Propositional Attitude Psychology as an Ideal Type

Justin Schwartz

1. Introduction[1]

Psychological explanation is widely thought to be a matter of ascribing intentional attitudes towards propositional contents representing facts or states of affairs. In the framework of propositional attitude psychology (PAP), as it is usually called, the attitudes include such states as beliefs and desires and the contents are given by propositions believed or desired. In this paper I take as my target the prevalent view that the content of intentional mental states is necessarily *propositional* in character. This view I reject holds that intentional psychology *is* just PAP, and that realism about mental states stands or falls with the thesis that agents actually instantiate propositions of a quasi-linguistic character interrelated in a way closely analogous to the sentential structure of natural language. This view is sometimes called the Language of Thought (LOT) Thesis, but following the usage of cognitive scientists, I call it Propositionalism.

Psychological theory, on this thesis, is a matter of describing the "language of thought." Thought is assumed to be structured very much like language, natural language to be the only model of a representational system powerful enough for capturing psychology. "There is no internal representation without an internal language," writes Jerry Fodor, ". . . [T]he language of thought cannot be a natural language. Nevertheless, facts about the latter provide us with some of our best data for inferences about the former" (Fodor, 1975 pp. 55, 100). (Fodor insists that thought is in a quasi-sentential language of thought rather than in sentences of a given natural language because, among other reasons, sentences in different natural languages can have the same content, e.g., "It's raining"; "Es regnet.") If Propositionalism is correct, denying that PAP is a descriptively adequate account of human behavior would then be tantamount to eliminativism or instrumentalism about the mental, i.e., to denying in some strong sense that people think and act on what they think. This is the conclusion drawn by some who deny the adequacy of PAP, such as Paul and Patricia Churchland, Stephen Stich, and Daniel Dennett.

I agree with the eliminativists that PAP (so understood) is strictly false. This does not commit us to elimination of the mental, however, because Propositionalism is also false. My proposal has two parts. I argue, first, that intentional states may have representational content of a nonpropositional — i.e., non-sentential — character and that those states need not be structured in the quasi-linguistic way that PAP holds. The correct psychology may be a nonpropositional intentional psychology (NIP). On the evidence so far, a theory of thought ought not be systematically modeled on natural languages, at least with respect to its syntax and inferential structure. It should be *in this sense* regarded as largely nonpropositional. If this is right, or even possible, PAP might be in principle replaced with our going out of our minds.

Second, I argue that, while false, PAP, unlike phlogiston theory, should not be eliminated outright. It should rather be regarded as an ideal typical approximation to a more correct theory in the way that we now regard classical mechanics (CM). Just as CM is false because it abstracts unrealistically from the real physical mechanisms which govern the dynamics of matter, PAP is false because it abstracts unrealistically from the psychological mechanisms which produce the actions of agents. In arguing for a similar conclusion, the Churchlands and others have invoked connectionist models drawn from recent work in artificial intelligence.[2] I sidestep the issue of connectionism and take my evidence from cognitive psychology, which seems to me less speculative and at any rate less fully explored.

Topoi **11**: 5—26, 1992.

Pace the Churchlands, however, neither theory, however, is *radically* false in a way that warrants their elimination and replacement. PAP still provides an approximate description of the real psychological mechanisms, just as CM provides a false but approximate description of the real physical mechanisms. (The psychological mechanisms may also be physical.) We may be fully realistic abut a nonpropositional system of mental representation, while still treating PAP as explanatory though false, in an ideal typical way. Thus we can retain PAP with its convenience and epistemic virtues.

Given the length of this paper I probably should not apologize for what is *not* in it, but some readers will miss direct or detailed discussion of some topics and arguments commonly discussed under the rubric of the language of thought. I don't explicitly address the details of such arguments for Propositionalism as presented by Fodor (1975, 1987) or Kim Sterelny (1990), although I do set them out, nor do I discuss the usual sort of arguments marshalled against Propositionalism which are summarized nicely by Patricia Churchland (1986, pp. 386—398).[3] The structure of my own argument is an inference to the best explanation. I note that languages of thought play a certain role in psychological theorizing and address the question of whether anything else could play that role as well or better. The burden of my case is that something can, namely nonpropositional mental representations. Sketching in how they can do this occupies the bulk of the paper; the suggestion is then that given the empirical evidence, this route is more promising than Propositionalism. The remainder of the paper is occupied with explaining how PAP can be retained as an ideal type.

The proposal to treat PAP as ideal typical has a number of advantages, some of which will become more clear below. I will note, here, that there are additional advantages which I do not have the space to explore. The proposal explains how we can attribute propositional contents to the nonlinguistic without paradox, and it gives us a unified account of how PAP works for both the linguistic and the nonlinguistic, i.e., ideal typically. The proposal is consistent with reductionism about psychology, in my opinion an advantage. In this context it also explains how we can ascribe the same propositional contents in the same way (ideal typically) to beings with different physical constitutions than ours without having to give up reductionism about the different NIPs which may be in our heads and theirs, i.e., it offers a line of response to the standard functionalist multiple realizability argument for the irreducibility of psychology (see Schwartz, forthcoming).

2. Proposition by any other name

Identifying PAP with a language of thought thesis will be controversial. Have I conflated the general thesis that psychology must advert to representational content of some sort, which we can call propositional content, with the specific thesis that this content is necessarily structured in quasi-linguistic form, as mental sentences which form an abstract system of representations distinct from the functional specification of mental states? There are two possible sources for this worry; one shallow and easy to handle, and the other deep and difficult.

The shallow problem arises from different disciplinary dialects for addressing similar problems. Psychologists, cognitive scientists, and some philosophers of psychology generally mean by "propositions" just quasi-linguistic structures as opposed to, e.g., mental images, mental models, and other internal systems not structured like sentences of a natural language. Other philosophers of mind, interested in philosophical explication, tend to mean by "propositions" just what animates sentences, images, models or other representations and makes them more than mere internal inscriptions.[4] In this idiom, "propositions" are just whatever specifies content. In *that* sense I do not deny that intentional psychology is propositional — in my terminology, intentional. The question then arises: how are these states animated? How do they specify content? According to Propositionalism they are animated, or do specify content, in virtue of being languagelike. I deny this and maintain that denying it does not eliminate the mental. As long as my target is clear, it doesn't matter if someone wants to use the other idiom and reserve "PAP" for what I call intentional psychology. Such a critic should read me as attacking LOT construed narrowly.

The deep and difficult motivation derives from a profound disagreement about the nature of psychology and about philosophy of mind. For the strain in philosophy of psychology I mainly consider, psychology is about *how people think*. Folk psychology is just a less rigorous and articulated sort of scientific psychology, but is part of the same explanatory project: the construction, testing, and application of hypotheses about what sort of internal states and processes produce overt

behavior. (Folk psychology has other purposes, but then so does scientific psychology.) In this debate, the dispute among the Churchlands, Stich, Dennett, Fodor, and me is about how good a hypothesis Propositionalism is. That view takes propositional attitudes to be hypothetical psychological states and considers the evidence for positing those states rather than others. Philosophy of psychology, here, is just the most speculative part of cognitive science, so call the shared assumption of these writers "cognitivism," the view that "intelligent behavior can (only) be explained by appeal to internal 'cognitive processes'" (Haugeland, 1981, p. 243). (Haugeland is right to parenthesize that "only." As I understand cognitivism it is consistent with an external contribution to internal cognitive processes.) Cognitivism is to philosophy of mind what naturalized epistemology is to the theory of knowledge.

Propositionalism, as a hypothesis about how people think, is rampant among cognitivists. David Hills, in his review of the literature, writes favorably of "the whole conception of mental representations or languages of thought," as if these were obviously the same thing, (Hills, 1980, p. 20). Field likewise argues for the necessity for a "representational system" which is basically a language of thought (Field, 1981). Fodor has recently restated his case for "why there still has to be a language of thought" (Fodor, 1987) and Propositionalism is defended at length in Sterelny (1990).

Many eliminativists accept this way of framing the issue, but hold that human behavior cannot be adequately explained in terms of propositional attitudes, and thus regard psychology *per se* as fruitless. Paul Churchland rejects the "central conception of cognitive activity [of folk psychology] as consisting in the manipulation of propositional attitudes" (Churchland, 1989, p. 11). In the context of an attack on Fodor, it is clear that what he rejects is what I call Propositionalism. "Why accept a theory of cognitive activity that models its elements on the elements of human language?" (Churchland, 1989, p. 16). Patricia Churchland takes as her stalking horse the "functionalist" (really, Propositionalist) view that

> . . . the intelligent organism is a sentential automaton, whose behavior is the outcome of a sequence of mental states (beliefs that *p*, desires that *q*, etc.), and the processing will be described in terms of the *semantic and syntactic* relations among the content-specifying sentences (Churchland, 1980, p. 188).

Elsewhere she comments: "it is doubtful that knowledge in general is sentential; rather, representations are typically structures of a different sort. . . . [R]epresentational structures are not sentences (propositions). . . ." (Churchland, 1987, p. 545). The parenthetical identification of propositions with mental sentences is very clear. In a different way, Stich also directs his fire against the view that beliefs are "mental sentences" with representational content (Stich, 1983). Unlike the Churchlands, who sometimes seem to countenance nonpropositional representation, what Stich would eliminate is representational content, so he attains a more radical position.[5] Among cognitivists the basic question remains: how do people think? Or at least: what internal mechanisms produce behavior?

For Radical Interpretation Theory, however, the strain of sophisticated behaviorism descending from Quine and represented by Donald Davidson, Robert Stalnaker or David Lewis, psychology is not about how people think; it is rather about *the nature of the mental.* Davidson, for example, sees "psychology as philosophy," *not* as science, where that means not just that it cannot generate precise predictions, but that it is a different sort of activity altogether — a normative and evaluative as opposed to a predictive and explanatory one (Davidson, 1980). Propounding scientific hypotheses about people's internal states is not the business of psychology or philosophy of mind. That business is rather finding system and order in our overt behavior, linguistic and other. PAP is not *psychology,* understood as a science of mind: it is more like literary criticism. Quine expresses the idea with the striking topiary image of the elephantine bush, whose internal structure is irrelevant. The issue is whether it is cut on the outside to look like an elephant. Its elephantinity reduces entirely to its external shape (Quine, 1960, p. 8). For Radical Interpretation Theory, internal cognitive processes don't count in a very strong sense. Settling accounts with the Radical Interpretation Theory is beyond my scope here. If a commitment to such a view is the source of someone's worries about framing the question of whether intentional psychology is *Propositionalist,* that will be the least of her doubts about the whole project.

