PACIFIC PHILOSOPHICAL Quarterly

Published by
the School of Philosophy
University of Southern California

April 1986 **Vol. 67, No. 2**

PACIFIC PHILOSOPHICAL QUARTERLY

PUBLISHED BY: The University of Southern California

EDITORIAL BOARD: Hartry Field, Barbara Herman, Brian Loar, Miles Morgan

MANAGING EDITOR: Raymond R. Neal

ADVISORY BOARD: Ernest W. Adams, *University of California (Berkeley);* Virgil C. Aldrich, *University of Utah;* Robert Audi, *University of Nebraska (Lincoln);* Theodore M. Benditt, *University of Alabama (Birmingham);* Gordon Brittan, *Montana State University (Bozeman);* Tyler Burge, *University of California (Los Angeles);* Nancy Cartwright, *Stanford University;* Roderick M. Chisholm, *Brown University; Wilhelm K. Essler, Johann Wolfgang Goethe-Universität (Frankfurt am Main, West Germany);* Philippa Foot, *Oxford University and University of California (Los Angeles);* H. P. Grice, *University of California (Berkeley);* Marsha P. Hanen, *University of Calgary;* David Harrah, *University of California (Riverside);* Karel Lambert, *University of California (Irvine);* Charles Morgan, *University of Victoria;* John Perry, *Stanford University;* Robert J. Richman, *University of Washington;* Wesley Salmon, *University of Pittsburgh;* Stephen Schiffer, *University of Arizona;* Krister Segerberg, *University of Auckland;* Zeno Vendler, *University of California (San Diego);* John M. Vickers, *Claremont Graduate School;* Richard Wasserstrom, *University of California (Santa Cruz).*

SUBSCRIPTION INFORMATION: Annual rates—Institutions, $30.00; Individuals, $16.00; Single issues: $7.50. Published quarterly; January, April, July and October. Address all subscription correspondence to: Managing Editor, *Pacific Philosophical Quarterly,* School of Philosophy, University of Southern California, Los Angeles, CA 90089-0451, U.S.A.

This journal is registered at the Copyright Clearance Center, Inc., P.O. Box 765, Schenectady, NY 12301, U.S.A.

The code on the first page of an article indicates the copyright owner's consent that copies of the article may be made for personal or internal use, or for the personal or internal use of specific clients. This consent is given on the condition that the copier pay through the Copyright Clearance Center the per-copy fee stated in the code for copying beyond that permitted by Sections 107 or 108 of the U.S. Copyright Law. (Articles published prior to 1979 may be copied for a fee of $1 per copy, also payable through the Center.) This consent does not extend to other kinds of copying, such as for general distribution, for advertising or promotional purposes, for creating new collective works, or for resale.

Pacific Philosophical Quarterly (ISSN 0031-5621); is published quarterly by the School of Philosophy of the University of Southern California. Second-class postage paid at Los Angeles, California, and at additional mailing offices. *Postmaster:* send address changes to Pacific Philosophical Quarterly, School of Philosophy, University of Southern California, 90089-0451, U.S.A.

Copyright © 1986 by University of Southern California ISSN 0031-5621

PACIFIC PHILOSOPHICAL Quarterly

Vol. 67, No. 2
April 1986

STALNAKER'S PROBLEM OF INTENTIONALITY*

BY

Stephen Schiffer

THOUGHTS have *content*, and so do words. A particular internal state may be a belief *that worms do not have noses*, and a particular sequence of marks or sounds may be a sentence which means *that worms do not have noses*. According to a certain familiar and widely held position, having content is a matter of being appropriately related to a content, and contents are propositions; that is, abstract, objective, language-independent entities that have essentially the truth conditions they have. Those relations to propositions that lead to the having of content are called *intentional relations*, and there are two kinds of them: mentalistic relations, most notably believing, desiring and intending, and semantic relations, most notably meaning; and it is further held by the familiar doctrine which I am describing that the semantic intentional relations reduce to the mentalistic intentional relations. As regards the reduction of the semantic to the psychological, the picture, somewhat simplified, is well conveyed by Stalnaker:

> One might begin an explanation of the nature of the semantic relation between a sentence S and a proposition x by explaining what it is for a *speaker* to mean x by uttering S. Then one might go on to explain what it is for a *sentence* S to mean x in terms of the existence of a convention in a population of speakers that S should be used to mean x. Conventions are to be explained in terms of the pattern of beliefs and intentions of the members of the population. If one had an adequate account of this kind, then one would have a reduction of a principal semantic relation to the mentalistic intentional relations, belief and intention.[1]

This comprehensive position on the nature of intentionality faces some outstanding questions:

(1) What more can be said about the nature of propositions? Are they

*This and the following two papers were given at a symposium on Robert Stalnaker's *Inquiry*, at the APA Central Division meetings, St. Louis, May 1986.

