theories, Freudian and Levi-Straussian, the results have consistently been stunning failures for both (see Harris, 1979; Erwin, 1993).

[8] These theories, in turn, are themselves supported by abductive reasoning, and still other theories.

[9] The sorts of group belief claims that are the least useful and most misleading are claims about individuals within a group, which are not the majority and don't control it. Such claims are often made to highlight certain surprising noteworthy practices of people in such groups. Someone might say, for example, "New York women wear sneakers to and from work and change into high heels at work." While this is certainly not true of the majority of women in New York, the prevalence of this practice may be surprising enough for someone to comment on, without it seeming to be worthwhile to think about how prevalent it is. The English language does allow enough vagueness in such matters, such that one isn't misspeaking by saying this. Still, such claims give limited information. It doesn't tell you which women do or how many do, and won't allow you to say much at all about what you can expect from New Yorkers. It is also easily confused with a claim about the majority. "The philosophy department fears castration," meant as a statement about some of the individuals in the philosophy department could also be confused with a claim about the majority of its members. It's also easily confused with a claim about the group as a whole. I'd recommend that social scientists not use "the x

believe y" statements at all for such non-majority non-holistic agent claims. Those sorts of collectivist claims are poor social science claims and poor journalism as well.

REFERENCES

AXELROD, R. & HAMILTON, D. (1981) The Evolution of Cooperation. *Science*, **311**, 1390–96.

BARKOW, J., COSMIDES, L. & TOOBY, J. (1992) *The Adapted Mind. Evolutionary Psychology and the Generation of Culture.* Oxford: Oxford University Press.

BEIDERMAN, I. (1987) Recognition by components, a theory of human image understanding. *Psychological Review*, **94**, 115–47.

BENNETT, J. (1990) Folk psychological explanations. In J. Greenwood (ed) *The Future of Folk Psychology.* Cambridge: Cambridge University Press, pp. 176–195.

BLOCK, N. (1993) Troubles with functionalism. In A. Goldman (ed) *Readings in Philosophy and Cognitive Science.* Cambridge, MA: MIT Press, pp. 231–254.

BOWER, G. & GLASS, A. (1976) Structural units and the reintegrative power of picture fragments. *Journal of Experimental Psychology: Human Learning and Memory*, **2**, 456–466.

BURT, R. (1982) *Towards a Structural Theory of Action: Network Models of Social Structure, Perception, and Action.* New York: Academic Press.

CHURCHLAND, P. (1988) *Matter and Consciousness.* Cambridge, MA: MIT Press.

CLARK, A. (1993) *Associative Engines.* Cambridge, MA: MIT Press.

COLBY, B., FERNANDEZ, D. & KRONENFELD, D. (1981) Toward a convergence of cognition and symbolic anthropology. *American Ethnologist*, **8**(3), 422–450.

DAVIES, M. & STONE, T. (1995) *Mental Simulation: Evaluations and Applications.* Oxford: Blackwell.

DENZAU, A. & NORTH, D. (1995) *Shared Mental Models: Ideologies and Institutions.* Unpublished.

DENNETT, D. (1987) *The Intentional Stance.* Cambridge, MA: MIT Press.

DENNETT, D. (1995) *Darwin's Dangerous Idea: Evolution and the Meanings of Life.* New York: Simon & Schuster.

DENZAU, A. & NORTH, D. (1995) *Shared Mental Models: Ideologies and Institutions.* Unpublished

EHRENCHEICH, B. (1991) *The Worst Years of Our Lives.* New York: Random House

ERWIN, E. (1993) Philosophers on Freudianism. In J. Earman (ed) *Philosophical Problems of the Internal and External Words: Essays on the Philosophy of Adolf Grunbaum.* Pittsburgh: University of Pittsburgh Press, pp. 409–459.

FAISON, S. (1997) China exports its own uncertainty. *New York Times*, October 26. 4 5.

FARRELL, W. (1993) *The Myth of Male Power: Why Men are the Disposable Sex.* New York: Simon & Schuster.

FODOR, J. (1994) *The Elm and the Expert: Mentalese and its Semantics.* Cambridge, MA: MIT Press.