3. Three kinds of psychology

Intentional psychology as a broader framework should be distinguished from PAP as a specific way of cashing it out. Consider first the Churchlands' and Stich's pro-

posed elimination of *folk* psychology. Now by itself this would not amount to radical eliminativism in the sense of denying that people think or have contentful representations causally relevant to the explanation of their behavior. It would not even necessarily amount to the elimination of PAP itself. The sorts of mental states or principles posited by a given folk psychology are not the only possible ones. Scientific psychological theories, such as decision theory or some cognitive psychologies such as Fodor's LOT, posit intentional mental states, or close enough; so do folk psychologies different from ours, such as the Greek, the medieval, or the Chinese. If we were to replace our current folk theory by some such alternative, we could only speak of more or less drastic revision of intentional psychology.

These sorts of theories may be regarded as different shorts of PAP; i.e., they share, broadly, the same quasi-linguistic framework. Our current folk psychology is just a member of the larger set of PAPs. All such theories explain action through the attribution of intentional attitudes towards propositional contents, where such contents are held to be languagelike in the relevant ways. What are these ways? To make a long story short, what makes a psychology *propositional* is that it invokes in its explanations sentential or quasi-sentential symbolic structures as internal states of agents. These have syntactic, inferential, and semantic aspects modeled on the sentential structure of natural languages.

Syntactically, they are composed from a finite vocabulary according to quasi-grammatical rules, and their interrelation and inferential structure is in some way governed by formal and inductive logic. Semantically, their representational character is a matter of their bearing meanings which determine their content insofar as it is determinate, in virtue of having, among other things, truth values and truth conditions in something like the way that sentences in a natural language have them. A particular theory in the PAP framework will invoke explicit or implicit generalizations such as the belief-desire generalization which Paul Churchland expresses formally as follows:

> $(x)(p)(q)$[((if x desires that p) & (x believes that (if p then q)) & (x is able to bring it about that q)) then (barring conflicting desires or preferred strategies), x brings it about that q)] (Churchland, 1989, p. 5),

where p and q are propositions describing some states of affairs. Thus quantifying over propositions, such a theory will operate within other constraints that individuate it as a distinct theory, including its peculiar generalizations and concepts, and within whatever constraints may operate on PAP generally, e.g., Davidson's (1980) Principle of Charity or Grandy's (1973) Principle of Humanity. Although at some high level of abstraction all PAPs may share a common core in virtue of which they are PAPs, nonetheless one such set of generalizations or constraints, or even one such theory, can be revised or abandoned without abandoning the framework of PAP.

One can, however, run the same move on PAP. This is just one way of explaining the actions of intentional agents by reference to attitudes of some sort towards representational contents of some character. Call the set of possible ways of so explaining behavior the framework of *intentional psychology,* where that is neutral on the nature of the attitudes and contents in question. In other words, intentional psychologies may be a larger set than PAP, including attitudes and representations of different sort than just beliefs, desires, and propositions. As Fodor (1987) admits, we may have an intentional psychology even if the narrower framework of PAP is incorrect, and in particular, whether or not contents are given by propositions of the specified sort, making thought languagelike in the relevant respects. Fodor calls this position "mere Intentional Realism." I will show that it need not be so "mere."

The specific attitudes (belief, desire, etc.) posited by folk or any other PAP are replaceable: nothing rests on cognitive attitudes being beliefs or conative ones being desires. More to the point, the representational states may not be propositional (quasi-linguistic). They may be more like maps or pictures, or more specifically, they may be or involve images, models, or other nonpropositional representations. Gilbert Harman puts it intuitively: "Not all human thought is in words. Our conception of ourselves in the world is more like a map than a story, and in perception our view of the world is more like a picture" (Harman, 1973, p. vii). Or as Nelson Goodman says in his criticism of a Propositionalist theory (Noam Chomsky's):

> What we call a language is a fairly elaborate and sophisticated symbolic system. . . . [B]efore anyone acquires a language he has an abundance of practice in developing and using rudimentary prelinguistic symbolic systems in which gestures and sensory and perceptual occurrences of all sorts function as signs. . . . (Goodman, 1972, p. 71)

While such systems may be antecedent and contributory

to language acquisition, we might not want to call them languages proper. "Acquisition of a language," Goodman sums up, "is acquisition of a secondary symbolic system" (*ibid.*) If having propositional contents is understood as distinct from, though analogous to, inner discourse carried on in something like a language of thought, I would add that any use of propositions is use of a "secondary symbolic system," or, in my terms, a secondary system of representations.

The elimination of propositions as the framework for representation, or the even outright elimination of PAP (which I do not advocate) would be innocuous as long as one had some other form of nonpropositional representation that fulfills the functional roles, including specifying content, which they fulfill in PAP. As long as we have some NIP with which to replace PAP, we remain in the intentional framework. Below I give examples of possible candidates for NIP. So far this is just a logical point. More radically, I propose that we perhaps *should* undertake some such replacement, at least for some scientific purposes. On my view, the contentful internal representations which are causally relevant to behavior, i.e., what Dennett, in the context of an attack on Propositionalism disparagingly calls "brain writing," are not — or not mainly — propositional attitudes (Dennett, 1978, pp. 39—52). Although these states may be to some extent representable by or translatable into propositional terms, if my thesis turns out right, a mature cognitive psychology will not systematically model them on natural language.

4. Eliminating the propositional character of content: Stich and Dennett

The first part of my proposal is that we can eliminate propositions without eliminating content or representational character. What this amount to may be illuminated by contrasting it with superficially similar proposals by Stephen Stich and Daniel Dennett, two writers who also doubt the descriptive adequacy of PAP, but take their skepticism in a direction that is eliminativist about representation or content itself. My concern here is not to refute their more radical eliminativism; it to show how my view differs and locate it more exactly.

In *From Folk Psychology to Cognitive Science,* Stich (1983) advocates a "syntactic theory of mind" which dispenses with the notion of content or representation. PAP, he holds, provides a good model or first sketch for understanding along quasi-linguistic lines the syntax and inferential relations among the functional states invoked in psychological explanation. He denies, however, that these states can be thought of as bearing semantic interpretation causally relevant to explanation of behavior. He attains this result by appeal to an "autonomy principle" (close to the "formality constraint" discussed below), according to which behavior is to be explained by internal states in the strict sense of "internal," abstracted from environmental context or causal history. For him, people are really much as we take computers to be, just syntactic engines crunching meaningless bytes without truth value or truth conditions.

I think this is wrong on both counts. I want to dispense with the quasi-linguistic logical syntax and inferential structured mental states we have. It seems to me a bad bet, in view of the empirical evidence, to take very seriously the analogy between the logical grammar of linguistic expression and that of psychological states. But, rejecting the formality constraint,[6] I would preserve the semantic notions of representational content captured by truth values and truth conditions, or close mimicries of these, whatever the syntactic and inferential structure of the states which have them. In short, my view is the logical complement of Stich's.

Stich and I agree that the glass is half full as far as PAP is concerned. We disagree which half is full. I say that it is the bottom half; Stich, the top half. In one sense I am more pessimistic than he. I doubt that PAP is a reliable guide to the syntactic structures and inferential patterns that a mature cognitive psychology would posit. In another sense, though, I am more optimistic. He says that what will fall away in such a psychology is causally relevant representational content. On my view, that is just what a more correct cognitive psychology should be expected to share with PAP.

In claiming, against Stich, that people do instantiate a representational system in which semantic content does explanatory work, I do not mean more than that people have internal states of some sort which are *representational* in some way and *systematically interconnected* in some manner. In particular, I do not mean that these internal states are necessarily relations to some abstract system of representation, where that implies that the internal states which constitute these states represent in virtue of relations they bear to a nonpropositional analog to a language of thought. Intentional states *may* be so related to such a representational system. The only relations they *must* involve, however, are their relations

to their objects and to each other. If some causal theory of representation is correct, what makes a belief that *p* a belief that *p* rather than that *q* or *r*, is a certain sort of causal role it plays. An internal state *N* is the belief that *p* in virtue of its functional role in a system in which the state of affairs *p* is causally relevant, in the appropriately belieflike way, to one's being in *N*. Representational character would be a two-place relation between an internal state and a state of affairs, not a three-place one in which the internal state mediates between a state of affairs and an abstract representational system. The "systematicity" I invoke is just whatever relations hold among the internal states. Nothing in my case against PAP depends on this particular sort of account of intentionality.

My proposal should also be distinguished from Dennett's view that the application of PAP to agents, or indeed of the concept of an intentional agent at all, is a matter of taking an "intentional stance" in which we treat certain systems *as if* they were intentional agents who realized propositional attitudes, but say that strictly speaking they do not (Dennett, 1978, pp. 3—22). Dennett's view is instrumentalist about intentional states with any representational content whatsoever, not just ideal typical about propositions. While he rejects the view that statements ascribing beliefs and other intentional states are useful falsehoods or lacking in truth-value, Dennett says they are like centers of gravity or classical mechanical phase spaces "in being *abstracta* rather than part of the 'furniture of the physical world' and . . . are *true* only if we exempt them from a certain familiar standard of literality" (Dennett, 1987, p. 72).

Unlike my proposal, Dennett's is a full-fledged form of radical eliminativism. That is because Dennett *identifies* intentional states with propositional attitudes, so that instrumentalizing the latter instrumentalizes the former. My argument, in contrast, turns on a distinction between intentional psychology and PAP; the proposal is that agents instantiate some NIP, and that we ascribe propositional attitudes mainly in an ideal typical manner as approximating ascriptions one would make in terms of the states posited by NIP. My treatment of PAP as ideal typical is similar to Dennett's treatment of it as involving *abstracta,* but since I hold that we instantiate nonpropositional representational states *of which* propositional attitudes are idealizations or abstractions, we can remain realistic about these states, although their nature is not antecedently known. They can be part of the "furniture of the world" — if materialism is true, the physical world — and their attribution can be as literally true as we like.

5. Propositions and intentionality

Rejecting Propositionalism allows us to stake out a possible, and I believe, novel third position between realism about PAP and eliminativism or instrumentalism about the mental. Why should anyone accept this position? The main reason is that there are promising candidates for NIP which can do duty for propositions in bearing content and which accord better with important results of empirical psychological inquiry. Below I discuss mental imagery and mental models. Parallel distributed processing, or the "connectionist" view, is another serious contender which I do not have space to consider. This undermines the main argument for propositions — that, as Margaret Thatcher used to say in a different context, there is no alternative.