Pacific Philosophical Quarterly 67 (1986) 87-97 0031-5621/86/0200-0087-$01.10
Copyright © 1986 by University of Southern California

very fine-grained things, say, entities with as much structure and as many components as the sentences which express them, or very course-grained things, say, sets of possible worlds, or perhaps something in between?

(2) How *exactly* are semantic intentional relations to be reduced to mentalistic intentional relations? How exactly are the notions of *speaker-meaning* and *convention* to be explicated, and what sorts of conventions are needed to explain the meaning of a sentence that has never been, and may never be, uttered?

(3) Let it be granted that all questions about *linguistic* representation reduce to questions about *mental* representation in the way the theorist supposes. How, then, is the theorist to account for mental representation? That is, what is the nature of the various mentalistic intentional relations—believing and the rest—that relate people or their inner states to propositions? I take this to be the really big question; the most pressing question facing the program that I have outlined. It is what Stalnaker calls *the problem of intentionality,* and I want to dwell a little on it.

Let us suppose with the theorist in question that believing is a relation to propositions, that, if you will, 'y' in the schema 'x believes y' is a genuine objectual variable the values of which are propositions. As regards the belief relation one can either hold with Brentano that the relation is irreducibly mentalistic, not explicable in any other terms; or one can hold the opposite, that believing is a relation that is intrinsically specifiable in naturalistic, or at least topic neutral, terms. The theorist I've been describing will not want to side with Brentano; he hasn't cashed in semantic facts for mentalistic facts only to say that that's the end of the line. Human beings, the theorist holds, are part of the natural order, and so are their mental states. Thus, the problem of intentionality for the program I have described is precisely the problem of giving a naturalistically acceptable reduction of the belief relation, and of all other mentalistic intentional relations. "The challenge presented to the philosopher who wants to regard human beings and mental phenomena as part of the natural order is," Stalnaker recognizes, "to explain intentional relations in naturalistic terms" (p. 6). But how is this to be done? To answer this is to solve the problem of intentionality. It is a particularly pressing problem for the program I've outlined, because it is arguable that, without such a solution, there can be no plausibility to the idea that believing *is* a relation to propositions.

II

Robert Stalnaker adheres to the described research program, and thus inherits its questions. He doesn't offer any answer to the second of the three outstanding questions—the one about the reduction of the semantic to the psychological—but in his book *Inquiry* he does offer an answer to the first question and a strategy for answering the third question; that is to say, a strategy for solving the problem of intentionality.

Stalnaker's answer to the first question, the one about the nature of propositions, is simple, but hairy. A proposition is any function from any non-empty domain of possible worlds into truth-values. In other words, a proposition is any non-empty set of possible worlds. (He also has interesting things to say about the metaphysical status of possible worlds, but they won't concern us.) The position is patently simple; what makes it hairy, of course, is that it entails that necessarily equivalent propositions are identical, and that therefore one believes all necessary propositions if one believes any, there being only one to begin with. Not surprisingly, Stalnaker thinks that his view of propositions enjoys "a compelling philosophical motivation" (p. ix), and that the obvious problems that it faces can be explained away.

Partly because of the time constraints imposed by the symposium,[2] I've decided to concentrate my remarks on Stalnaker's strategy for solving the problem of intentionality; but I can't resist getting in a quick two cents about the two matters just touched on.

First, a comment about the philosophical motivation for the possible worlds framework. This is supposed to be given by what Stalnaker calls "the pragmatic picture," and which he sketches thus:

> Rational creatures are essentially agents. Representational mental states should be understood primarily in terms of the role that they play in the characterization and explanation of action. What is essential to rational action is that the agent be confronted, or conceive of himself as confronted, with a range of alternative possible outcomes of some alternative possible actions. The agent has attitudes, pro and con, toward the different possible outcomes, and beliefs about the contribution which the alternative actions would make to determining the outcome. One explains why an agent tends to act in the way he does in terms of such beliefs and attitudes. And, according to this picture, our conceptions of belief and of attitudes pro and con are conceptions of states which explain why a rational agent does what he does. (p. 4.)