GEERTS, C. (1973) *The Interpretation of Cultures: Selected Essays.* New York: Basic Books.

GILBERT, M. (1989) *On Social Facts.* Princeton: Princeton University Press.

GILBERT, M. (1996) *Living Together.* New York: Rowan and Littlefield.

GOLDMAN, A. (1993) The psychology of folk psychology. *Behavioral and Brain Sciences*, **16**, 15–28.

GORDON, R. (1986) Folk psychology as simulation. *Mind and Language*, **1**, 158–71.

GRUNBAUM, A. (1984) *The Foundations of Psychoanalysis: A Philosophical Critique.* Berkeley: University of California Press.

HARRIS, M. (1979) *Cultural Materialism.* New York: Random House.

HUTCHINS, E. (1995) *Culture in the Wild.* Cambridge, MA: MIT Press..

JOHNSON-LAIRD (1993) *Human and Machine Thinking.* Hillsdale, NJ: Lawrence Erlbaum and Associates.

JONES, T. (1988) Interpretist Social Science and the "native's point of view": A closer look. *Philosophy of the Social Sciences*, **28**(1), 32–68.

JONES, T. (2000) Ethnography, Belief, Description and Epistomological Barriers. *Human Relations*, **53**(1), 117–152.

LAUDAN, L. (1996) *Beyond Positivisim and Relativism.* Boulder: Westview Press.

LEVI-STRAUSS, C. (1963) *Structural Anthropology.* New York: Basic Books.

LIFTON, R. (1975) *Explorations in Psychohistory: The Wellfleet Papers.* New York: Simon & Schuster.

LINDHOLM, C. (1981) Leatherworkers and love potions. *American Ethnologist*, **8**(3), 512–25.

LYCAN, B. (1988) *Judgement and Justification.* Cambridge: Cambridge University Press.

MARR, D. & NISHIHARA, H. (1978) Representation and recognition of the spatial organisation of three-dimensional shape. *Proceedings of the Royal Society of London B*, **200**, 269–94.

MORRIS, T. (1994) The neural basis of learning with particular reference to the role of synaptic plasticity: where are we a century after Cajal's speculations? In N.J. Mackintosh (ed) *Animals Learning and Cognition.* San Diego: Academic Press, pp. 135–83.

NELSON, R. & WINTER, S. (1982) *An Evolutionary Theory of Economic Change.* Cambridge, MA: Harvard University Press.

OHNUKI-TIERNEY, E. (1984) Monkey performances: a multiple structure of meaning and reflexivity in Japanese culture. In E. Bruner (ed) *Text, Play, and Story.* Washington DC: American Ethnological Society, pp. 289–314.

PALMER, S. (1977) Hierarchical structure in perceptual representation. *Cognitive Psychology*, **9**, 441–74.

PELLS, R. (1997) *Not Like Us. How Europeans have Loved, Hated, and Transformed American Culture Since World War II.* New York: Basic Books.

QUINE, W. (1960) *Word and Object.* Cambridge, MA: MIT Press.

QUINE, W. (1970) On the reasons for the indeterminacy of translation. *Journal of Philosophy*, **67**, 178–83.

REYNOLDS, G. (1968) *A Primer of Operant Conditioning.* Glenview, IL: Scott, Foresman.

ROSENBURG, A. (1988) *Philosophy of Social Science.* Boulder: Westview Press.

SAPIR, D. (1981) Leper, hyena, and blacksmith in Kujamaat Diola thought. *American Ethnologist*, **8**(3), 526–43.

SATZ, D. & FEREJOHN, J. (1994) Rational choice and social theory. *Journal of Philosophy*, **9**(102), 71–87.

SCHAMA, S. (1995) *Landscape and Memory.* New York: A.A. Knopf.

SNOW, C.P. (1959) *The Two Cultures and the Scientific Revolution.* Cambridge: Cambridge University Press.

SQUIRE, L. (1987) *Memory and Brain.* Oxford: Oxford University Press.

WILSON, M. (1959) *Communal Rituals of the Nyakyusa.* Oxford: Oxford University Press.