Propositions offer a way to explain intentionality and human reasoning ability. It is in virtue of intentional states being attitudes towards propositions that they have content or can represent things in the world — or not in the world — and thus have these represented things causally affect our behavior. The idea is that representational capacity depends on our ability to instantiate and take attitudes towards propositions. Crudely, the fact that propositions can be about things makes it possible for intentional states to represent their objects. That propositions interrelate syntactically, logically, and inferentially allows us to manipulate these representations in useful ways.

Propositions (linguistically construed) represent in virtue of their semantic properties. On a standard view, they have meanings constituted by truth-conditions which can be satisfied by facts or states of affairs in the world; this gives them truth value and potential for referring to or being about these states of affairs. Contrast the proposition "$F = ma$" with the physical event of there being a certain quantity of force exerted proportional to the mass and acceleration of a particle. The latter has no truth value or truth conditions and does not represent. The former is about and represents the latter in virtue, partially, of its having that event as (part of) its truth conditions, and it is true in virtue of its truth conditions obtaining. The question then arises for cognitivists: is there any nonlinguistic form of representation which will allow us to capture intentionality in

this way or one that is just as good? Below I argue that there are several.

Propositions have a syntactic aspect as well, by which I mean that they are constructed by certain rules of order or well-formedness from a finite number of components which play the same functional role in each occurrence, allowing us to pick out particular elements as the same in different occurrences, and to determine whether the expression so constructed meets the syntactic conditions that permit it to be truth-valuable or to satisfy a set of truth conditions. Thus in "$F = ma$," "=" is the mathematical equality function. The syntactic rules governing the "=" inscription require one to place appropriate quantities (e.g., "F," designating the quantity force, or m, designating classical mass) on each side of it. The expression "$F =$" is not well-formed and cannot be evaluated; "$F = m^2a$" is well-formed but false because the quantities in question fail to satisfy the truth conditions for the expression; "$F = ma$" is well-formed and, if *CM* is correct, true. (It happens that *CM* is not correct because there is nothing answering to "m" or "F.") For cognitivists, the questions are, first, need we posit any syntactic structure to explain human behavior, or could we get by without such fine-grained differentiations? Second, if we need some syntactic structure, must it be quasi-linguistic? Below I assume that the answer to the first question is yes and argue that the answer to the second is no.

Finally, propositions have an inferential structure: they enter into logical and evidential relations with one another governed by certain inferential or computational rules. This is important to the notion of content or representation. The proposition that p can imply, contradict, or be consistent with q; q can be evidence for or tend to disconfirm r. What we believe, or, in general, what states we ascribe, seems to be in part a matter of these logical and evidential relations, and the relations depend in interesting ways on the content of the propositions. This puts a certain holism into mentalistic ascriptions. One ascribes mental states in related, systematic sets, and, other things being equal, revises attributions to make the logical and evidential relations come out right, as far as that is consistent with evidence of error or ignorance. The question for cognitivists is whether the inferential structure that people use to think is the same as that governing natural language, i.e., the same as they use to talk. Below I argue that it is not.

In short, the key idea of PAP is that intentional agents are essentially cognitive systems who have contentful psychological states which are causally relevant to their behavior. The agents are such because these states are attitudes towards logically and evidentially interconnected syntactically differentiated propositions which provide the content of those states. It is held that propositions are the only things which fit the bill, so characterizing something as an intentional agent requires appeal to propositions. If an agent S believes that p, where p is some state of affairs or fact, it is in virtue of S's being able to have such propositions as the content of a belief, and thus as part of the intrinsic structure of S's cognitive system, that S is a cognitive system at all. The notion of a propositional content and an intentional agent are thus linked. Failing a better story, intentional psychology is exhausted by PAP. This, of course, is just Propositionalism.

6. Nonpropositional representation

Any alternative to Propositionalism must satisfy three conditions. First, it must be different enough from PAP to count as an alternative. Second, it must have a semantics of sufficient power to capture a notion of representational content. Third, it must have an inferential structure flexible enough to explain how people reason as successfully as they do, and why they fail where they do. It need not have a syntactic structure at all, but if it does, it must do the work that the model of linguistic syntax does. There are alternatives which meet these conditions and which are more adequate to the empirical evidence than Propositionalism.

By way of illustration, consider two possible NIPs: mental imagery, which, though insufficient to provide a comprehensive alternative framework for psychology, allows us to see how such a framework might work, and mental models, which Philip Johnson-Laird contends does provide such a framework. (Connectionism provides another such possible framework.) If these are possible NIPs, we don't need propositional attitudes to have minds — even if PAP *happens* to be the way we have them. The negative evidence uncovered so far seems to count strongly against PAP as the best story, whether or not it strongly supports positive claims made by advocates of other models. As far as we can tell just now, some NIP is more likely to be right than PAP.

What it is for a form of representation to be nonpropositional is a matter of degree. No one property marks off the propositional from the nonpropositional.

Propositionalism holds that the only model powerful enough to describe a system of mental representation is natural language. Proposed alternative models of representation may diverge from this model in various ways, being more like, e.g., maps, pictures, or models than like language. To invoke the semantic-syntax distinction, two such ways would be to have a semantics or a syntactic structure radically different from that of natural language. I favor divergence on the syntactic and, further, the inferential dimensions, and will consider alternatives to thought as being governed by rules of deductive and inductive logic. There is strong empirical evidence that intentional agents merely approximate formal logical and inductive rules in their actual behavior. As to syntax, the fact that we report our mental states in natural languages structured with a certain syntax gives only very weak reason to suppose that they are produced according to the rules in which we report them. The semantics of a nonpropositional form of representation, I suspect, might not diverge much from that of natural language; if it did, so much the better from my point of view. That would only strengthen the disanalogy with propositions. And as long as such an alternative had a semantics, psychology would not skirt Stich's radical eliminativism.

7. Mental imagery

One useful approach is through an overview of the ongoing discussion around mental imagery, as articulated by Stephen Kosslyn (1980) and developed into a (nonimagistic) theory of mental models by Philip Johnson-Laird. This research program promises to tie together in interesting ways with neurophysiology on one hand and on the other with the empirical study of human reasoning abilities by Richard Nisbett, Lee Ross, Daniel Kahneman, and Amos Tversky (see Johnson-Laird and Wason, 1977). I don't propose a reversion to a modern-day version of classical empiricist impressionism or the "Idea" theory of Locke, Berkeley, and Hume, according to which thought is mainly the association of mental images. (Some psychologists do slip into this view, notably Kosslyn.) My point is just that there is good evidence that at least some of our thinking is imagistic in a nonpropositional way, and that seeing how images can do the semantic work a theory of representation needs shows that we don't need to supplement them with propositions *for that reason.* We may need propositions, or at any rate have them whether we need them or not, for other reasons. Moreover, Johnson-Laird's nonimagistic mental model theory suggests a way of extending the ideas of the imagery theory into areas where it might be thought that one needed propositions.

After setting out the case for and implications of mental imagery, I deal with some objections purporting to show either that imagery can't meet the conditions mentioned above and therefore fails to count as intentional, or either collapses into or is epiphenomenal upon propositional representation. These replies, although directed specifically at the objections to mental images, will stand in for similar objections to NIPs in general. Finally, I discuss mental models as a more promising nonpropositional, nonimagistic account and briefly address specific objections against it.

Daniel Dennett has savaged the idea that we can know from introspection that we have mental images (Dennett, 1969, 1981).[7] Even so, the evidence of introspection is not *worthless,* any more than the evidence of other kinds of immediate perception. It is just to be taken with a grain of salt. The existence of "quasi-pictorial" mental images is, however, strongly indicated by experiments which show, e.g., that the time it takes to identify whether two figures presented to subjects are identical is a direct function of the number of degrees of difference the two are rotated from each other (Shepard and Metzler, 1977, pp. 532—536).[8] Patricia Kitcher, in the context of criticizing the sort of anti-Propositionalist moral the Churchlands and I wish to draw from the case, puts the point: "since the relation between a linguistic representation and the thing represented is arbitrary, we would not expect the length [of time] of the symbolic manipulation to vary with the orientation of the stimulus object" (Kitcher, 1984, p. 95). Moreover, imagery finds neurophysiological support, among other places, from Martha Farah's discovery that image-generation capacity can be destroyed in isolation by brain lesions (Farah 1984). Kosslyn's imagism is further confirmed by the fact that, contrary to what one might have expected from brain laterality studies, although in accord with the view that Kosslyn has located an important aspect of human reasoning ability, this imaging component is located in the left posterior hemisphere of the brain, that associated with reasoning and analytical thought (Gardner 1985, p. 331). As far as retaining intentionality goes, we could in

principle replace the propositional apparatus by an imagistic one without denying that people think.

7.1. *Translatability and the primacy of propositions*

A first objection is that mental imagery is really propositional in character or that even if not, propositions are *primary* in such a way that they provide the basic form for any mental representation. If so, imagery fails to meet the first condition: sufficient difference from Propositionalism. A common argument to this effect turns on the claim that any information representable by mental images (or any form of nonpropositional representation) can be translated into or expressed in propositional form (Baylor, 1971; Pylyshyn, 1973; Anderson, 1978). Anderson presents a theorem based on the translatability of any information into a binary form computable by a Turing machine. Simplicity considerations then dictate, he contends, that we postulate only propositions — so construed — as the form of mental representations.

One may doubt that Turing-machine computability captures the quasi-linguistic notion of a proposition. This is the basis of now generally accepted arguments against Turing-machine functionalism. Even granting translatability into this or some other form, however, it does not warrant elimination of the translatable, or its demotion to "merely secondary" status. Translatability into propositional form no more makes images propositions than the fact that this sentence may be translated into Chinese makes it a Chinese sentence. The possibility of a Chinese translation of any particular proposition we state in English does not give us grounds for denying that English exists or assreting that Chinese is "primary." To the extent that mutual translation among natural languages is possible, they would on this view self-destruct like the two cats of Kilkenny. If it is thought pernicious for my point that this example relies on expressing the same proposition in different *natural languages,* consider Harman's map analogy. That the information contained in a map may be expressed propositionally does not mean that, in using a map to find our way about, what we use is a set of propositions. Finally, we ought to be cautious about granting translatability. Even inter-linguistic translation is treacherous. With nonlinguistic-linguistic translation the situation is even worse.

Zenon Pylyshyn (1973, 1981) has criticized Kosslyn's theory on the grounds that imagistic representation is "cognitively penetrable," i.e., what one imagines or what images one entertains is affected by what propositions one holds true. This is suppose to show that propositional representation is more basic. But it appears that the reverse is also true, i.e., that what propositions we believe is not affected by what images one entertains. Some of Nisbett and Ross' work on the effects of vividness of inference supports the idea that propositional belief is thus "imagistically penetrable" (Nisbett and Ross, 1980, pp. 74—76, 132—33, 290—91). If so, "cognitive penetrability" shows nothing about which form or representation is more basic.

Fodor suggests that while there are mental images, they are "constructed" to "accord with descriptions." That is, "we have access to a computational system which takes a description as input and gives, as output, something [imagistic] which satisfies that description," where this description is linguistically structured. Linguistic expressions, then, are primary; all images can do is to "facilitate" our "performance of certain kinds of tasks," e.g., "using the image rather than the description permits the subject to do the job of perceptual categorization in parallel rather than in series; he can check letter case and type at the same time" (Fodor, 1981b, pp. 80—81). But do we learn about letter case from a *description*? Could we give an adequate one for letter case? This is an enormous problem in building scanning machines which will type from manuscripts; if one has such a description, there's a fortune to be made in the office electronics industry. We *do* have the capacity Fodor describes, but it is evidently not primary to imaging.

7.2. *Some semantics for nonpropositional representation*

A more difficult argument concerns the need for a semantics in accounting for content. Propositionalism provides a reasonably well-articulated story of how intentional states can be contentful. There is little doubt that propositions are if anything is, and the strengths and weaknesses of various proposals for how this might work are well-known. No such situation holds with images or any other NIP. We have not a glimmer of a semantics for these, and therefore must go with Propositionalism as the only account of content. The objection, in short, is that imagery cannot meet the second

condition set out above: that NIP have a sufficiently powerful semantics to do what PAP does. Two routes are possible: one, which I undertake here, is to reply directly that a powerful semantics is possible; the other is to suggest, as Nelson Goodman does in the spirit of classical pragmatism, that it is beside the point for much of our cognitive activity, which should be seen as aimed at attainment of our purposes rather than representation of truths. This latter proposal deserves more attention than I can offer it here.

I confine myself to whether images in particular could have a specific basic semantic property, *viz.*, truth value. I choose this target because the best Propositionalist argument I know to the negative is offered by Fodor with specific regard to whether images can have truth values. It will be evident enough, I hope, how the sort of considerations I adumbrate here could be extended to other kinds of nonpropositional representation and perhaps to other semantic properties. "Having a thought *cannot* be simply a matter of entertaining an image," Fodor argues,

> For thoughts are the kinds of things that can be true or false. They are thus the kinds of things that are expressed by *sentences*, not words. And while . . . it makes a sort of sense to imagine a representational system in which the [imagistic] counterparts of words resemble what they refer to, it makes no sense at all to imagine a representational system in which the counterparts of sentences do. . . . I see no way of construing the notion that there might a language in which *truth* is defined for icons instead of symbols; in which, that is, "formulae" of the system are true of what they resemble. The trouble is *precisely* that icons are insufficiently abstract to be the vehicles of truth (Fodor, 1981b, pp. 65—67).

Fodor's worry about abstraction seems to derive from a worry about vagueness as a source of a sort of indeterminacy. To assign truth-values to property-ascribing representations, one has to know which property is ascribed. But

> any picture of a thing will, of necessity, display that thing as having indefinitely many properties; hence pictures correspond (and fail to correspond) in indefinitely many ways to the things they resemble. Which of these correspondences is the one that makes the picture true? (Fodor, 1981b, p. 67)

Fodor grants that images may refer (a semantic property) in virtue of resemblance, but he distinguishes between "thinking *of* something" and "thinking *that* something." The latter, he says, is what is important for truth value and thus for content.

Fodor's objection applies, *mutatis mutandis,* to any form of representation in which representing a state of affairs is in part a matter of one's having a mental structure which is, as Johnson-Laird says of mental models, "analogous to the structure of the corresponding state of affairs in the world" (Johnson-Laird, 1983, p. 156). Even so, the objection is misguided.[9] Set aside the premise which would have the imagist construct an "Iconic English," with words being replaced by images. The syntax of imagistic thinking might differ sharply from any discursive representation in ways relevant to its semantics. One point of introducing images or other nonpropositional representations is to explain phenomena which resist explication on the model of natural language. Iconic "English" just recreates that model with different sorts of "words."

More seriously, Fodor claims that images or icons can only refer to particular objects and cannot fill the role of propositions in specifying what propositions do as the objects of that-clauses, i.e., intentional contents which described states of affairs. This claim is false. Over my desk is a reproduction of a Vermeer depicting, as the title indicates, a "Young Woman with a Water Jug." It represents (among other things) the state of affairs which we can express propositionally by saying that *the young woman is holding a water jug.* If (as is not so) single images could not represent in this way, there is no reason to suppose that imagistic thinking would have to be in single disconnected images rather than in connected, sequentially ordered sets of them — in mental movies, as it were. Something like this, not necessarily imagistic, has been urged by schemata theorists, at least for some sorts of thinking (Neisser, 1967). Such a mental quasi-movie could nonpropositionally represent quite complex processes and states of affairs, as silent movies once did and better films do today. It could allow for property ascription: in one frame the woman is depicted as not holding the jug; in the next, as holding it.

Fodor's trump is the "lack of abstractness" of imagery. He offers the following example:

> . . . what would it be like to have a representational system in which the sentence "John is fat" is replaced by a picture? Suppose that the picture that corresponds to "John is fat" is a picture of John with a bulging tummy. What picture, then, are we going to assign to "John is tall?" The same picture? If so, the representational system does not distinguish the thought that John is tall from the thought that John is fat. A different picture? John will have to have some shape or other in whatever picture we choose, so what is to tell us that having the picture is having a thought

> about John's height rather than a thought about his shape? (Fodor, 1981b, p. 66)

If mental images or other nonpropositional representations will not give us clear and determinate contents for thoughts so that we can make such distinctions, then they will have no determinate truth values, and thus cannot serve for primary mental representations.

Now Fodor must think that propositions are in better shape as regards determinacy. This is a highly contentious claim. Without any appeal to Quine's (1960) indeterminacy of translation, the issue of coextensive predicates recreates Fodor's puzzle in a propositional context. All chordates are renates: "*X* has a heart" and "*X* has kidneys" are coextensional. Does this mean that there is no way of distinguishing between the thought that dolphins are renates and the thought that dolphins are chordates? The question appears on all fours with Fodor's. If we consider identificatory reduction, which gives coextensive correlates for much of what there is, Fodor's purported problem would be quite pervasive for propositions. How would one distinguish between the thought that rain is water and the thought that rain is H_2O? (The referential opacity of propositional contexts is no help here — at least, not where one knows that the coextension holds, as we do with water and H_2O.) Generally, if indeterminacy of this sort counts against images, it counts equally if not more so against propositions. If it does not count against propositions, neither does it count against images.

A more specific argument pertains to *how much* indeterminacy there is in imagistic representation. Granting that there is some, is there enough to block assignment of truth values? If Fodor's argument goes through, we can say nothing whatsoever of any degree of specificity about what Vermeer's "Young Woman with Water Jug" represents. This is patently false. The painting represents a young woman holding a water jug. It also gives us a good deal of additional information: that she is standing by a window; that the jug is in a flat bowl on a red tablecloth; that she is wearing a white headdress and flaring waistcoat. It is not *just* a representation of a young woman holding a water jug, but whatever else it is, it is also that, and insofar as the painting is taken as reportage, if Vermeer's model was holding a water jug, it is true; if not, then false — not that truth value is the most important thing about this particular representation. The indeterminacy nonetheless allows for assignment of truth values.

It may be, though, that Fodor's point is rather this. The old saw has it that "A picture is worth a thousand words"; here, the worry is that images convey *too much* information. Which of the various fine-grained propositions the picture represents is intended? By using the appropriate proposition, one can represent *just* the state of affairs of John's being fat or of the woman's holding a jug, without any additional information being conveyed, while one cannot do this with images, nor, perhaps, with mental models or other forms of nonpropositional representation.

First, however, images are not necessarily as coarse-grained as Fodor seems to think. Black and white images represent shapes and contrast without representing any particular colors. Stick figures represent the position of limbs without representing fatness or thinness. A map can represent relative spatial positions of areas without representing them as being of a certain shape or size. While there are limits on what can be represented in a given style or medium, one can devise modes of picturing which represent things almost as abstractly as one likes. And what we can devise, perhaps, nature can evolve. Say, though, that one can attain a unique sort of abstraction or fine-grainedness with propositions. It would not follow without further argument that one needs *just this sort* of abstraction for truth value. One can probably attain a different but no less unique degree of abstraction with other kinds of representation. With regard to the issue of truth value, the point is that even if propositions are especially fine-grained in a way that no image can be, nonetheless such nonpropositional representation may be fine-grained *enough* to bear truth values. What they represent may be somewhat indeterminate in that different propositions may satisfy them — as "*X* is equilateral" and "*X* is equiangular" are satisfied by the same set of triangles — but, to anticipate my proposal that PAP is ideal typical, that may be OK in *psychology* where explanation of behavior need not depend on as fine-grained an ascription as we can make with propositions.

Perhaps, though, Fodor intends the Wittgensteinian idea that images or pictures are not, as one might say, *intrinsically representative*. Wittgenstein says

> I see a picture: it represents an old man walking up a steep path leaning on a stick. — How? Might it not have looked just the same if he had been sliding downhill in that position? Perhaps a Martian would describe the picture so. I do not need to explain why *we* do not describe it so (Wittgenstein, 1958, p. 54e).

One wishes Wittgenstein had not been so coy about explaining why not. Without any pretense that he would think I had it right, I submit that the reason is that someone sliding downhill is not likely to be in the *same* position as one struggling uphill. (We lean forward going up hills and backwards going down them.) Given what we know about traversing steep paths, we take a picture to represent an old man climbing up in virtue of its representing him as having some of the properties which people have when struggling uphill. When, as in cartoons, someone is represented as sliding downhill in the same position one normally assumes when walking uphill, it's supposed to be funny. (The knowledge in question need not be descriptive or propositional; it may be motoric, as one knows how to ride a bicycle.) If the Martian reads the picture incorrectly, it may be because he does not know how people climb steep paths down here.

If Fodor means that images aren't intrinsically representative in virtue of the structural analogies between the image and what it represents, but only represent in a context of other knowledge of causally relevant factors, of course! The same goes, however, for words and sentences. Outside the right context, an inscription (whether external or "brain-writing") isn't even an image or a proposition. It's not representative at all, but, as Wittgenstein says, "dead." An image may represent in virtue of structural analogies with what it represents, but if those analogies occur accidentally, the inscription is not an image of what it's analogous to. If we find on Mars a perfect wind-etched replica of Mount Rushmore, the shapes there no more represent Washington, Jefferson, Lincoln, and Teddy Roosevelt than the equally miraculous wind-etched inscription below, orthographically replicating the Lincoln Memorial text of the Gettysburg Address, represents a statement of the reasons for which Lincoln thought the Union fought the Civil War.

None of this shows that imagistic representation is primary in the way that Propositionalism holds propositions to be. Indeed, there are good empirical grounds for thinking otherwise. These have to do with how well imagistic theories account for observed human capabilities, not with whether images can have semantic properties. That they can and do. Thus the field is open for some form of nonpropositional representation — not necessarily an imagistic one — which might displace propositions in our psychological theory as the primary bearers of content.

8. Mental models and semantics

One such is suggested by Philip Johnson-Laird (1983, 1988). Building on studies of human difficulties with formal deductive reasoning, he argues that explanation of human cognitive and reasoning capacities and weaknesses makes it necessary to postulate "mental models," nonimagistic structural analogs for external phenomena. Rather than attributing to subjects inadequate grasp of the rules of some formal logic, Johnson-Laird suggest that people normally reason by use of models which do

> not have an arbitrarily chosen syntactic structure, but . . . [a structure] that plays a direct representative role since it is analogous to the structure of the corresponding state of affairs in the world — as we perceive or conceive it" (Johnson-Laird, 1983, p. 156).[10]

Mental images might, he suggests, correspond to quasi-perceptual imagined "views" or projections of models.[11]

What sort of thing might such a model be? The chemist's use of a ball-and-stick model of atomic structure is an external example; the Newtonian physicist's model of a billiard-ball universe would perhaps be an internal if highly self-conscious one. To clarify the idea, Johnson-Laird contrasts the following ways of getting out of a maze:

> You come to a turning . . . You recognize that you have been at this point before, and, in your imagination, you turn right, proceed down an alley, and are confronted by a dead end. So this time you decide to turn left. What you did, I assume, was to reconstruct a route through the maze on the basis of a mental model of it. You may hardly have experienced any imagery at all. . . . In [any] case, there was nothing "propositional" . . . about your reasoning; there was no process based on the representation of verbal propositions. . . . However, you might have made your decision in a very different way. You might have recalled that the way to get out of the maze is to keep turning left. . . . This method makes use of a mental representation of verbal propositions (Johnson-Laird, 1983, pp. 154—55).

The theory is explicitly semantic: the formal analogies between the models and their objects are supposed to explicate the notions of truth and reference. Johnson-Laird has developed an elaborate model-theoretic "procedural" semantics for mental models, the evaluation of which is a topic for another study. Some fans of the "formality constraint" (see below) will find the invocation of semantics obnoxious. Others still find the semantics inadequate, but not fatally so. A resemblance-

based model-theoretic semantics could be amplified or replaced, say, with a causal theory.

What is to the point here is that on Johnson-Lairds' view neither the semantics nor the syntax of thought, much less its inferential structure, is propositional. Rather than operating on abstract propositions in which terms designate sets of individuals falling under certain property ascriptions, mental modeling involves operations on a structure composed of internal states which stand for specific individuals who satisfy certain descriptions. Thinking involves shuffling about tokens in thought experiments, not drawing (propositional) conclusions from (propositional) premises according to the rules of logic.

In the maze example, one creates in one's head a scale-model map of the maze in which specific tokens stand in relation to one another spatially (or in some appropriate representation of spatial relations) as the paths in the maze stand to one another. This is quite different from encoding the structure of the maze in a quasi-linguistic manner, where one instantiates or bears some relation to the proposition, "One can get out of the maze iff one always turns left," and when faced with a choice about which way to turn, one reasons thus: "I want to get out of the maze; to do this I must turn left; so I should turn left."

8.1. *Syntactic structures*

Johnson-Laird's proposal has implications for both the syntactic and the inferential structure of human thought. Now there is no a priori reason why a system of non-propositional representation would require syntactic differentiation at all, much less a sort modeled on the syntax of natural language. The components of representations and their ordering might not matter as long as they provided enough overall structure to do the causal work they should in producing behavior. Someone who claimed this might maintain something like what Fodor calls the "fusion story" (Fodor, 1981a, p. 179) or Field the "orthographic accident" story (Field, 1981). If intentional states are thus "fused" and not syntactically differentiated, Propositionalism fails spectacularly as an account of intentional psychology. Not only would mental states not have the syntax of natural language, they would have no syntax at all. That would be fine by me, but Fodor's and Field's empirical objections to this view seem persuasive. If they are right that thought is compositional, intentional psychology will probably have a syntactic structure to which we advert in thinking.

What Johnson-Laird's theory suggests, however, is that the syntactic structure in question need not be analogous to that of a natural language, whatever that might be. The rules which describe the differentiation and ordering of the elements of a system of mental models might diverge widely from those governing linguistic expressions. The chemist's red ball may represent carbon in whatever different molecular models or places in the same model in which it occurs, and the order the red ball may have in relation to colored balls representing other sorts of atoms may be strictly regimented. The rules describing the relations here, though, are those of *chemical theory,* not English. And if people do not (or do not mainly) think in English or in a language of thought, but rather by the manipulation of mental models, there is no antecedent reason to suppose that the rules describing the manipulation of such models are like the syntax of natural language. If, moreover, Johnson-Laird is right that the inferential structure of human thought is not logical, it is plausible as well as possible that the syntactic structure of thought is nonlinguistic.

8.2. *Inferential structure*

If thinking is largely a matter of operations on mental models, this would account for the difficulty people have in using conscious and explicit rules of logic. For if mental logic accurately described human psychology, it is hard to understand why logic is so hard to *learn.* The mental models theory also better accounts for the systematic deductive errors to which reasoners are prone (Johnson-Laird, 1983; Kahneman, Slovic, and Tversky, 1982). On Johnson-Laird's account, the actual operations the mind performs on models often fail to preserve truth or to produce the conclusions indicated by inductive probability. His theory predicts better than a "mental logic" picture the actual (good and bad) inferences people will make. The theory accounts for two features of actual reasoning on which mental logic theories run aground. One concerns the content-sensitivity of human inference, i.e., the fact that (to shift to the case of overt linguistically expressed argument), what the premises *say* affects our ability to draw proper conclusions in ways one would not expect if we were

operating on content-free logical forms. The other concerns the result that the difficulty of reasoning problems (measured by time of solution) correlates nicely with the number of model elements and models required to run a mental "simulation," regardless of whether the linguistic expression is simple or complex.

The "mental models" account was originally designed to provide a more empirically adequate theory of deductive inference. The Propositionalist can say that people do reason with first order logic and Bayesean probability theory, only rather poorly, but Johnson-Laird's account has the virtue of explaining why we are so bad at reasoning according to self-conscious norms; namely, because we do not instantiate those norms in our functional wiring. It thus offers a way of meeting the third condition for an adequate NIP, that it account for human reasoning capacities (and incapacities). Using mental models, Johnson-Laird shows how various kinds of nonpropositional representation can bear logical and evidential relations to one another in virtue of their syntactic structure, i.e., how one can reason using mental models without recourse to logical rules operating on propositions.

I cannot here spell out Johnson-Laird's theory of inference in detail. Consider syllogistic reasoning as an example. The syllogistic comprises a list of 64 valid and invalid inferential patterns, e.g.:

> All artists are beekeepers.
> All beekeepers are chemists.
> So, all artists are chemists.

Johnson-Laird argues that it is possible effectively to test the validity of such inferences without recourse to logical rules as follows (Johnson-Laird, 1983, pp. 64—101). One creates a model of the situation described in each premise by imagining individuals who fill the relevant roles and combining information from each step. One need not use "actors"; a model can be constructed using imagined chips or other symbols — even words. (The number of individuals imagined does not matter.) Combining the information in the premises, one gets something like this:

Artist = beekeeper = chemist
Artist = beekeeper = chemist
Artist = beekeeper = chemist
(beekeeper) = (chemist)
(beekeeper) = (chemist)
(chemist)

One tests the conclusion by looking for a counterexample, an interpretation of the premises that rules out the state of affairs modeled in the conclusion. Here one can see by simple inspection of the model that the conclusion follows. More complex syllogisms will require more than one model. Finding a counterexample refutes the model; not finding one, of course, may just mean one hasn't found the problem yet. The proposal accurately predicts which of the 64 syllogisms will be most difficult for people to solve in terms of (among other things) the number of models testing them requires (Johnson-Laird, 1983, pp. 101—104).

Johnson-Laird argues that such an approach can be extended to nonsyllogistic inference and many other aspects of thought: "Vision yields mental models, and the control of movement depends on mental models. Similarly . . . the process of understanding leads to models of states of affairs that are described" (Johnson-Laird, 1988, p. 231). Reasoning then does not require propositions or rules of inference. Moreover, the hypothesis that people do reason in something like this manner is supported by experimental evidence.

8.3. *Mental muddles?*

Lance Rips offers three objections to any mental models theory as an alternative to Propositionalism. The first is that such theories violate the "formality constraint" that cognitive psychology must do without semantic notions like truth and reference and confine itself to structural features of representations internal to the representations themselves (Rips, 1986, p. 265). But, as the programmers say: that ain't a bug, it's a feature. As Stich (1983) shows, this sort of psychological individualism amounts, for all intents and purposes, to radical eliminativism about the mental. One ends up discussing how content-free computational operations in a language of thought produce bodily motions of an organism, narrowly construed and without reference to the external environment or causal history of the agent. That is a legitimate topic of inquiry, but *contra* Stich, not the only one; intentional psychology is also legitimate. So cognitive psychologists ought not respect the formality constraint if they want to keep intentional psychology at all. The mental models theory may not provide an adequate semantics, as Rips further objects — in Johnson-Laird's version, because of its neglect of the causal dimension, it does not. It might be supplemented,

however, with a causal (or some other) story about how the models represent, and saying that the account suffers from inadequacy in doing the job is a different matter from saying that the job ought not be done at all.

Rips' second objection it that the theory is psychologically unrealistic because with complex domains involving many variables, "mental simulation may be too hard" (Rips, 1986, p. 270). He cites evidence from research in AI modeling of simple dynamical behavior of balls that "the burden of building a *complete* description of possible states is too onerous outside very small domains and is too restrictive a style to capture all of the ways that people use qualitative physical knowledge" (Forbus, 1983, cited in Rips 1986, emphasis added). As Rips expresses the objection,

> Unless you know a great deal about a system, even qualitative simulations are off limits, given the normal memory and processing limits humans have to contend with. For any system that is the least complex, we're thrown back on crude rules of thumb that are far from literal simulation. . . . I suspect that although most people know a little about what the parts of their car do, if anything goes wrong they have to rely on such simple-minded [and propositional] heuristics as "if nothing happens when I turn the key and if it's a cold day, then maybe it's the battery" (Rips, 1986, p. 271).

As Rips notes, *complete or fully adequate* simulation is likely to be difficult or impossible given our computational limitations, so that such modeling will not give good results. Johnson-Laird will agree: it is a feature of mental models theory that it explains human cognitive limitations and failures. Whether the theory can also explain the successes is an empirical question. It is not decided by observing that we cannot be ideal cognizers if it is true. We're not. That's part of the point.

Rips' final objection is that mental models theory collapses into Propositionalism. The models

> have to respect properties that seem much like those of standard logical syntax. The equal sign in the syllogism example has to be interpreted in a way that enforces the idea that the two tokens refer to the same individual. . . . [So] manipulation of mental models isn't fundamentally different from manipulation of mental propositions. To the extent that rules that operate on models are sensitive to these logical constraints, they just *are* inference rules (Rips, 1986, p. 279)

Here we can invoke a distinction between rule-governed and rule-corresponding behavior. Philosophers have long distinguished between rule-*following* and rule-governed behavior: matter is governed by the dynamical laws; people follow traffic regulations. The distinction I suggest is between behavior produced because the subject is governed by a certain rule, and the same behavior produced in a different way. The fact that manipulation of mental models is sensitive to logical constraints in that it produces similar outcomes to manipulation in accord with logical rules tells us nothing at all about how manipulation of mental models proceeds. The same effect may have different causes. Indeed, one would expect, since logic really is normatively superior, that thinking would face strong pressures to roughly approximate it. Those pressures need not operate through bringing humans to instantiate mental logic.

There is, however, a deeper reply, which Rips himself acknowledges without appreciating its force. "Of course the inference rules that mental models employ are nonstandard," he admits, "and such rules may certainly be worth study" (Rips, 1986, p. 279). The real point of the mental models theory is not that thought proceeds independently of inference rules, but that it's not logical. Certainly there are regularities in the manipulation of mental models, and if Rips wants to call these regularities "rules," he can have the expression. Since these regularities only approximate the rules of logic at best, mental models does not collapse into the mental logic version of Propositionalism. It's not that people think in logic, but badly. The evidence suggests that we don't thing in logic *at all.*

9. A place for propositions

It would be premature to advocate Johnson-Laird's mental models theory (much less Kosslyn's imagism) as the correct cognitive psychology. Johnson-Laird's is not, moreover, the only such story. Connectionism is another contender, perhaps integrable with the mental models theories. Any empirical theory of deductive reasoning also awaits integration with the parallel work on inductive reasoning by Kahneman, Tversky, Nisbett, and Ross. Here Johnson-Laird's theory is a particularly promising candidate; certainly his results demand explanation. To whatever extent any of these accounts proves tenable, they suggest that Propositionalism is disconfirmed — and in the way I urge, not in the way that Stich urges.

In drawing rings around the fact that people instantiate representations of a nonpropositional character, I should not want to deprecate the fact that humans,

whatever else they are, are language users, and some human cognition involves linguistic representation. The place of propositions, I suggest, is quite naturally in accounting for whatever part of our thought *is* linguistic in character, and surely this is not negligible. Theoretical science, for example, involves a large propositional component. On a commonsense level, we surely do sometimes if not always think in a natural language.

To thus countenance propositions is not to take back the claim that Propositionalism is false. People may sometimes shuffle about sentencelike structures without thought being propositional through and through, all the way down, or at the bottom. Allowing propositions to play a literal role in such explanations in no way countenances PAP as a general framework for explaining behavior. Propositionalists like Fodor similarly countenance mental imagery as a convenient way of manipulating propositionally represented information. In the same spirit I can allow that some thought is propositional but deny that the primary forms of representation are propositional or that thinking is, for the most part, performing logical and inductive operations on propositions. Thus we can take propositional representation literally without taking PAP literally.

10. The ideal typical character of propositional attitude psychology

There is strong empirical evidence that people do not mainly think by depicting states of affairs in sentencelike psychological structures governed by mental logic. At least the syntactic and inferential dimensions of PAP will probably have to go. If it is false, what relation might the propositional attitude framework bear to a successor framework, insofar as we can say without actually having such a successor in hand? One might say with the Churchlands and Stich that it should simply be abandoned, either altogether or for serious scientific purposes. If it's false it's false, they say, so over the side with it. Unlike them, however, I am not interested in abandoning intentional psychology along with PAP, but in replacing one sort of intentional psychology with another.

In part this is because the explanation of action by the ascription of propositional attitudes is far more persistent and successful than one would expect if the framework really were "radically false" in the way that Paul Churchland claims. PAP is deeply entrenched, highly successful, and virtually universally used in ordinary — and now that behaviorism is out of fashion — scientific psychological interpretation, and is so because of its great explanatory and predictive power and other epistemic virtues: simplicity, coherence, and flexibility. Any successor framework should account for these features, and if possible, preserve them.

10.1. *Ideal type explanation*

The example of classical mechanics (CM) gives us a model of how to treat an entrenched and (within broad limits) successful theory which is nonetheless strictly false. The second part of my proposal is that propositional attitude psychology is best understood as an ideal typical approximation to whatever sort of representation people actually instantiate, in something like the way that CM is an ideal type approximation of relativistic mechanics (RM). Reference to propositional attitudes would then be understood on the lines of the ideal typical way that we take reference to classical mass in light of relativity theory. We would hold that in the most exact sense agents do not generally represent the world propositionally, but the talk of such representation would be, like classical mass, admissible for explanatory purposes as a handy approximation to a correct description of actual representation.

How ideal type explanation works is little studied and ill-understood. While the issue has received some attention, I cannot pretend to have an articulated account in which to frame the proposal that psychology be regarded as ideal typical. I shall have to proceed by way of analogies. CM is strictly false in that there is nothing that has the properties it ascribes to its central posits, e.g., classical mass, force, or absolute space and time; its basic laws, e.g., Newton's Second Law, are therefore false. These concepts are replaced with relativistic concepts of mass-energy, geometricized inertia, and curved relativistic spacetime, governed by different laws. We do not, however, treat CM as we did phlogiston or caloric theory — abandoning any scientific use of the theories along with their entities and laws. We use CM for most purposes because it is mathematically simpler and perfectly adequate to dealing with macroscopic objects over short distances and at low velocities.

Nor do we regard CM in the purely instrumental way that Milton Friedman (1979) suggests we take rational actor theory in economics. Rather, we treat it as an

idealized description of the relativistic mechanisms under certain conditions. The initial and boundary conditions under which we use CM, where phenomena occur in spacetime regions small enough to be treated as flat and have velocities negligible compared to that of light, very closely approximate the conditions under which CM (or a close analog) is a limit case of RM, namely, where spacetime is flat and c is infinite.[12] Moreover, on a causal theory of reference we can say that CM refers to mass-energy when it posits mass; the description was wrong (not catastrophically, or the theory would not have been such a success), but the designation was right (Field, 1973; Putnam, 1975). This helps explain why CM works as well as it does despite being false, and underwrites our use of it in an ideal typical but full-bloodedly explanatory manner.

On this analogy, PAP likewise would be ideal typical rather than black-box instrumental. Reference to propositional is pragmatically convenient, and good enough for most purposes — speaking with one's spouse, persuading the electorate, discussing psychological theory. Moreover, it offers an idealized description of the actual mechanisms in virtue of which people represent which, though no more strictly accurate than CM, is close enough to account for the fact that PAP does as well as it does. These analogies might include the following:

(1) NIP may preserve the central attitude-content distinction basic to the structure of PAP;

(2) NIP has a semantic or representational structure which closely mimics the truth-conditional (or possible-worlds) semantics of natural language, i.e., NIP posits internal states which, though not propositional attitudes, have representational content;

(3) The inferential structure of NIP mimics, though less closely, the logical structure of natural language. It often fails to preserve truth or come out with the best probabilistic conclusions, but it does so often enough to allow us to get around.

Doubtless there will be other analogies as well. As far as *explaining* behavior goes, PAP is no worse off than any other sort of ideal type explanation in the physical or social sciences. It gives us explanatory information — albeit in idealized form — relevant to understanding the mechanisms which produce action. Physicists' appeals to frictionless surfaces or point masses are not essentially different. In this sense Dennett's talk of *abstracta* is just right.

What makes PAP ideal-*typical* and not just an *idealization* is its implicit invocation of an ideal model or *type* of agent, who, minimally, thinks in propositions and whose thought is governed by logic. To put it another way, this ideal type is such that its behavior is produced by internal states the structure and relations of which are analogous to the sentences of natural language. The difference between a mere idealization and a full ideal-typification (to coin an ugly expression) is a matter of degree: if one takes a complex phenomenon, idealizes it sufficiently, and characterizes the result in terms of an abstract theory which onesidely accentuates certain features, one has an ideal type. The characterization is vague, but it reflects the difference between, e.g., abstracting from friction and a full-blown appeal to Newton's laws. One might regard cases like the former as degenerate ideal types.

This argument presupposes a realist account of explanation which sees explanation as a matter of elucidating the causal or probabilistic mechanisms which produce the explananda. I do not defend such an account against its main rivals, the empiricist and pragmatic accounts. But I do not need to. Pragmatists about explanation, who hold that explanation is entirely relative to context and purpose, will be happy to tolerate PAP in many contexts and for many purposes. Empiricists who think that the deductive-nomological model of explanation stands, and that explanation and prediction are therefore perfectly symmetrical, can acknowledge that ideal type explanations allow for adequate prediction, and are therefore to count as explanatory in this sense. Their premises will be false but approximately true.

10.2. *The rational actor analogy*

Some further analogies may help make the proposal a little less bald and a little more plausible. Consider the use of rational actor theory in social science; roughly, the procedure of predicting and explaining behavior by regarding people as if they were ideally rational self-regarding utility maximizers. Rational actor theory is fairly predictive but ideal typical, in that people are not like that. Actual agents only approximately instantiate the motivational and cognitive structures of ideally rational economic agents. Their motivations and cognitive capabilities are far more complicated but it is convenient to abstract from these for many purposes.

As long as actual structures tend to result in behavior which approximates that of *homo economicus*, no harm

is done: we can use rational actor theory to predict behavior. If the actual structures do not deviate too far from those of *homo economicus,* we allow that appeal to ideal rational actors is explanatory as well as predictive. As a matter of empirical fact, these approximations hold to a high enough degree that for many purposes we can use rational actor theory to predict and explain human action. One way to understand PAP is as working in much the same way as the rational actor model in social science. Ascribing propositional attitudes to agents is also a fair way to predict behavior, and no less explanatory in principle, even if agents do not actually instantiate exactly these structures, than attributing to them the cognitive and motivational structures of ideally rational actors. As long as their actual representational structure approximates ideal typical propositional ones closely enough and results in behavior enough like that predicted by PAP, we can allow the use of the latter much of the time without worries about what is actually "in the head"; as with rational agent theory, for certain purposes we can abstract from what is in the head and ascribe propositional attitudes to predict and explain behavior.

The analogy further allows us to understand how we can use an ideal typical model of something for which we lack a more adequate theory. Given our current relativistic perspective in physics, we see that pre-relativity classical physicists did the same thing; only they did not know, nor for the most part believe, that CM was ideal typical. In the case of rational actor theory, it is not as if we have an adequate theory either of human behavior in general or of the specific sorts of behavior which we use rational actor theory to explain, and in many cases we have a shrewd guess that any such theory might be too cumbersome to use for many purposes in terms of necessary evidence and complexity. Rational actor theory allows us to make testable predictions based on deductive inference from a handful of postulates and assumptions, and these are for many purposes good enough to warrant use of the theory. PAP likewise permits us to explain behavior on the good hypothesis that people do represent themselves and the world around them without our knowing antecedently just how they represent it.

It may be objected that rational actor theory presupposes PAP, and may even be understood as the purest and most coherent form of it, implicitly definitive of the concepts of belief, desire, and so forth. Davidson (1980) holds something like this view. If so, that would make the analogy better from my point of view. For if rational actor theory is ideal typical, and PAP is defined by or (in its strictest form) identical to rational actor theory, then PAP will be on all fours with rational actor theory, i.e., ideal typical.

10.3. *What's right about functionalism*

Another analogy may help explain why, if PAP is ideal typical, people approximate it as well as they do; why, that is, PAP should be regarded as ideal typical and not as black-box instrumental. This analogy will help us see what's right about functionalism as a theory of mind. Functionalists like to compare mental states to the states of artificial computational devices such as adding machines by way of clarifying the notion of functional role. The point is that there is a sense in which it doesn't matter what a thinker is made of, any more than it matters what an adder is made of; so long as they respectively think and add. And there are many ways to build something which might add, and perhaps not a few to build something that might think. What makes an internal state of a person the thought that p (rather than that q) is its functional role, not is material realization, just as what makes the internal state of an adding machine an instance of adding $2 + 2$ is *its* functional role, not its material realization.

Many functionalists suppose the analogy to show that mental states are no more reducible to the physical states which instantiate them than arithmetic is to adding machine states. In Schwartz (forthcoming), I argue that this is a mistake. Here I appropriate the analogy for another purpose. On one view, adding machines instantiate arithmetical states, i.e., functions such as addition performed on integers. On the other, adding machines instantiate computational or machine states which mimic arithmetical states but are not identical with them. If so, they instantiate only a machine model of a bounded arithmetic. One might prefer the latter view because adding machines make mistakes, while true arithmetic functions are never wrong, or because the computational ability of any adding machine is limited by the size of its memory, and no such restriction if found in real arithmetic. Assume the second account for the sake of argument.

We might construe the relation between propositional attitudes and NIP as being analogous to that between arithmetical states and the computational states

the adding machine actually instantiates. The agent's internal states might be regarded as the computational or machine states which mimic, to some degree of approximation, the functions and states of PAP. The actual way we represent things would operate on states which acted in some way like the propositions and attitudes one ascribes in PAP, but one would no more literally ascribe the latter than one would say that an adding machine literally instantiates arithmetic. Just as we can use arithmetic to explain the behavior of adding machines, we can use PAP to explain the behavior of agents who do not instantiate it.

Alternatively, we might take PAP to be a *mimicry* of NIP. There the latter would be like arithmetic and the former like the machine or computational states. The point would not be that people actually instantiate PAP and only approximate some other sort of representation. The analogy here is rather the relation of mimicry rather than that of instantiation. Just as the adding machine performs certain nonarithmetical processes which are structurally sufficiently similar to a bounded arithmetic to enable us to calculate with them, so propositional attitude interpretation ascribes propositional attitudes which are structurally similar enough to NIP states to enable us to predict and explain human behavior in those terms.

If people actually instantiate NIP and not PAP, the structures of the two will occasionally diverge, and in some cases the analogies will entirely break down. For many purposes, though, PAP will serve as a powerful "computer" for predicting and explaining action. In addition, examination of the similarities may be a guide to research, just as the structure of arithmetic could be a guide to investigating how actual adding machines work. Speaking, as I have, of the syntactic and semantic structures of nonpropositional representation involves an appeal to this sort of analogy: the notions of syntax and semantics are derived from the study of natural languages, applied by analogy to propositions, and thence to psychological processes which may have features which are in some way like those of a natural language's syntax and semantics. On my view, the semantics is more analogous than the syntax.

This is less an a priori account of the nature of thought than an empirical research project. Unlike arithmetic and machine or computational states, we have no generally accepted or fully articulated story about NIP, which is the name of a problem rather than a theory. All we have is some indication that PAP may not be descriptively adequate, and the outlines of a better account of NIP. What's right about functionalism, then, is that functional explanation, construed as ideal typical abstraction from physical realization, is a legitimate form of explanation and defines a legitimate subject for inquiry.

10.4. *What's in the head*

My remarks about "the actual representations in the head" may raise the eyebrows of those (like myself) impressed by the arguments for externalist theories of content. Tyler Burge (1979) and Hilary Putnam (1975) develop examples which may be read as suggesting that ordinary psychological interpretation involves something like overriding what, if anything, may be "in an agent's head" and ascribing contents dictated by certain social conventions (Burge) or by the physical environmental (Putnam). I am not here arguing for externalism about content, although I accept it; I hope to be allowed to peel their examples out of context to make my point.[13]

Putnam professes not to be able to tell the difference between beeches and elms, deferring in these matters to experts in botany by the convention of the linguistic division of labor. Yet if he makes a statement using these terms, e.g., "Those beeches are lovely," we ascribe to him beliefs about beeches, even though what's in his head is some representation of a grove of trees of aesthetically pleasing but indeterminate nature. Burge similarly suggests that we ascribe to people with only incomplete knowledge or even deviant understanding of the terms they use the contents socially accepted as representing the complete, standard, and correct uses of the terms. Suppose someone wrongly thinks that "arthritis" is an inflammation not only of the joints but also of the muscles, and sometimes says things like, "That arthritis in my thighs is flaring up again." We are inclined to say that he has beliefs about arthritis and not something else, but which are false as we use the term.[14]

One way (not Putnam's or Burge's) to interpret this set of cases is as illustrations of the ideal typical character of ascription of propositional content. We might take it to show that an attribution of the belief that *p* expresses our conviction that the agent has a mental state which is close enough in content to the naturally or socially determined content of *p*. *S* believes *something or other* which for our purposes can be described as the

belief that *p*. We do not much care, for the purposes of interpretation in PAP, about the states that are actually in the head or the precision and accuracy of an agent's actual mental contents. We give them the benefit of the doubt even when (as with Putnam's ignorance of botany) it is manifestly undeserved, ascribing to them contents which they do not instantiate, but which simplify the pragmatic purposes of explaining, predicting, influencing, and evaluating behavior.

However, although we may for many purposes use PAP ideal typically, we nonetheless think that there is, after all, some representation or other, of perhaps a vague, indeterminate, or even erroneous character which enters into the explanation of the agent's actions, including linguistic behavior such as praising the beeches or complaining about one's arthritis. It is just that for many purposes we do not care what these representations are, or insofar as we do, we only care that they are not standard, correct, or exact. Thus we may just say to Putnam, "Those are elms," without bothering to try and sort out what he literally had in mind. Or we may just agree that the scenery is breathtaking. It is only in special cases — if we are engaged in cognitive psychology, say, or in the explanation of why someone acted strangely (say, taking arthritis medicine for a muscle ache) — that we bother with trying to determine the character of what's *really* in the head. To deny that there is in fact something or another in the head, fuzzy or confused as it may be, and even partially determined by social convention or physical reality, is just radical eliminativism, and Putnam and Burge reject *that* quite as decisively as I do. If my reading of the Putnam and Burge examples is right my proposal may not eliminate anything we do not already in fact treat ideal typically.

This account explains a great deal of the indeterminacy in propositional attitude ascriptions. For different (even strictly incompatible) ascriptions may for different purposes equally well approximate the actual representations NIP would posit if we had it. (This is not to say that NIP would posit fully determinate contents either). And yet the account allows us to retain a robust notion of internal representation.

The claim that we characterize content ideal typically in terms of propositional attitudes might be turned into an objection to my proposal. Perhaps what we do (it might be said) is to characterize *contents* in terms of attitudes towards propositions, where this is not a way of talking about internal states *at all*. Whatever form representation takes in the head, it involves taking attitudes towards contents which may be described in quasi-linguistic terms as having a propositional structure. No suggestion need be involved that such contents are inscribed in an agent's brain, or that any contents be so inscribed. Brian Loar (1981) advocates a theory of this sort, in which nonpropositional internal states are "indexed" to external sentences in a natural language. Now insofar as the claim is that contents *may be described* this way, I have no quarrel with it. Of course they can: that is why PAP is a plausible (if ideal typically) psychology. But insofar as this is meant as a way of construing *psychology,* I find the objection obtuse. What psychology does is to ascribe internal states with certain causal powers to produce behavior, and here we must ask if the ascription is correct. My claim is that with PAP, the answer is "only approximately."

11. Conclusion

Propositional attitudes (understood quasi-linguistically), then, are not essential to intentional psychology. We can eliminate such attitudes, denying that human beings think in a languagelike language of thought, without thereby denying that they do think. This is fortunate, as the empirical evidence is strong that our thinking diverges from the Propositionalist model on the syntactic and inferential dimensions. Nonetheless we need not lose the semantic properties which (if I am right) are what makes thinking thinking in the first place. Even an imagistic psychology could be semantically rich, although it is far more likely that we instantiate a mental models psychology — or perhaps, though I have not discussed this, a connectionist one. My idea is that we go irrealist about propositional attitudes, construed as languagelike internal states, but help ourselves to PAP for most purposes in the same way that we go irrealist about classical mass and help ourselves to CM.

There is more payoff in this than just a better, more empirically adequate psychology (as if that were not enough). We understand more clearly what we care about in the intentional, where that is a matter of representative (not propositional) character. I do not claim that the Mental is the Intentional, because that sort of proposal doesn't seem productive on the basis of the history of similar claims. My claim is rather that intentionality is what we are least willing to revise or

eliminate in our strong commitment to the idea that people think and act on what they think. We also understand more clearly what sort of theory PAP is, namely, an ideal typical one. Functionalist psychology, properly construed, will be ideal type explanation of human behavior in mentalistic terms. With this comes a host of problems about the nature of ideal type explanation, but if we take the route I suggest, at least we will be debating the right issues and not making a priori claims about the Nature of Mind in general or the Irreducibility of the Mental.

Notes

[1] Thanks are due for comments, criticism, and ideas to Kathy Akins, Anne Bezuidenhuit, John Bickle, Paul Boghossian, Allan Gibbard, David Hills, Jaegwon Kim, Calvin Normore, Peter Railton, Kendall Walton, and Mark Wilson.

[2] See Rumelhart and McLelland, *et. al.* (1986), and a burgeoning literature summarized in part in Churchland (1986) and Churchland (1989).

[3] I find these objections weak, although the usual sorts of replies — urged, for example, by Sterenly (1990, pp. 160—197), seem inconclusive. I avoid this debate because the merits of Propositionalism can be assessed only by comparing it to a reasonably well-articulated rival story, which I here try to present.

[4] See Wittgenstein (1958, p. 128e (§ 432)): "Every sign *by itself* seems dead. *What* gives it life? — In use it is *alive*." Wittgenstein's answer seems to be that "the *use* [is] its life." Other answers are possible.

[5] In Churchland and Churchland (1983), these writers, citing Stich, subscribe to the formality constraint discussed below and draw radical eliminativist conclusions from it. While in my view a correct inference, this seems inconsistent with other things they say, such as those quoted here.

[6] Below I sketch an argument against this constraint, but a full-fledged discussion (which I hope to provide another time) is impossible here.

[7] Aside from this sensible observation, Dennett, of all people, offers an a priori argument that there cannot be mental images (Dennett, 1969, pp. 136—87); as Fodor remarks, this is rather embarrassing (for Dennett), since there is good empirical evidence for mental imagery (Fodor, 1981b, p. 76).

[8] The classic work was done in the 1970s by Roger Shepard and his colleagues (Shepard and Cooper, 1982). Kosslyn and his colleagues have extended these results to a wide variety of cases. See Kosslyn (1980); Kosslyn *et. al.* (1981); Kosslyn (1981, pp. 207—244).

[9] It seems a fair criticism of the sort of Humean impressionism to which Kosslyn sometimes appears to adhere as an account of content. Fodor seems to offer it as a criticism of the possibility of nonpropositional representation in general.

[10] The final clause need not smuggle back in reference to propositions if perceiving or conceiving states of affairs is a matter of having mental models of them.

[11] Kahneman and Tversky, who do parallel work on inductive reasoning to Johnson-Laird's on deductive reasoning, also like mental models. See Kahneman and Tversky (1982, pp. 201—210).

[12] Mark Wilson warns me that the mathematical relation of CM and RM is rather more complex than I suggest here. However, I think as a sketch this crude statement will do, although a detailed study might be revealing in understanding precisely how the idealization works.

[13] This idea was suggested to me by Kathy Akins. She assures me that what I got out of her ideas was not at all what she meant. If I am right, however, we can override her denial.

[14] The intuitions invoked are delicate. For example, if Putnam in his ignorance praises the beeches while pointing at elms, we might ascribe him a false belief about beeches, or a true belief about elms, while taking it that "beech" means "elm" in his vocabulary. Others might read the Burge case as involving not a false belief about "arthritis" but a true belief whose content is something like "My thigh aches again." These interpretations equally make my point.

References

Anderson, John R.: 1978, 'Arguments concerning representations for mental imagery', *Psychological Reviews* **85**, 249—77.

Baylor, G. W.: 1971, 'Programs and protocol analysis on a mental imagery task', *First Joint International Conference on Artificial Intelligence.* n.p.

Burge, Tyler: 1979, 'Individualism and the mental', in *Midwest Studies in Philosophy* IV, *Studies in Metaphysics,* ed. Peter A. French, Theodor E. Uehling, Jr., and Howard K. Wettstein, University of Minnesota Press, Minneapolis, pp. 73—122.

Churchland, Patricia S.: 1980, 'A perspective on mind-brain research', *Journal of Philosophy* **77**, 185—207.

Churchland, Patricia S.: 1986, *Neurophilosophy,* MIT Press, Cambridge, Mass.

Churchland, Patricia S.: 1987, 'Epistemology in the age of neuroscience', *Journal of Philosophy* **84**, 544—552.

Churchland, Paul M.: 1989, *A Neurocomputational Perspective,* MIT Press, Cambridge, Mass.

Churchland, Paul M. and Churchland, Patricia S.: 1983, 'Stalking the wild epistemic engine', *Noûs* **17**, 5—18.

Davidson, Donald: 1980, *Essays on Actions and Events,* Clarendon Press, Oxford.

Dennett, Daniel: 1969, *Content and Consciousness,* Routledge and Kegan Paul, London.

Dennett, Daniel: 1978, *Brainstorms,* Bradford Books, Cambridge, Mass.

Dennett, Daniel: 1987, *The Intentional Stance,* MIT Press, Cambridge, Mass.

Farah, M. J.: 1984, 'The neurological basis of mental imagery: A computational analysis', *Cognition* **18**, 243—61.

Field, Hartry: 1973, 'Theory change and the indeterminacy of reference', *Journal of Philosophy* **70**, 462—81.

Field, Hartry: 1981, 'Mental representation', in *Readings in Philosophy of Psychology,* Vol. 2, ed. Ned Block, Harvard University Press, Cambridge, Mass., pp. 78—114.

Fodor, Jerry A.: 1975, *The Language of Thought,* Thomas Crowell, New York.

Fodor, Jerry A.: 1981a, *Representations: Philosophical Essays on the Foundations of Cognitive Science,* MIT Press, Cambridge, Mass.

Fodor, Jerry A.: 1981b, 'Imagistic representation', in *Imagery,* ed. Ned Block, MIT Press, Cambridge Mass., pp. 63—86.

Fodor, Jerry A.: 1987, *Psychosemantics,* MIT Press, Cambridge, Mass.

Forbus, K. D.: 1983, 'Qualitative reasoning about space and motion', in *Mental Models,* ed. D. Genter and A. L. Stevens, Lawrence Erlbaum, Hillsdale, N.J., pp. 53—74.

Friedman, Milton: 1979, 'The methodology of positive economics', in *Philosophy and Economic Theory,* ed. Frank Hahn and Martin Hollis, Oxford University Press, Oxford, pp. 18—35.

Gardner, Howard: 1985, *The Mind's New Science: A History of the Cognitive Revolution,* Basic Books, New York.

Goodman, Nelson: 1972, *Problems and Projects,* Hackett, Indianapolis.

Grandy, Richard C.: 1973, 'Reference, meaning, and belief', *Journal of Philosophy* **70**.

Harman, Gilbert: 1973, *Thought,* Princeton University Press, Princeton.

Haugeland, John: 1981, 'The nature and plausibility of cognitivism', in *Mind Design,* ed. John Haugeland, MIT Press, Cambridge, Mass., pp. 243—81.

Hills, David: 1981, 'Mental representation and languages of thought', in *Readings in Philosophy of Psychology,* Vol. 2, ed. Ned Block, Harvard University Press, Cambridge, Mass., pp. 11—20.

Johnson-Laird, Philip N.: 1983, *Mental Models,* Cambridge University Press, Cambridge.

Johnson-Laird, Philip N.: 1988, *The Computer and the Mind: An Introduction to Cognitive Science,* Harvard University Press, Cambridge, Mass.

Johnson-Laird, P. N. and Wason, P. C. (ed.): 1977, *Thinking: Readings in Cognitive Science,* Cambridge University Press, Cambridge.

Kitcher, Patricia: 1984, 'In defense of intentional psychology', *Journal of Philosophy* **81**, 89—106.

Kosslyn, Stephen Michael: 1980, *Image and Mind,* Harvard University Press, Cambridge, Mass.

Loar, Brian: 1981, *Mind and Meaning,* Cambridge University Press, Cambridge, Mass.

Neisser, Ulrich: 1967, *Cognitive Psychology,* Appleton Century Croft, New York.

Nisbett, Richard and Ross, Lee: 1980, *Human Inference,* Prentice-Hall, Englewood Cliffs.

Putnam, Hilary: 1975, *Mind, Language, and Reality: Philosophical Papers,* Vol. 2, Cambridge University Press, Cambridge.

Pylyshyn, Z. W.: 1973, 'What the mind's eye tells the brain: A critique of mental imagery', *Psychological Bulletin* **80**, 1—24.

Pylyshyn, Z. W.: 1981, 'The imagery debate: Analog media versus tacit knowledge', in *Imagery,* ed. Ned Block, MIT Press, Cambridge, Mass., pp. 151—206.

Quine, W. V. O.: 1960, *Word and Object,* MIT Press, Cambridge, Mass.

Rips, Lance: 1986, 'Mental muddles', in *The Representation of Knowledge and Belief,* ed. Myles Brand and Robert M. Hamish, University of Arizona Press, Tucson, pp. 258—286.

Schwartz, Justin: forthcoming, 'Who's afraid of multiple realizability?', in *Closing the Gap: Symbolic vs Subsymbolic Processing,* ed. John Dinsmore, Lawrence Erlbaum, Hillsdale, N.J.

Shepard, R. N. and Metzler, J.: 1977, 'Mental rotation of three-dimensional objects', in Johnson-Laird and Wason, 1977, pp. 532—35.

Sterelny, Kim: 1990, *The Representational Theory of Mind,* Basil Blackwell, Oxford.

Stich, Stephen P.: 1983, *From Folk Psychology to Cognitive Science: The Case Against Belief,* MIT Press, Cambridge, Mass.

Wittgenstein, Ludwig: 1958, *Philosophical Investigations,* 3rd ed., transl. by G. E. M. Anscombe, MacMillan, New York.

Department of Philosophy
Ohio State University