What puzzles me is how this pragmatic picture can be taken to justify Stalnaker's conception of propositions when, as far as I can tell, it is a picture of rational activity that is consistent with virtually any philosophical theory of propositional attitudes that I can think of. (It should be noted that Stalnaker also speaks of his conception of propositions as being motivated by the "pragmatic-causal picture" (e.g. p. 81), where that, I take it, is the conjunction of the pragmatic picture just glossed and the Wisconsin-style causal analysis of mental representation that we'll come to shortly. But that causal theory can provide no independent motivation for Stalnaker's conception of propositions; for, as will be evident when we get to it, the causal theory entails that a person who believes anything also believes every necessary proposition, and it's pretty hard to see how anyone who hadn't

already accepted the possible worlds framework could find this acceptable.)

My other quick comment concerns the methodology for fending off the *prima facie* counterexamples to the possible worlds construal of propositions. Consider these two utterances:

[i] Ralph believes that $1^2 + 1^2 = 4$.

[ii] I believe that I'm a paragon, but that that guy [pointing to what I fail to realize is a photo of myself] is a moral degenerate.

Intuitively, one wants to say that these may both be true, and that each ascribes belief in an impossible proposition. Now Stalnaker can't accept this conjunction of claims about [i] and [ii],[3] and his strategy with such cases is to claim that, despite first appearances, the 'that'-clauses don't refer to impossible propositions. In the case of [i] Stalnaker favors a metalinguistic maneuver, and I'm not sure what he'd say about [ii]. But for each such *prima facie* counterexample he finds an unproblematic, if unexpected, proposition to be the real referent of the 'that'-clause. My niggling comment here is just that I'd feel a lot more comfortable about Stalnaker's ability to fend off counterexamples if his solution in each case were generated by a general semantic theory of 'that'-clauses which *we* could use to determine the proposition referred to in any given propositional attitude ascription. Such a theory would show how the referent of any give 'that'-clause was jointly determined by (a) semantic values assigned to its parts by a compositional semantics for the language and (b) antecedently specified contextual factors. I might add that I don't see why this should be such a difficult thing to do, if the possible worlds framework is correct.

III

I turn now to Stalnaker's answer to the third question, to his account of the belief relation, and thus to his solution to the problem of intentionality. For the problem of intentionality is, he tells us, "the central philosophical problem that an account of mental representation must solve" (p. 7).

To solve the problem of intentionality relative to the possible worlds framework requires that one give a materialistically adequate account of the belief relation, where that is taken to be a relation between a person and a non-empty set of possible worlds. Actually, Stalnaker doesn't offer a solution to the problem of intentionality; what he offers is "only a strategy—only the bare outlines of an account of intentional mental states" (p. 19). The question I shall be asking is whether this strategy is at all promising.

Stalnaker's strategy is a strategy for explicating the belief relation. Although he doesn't put it in just these terms, that strategy begins by dividing the task of defining 'x believes p' into two subtasks.

The first subtask is to give a naturalistic account of *belief states*; that is, to say in naturalistically acceptable terms what it is for an internal state of a person to be a belief.

The second subtask is to define a relation R, again in naturalistic terms, so that the final definition of 'x believes p' will take this form:

x believes p iff x is in some belief state s such that R(x, s, p).

Stalnaker says virtually nothing about how the first subtask is to be accomplished. He does say that "beliefs are *beliefs* rather than some other representational state, because of their connection, through desire, with action" (p. 19), and this suggests that the notion of a belief state is somehow to be functionally defined in tandem with the notion of a desire state; but we are given no indication of how to go about constructing such a definition, or of what the functional criterion for being a belief would be. At the same time, Stalnaker does advance a general theory of rationality that would greatly constrain the needed account of belief states. For, if I understand him correctly, Stalnaker maintains that an ideally rational believer would be in only one belief state. The content of this single belief state would be given by a set of possible worlds—intuitively, the set of possible worlds not ruled out by the way the agent conceives the world to be. Particular propositions believed by the ideally rational believer would then be properties of the belief state. This agent believes a proposition just in case it is true in every possible world in the set of possible worlds that represents his single belief state. Now in order to see how ordinary mortals can profit from deductive inquiry, we must see the ideally rational believer as an ideal of rationality that probably never obtains. In practice, believers will have a large number of concurrent but separate belief states. These separate belief states, Stalnaker says, are like "separate centers of rationality" (p. 87); each such belief state may be represented by a set of possible worlds, and a person will believe any given proposition p just in case p is true in every possible world in *some* belief state that the person is in. It follows that a person will believe every proposition entailed by every proposition he believes, but because he may be in separate belief states, it doesn't follow that he believes any proposition which is a deductive consequence of a set of propositions, each one of which he believes. It's this last fact, incidentally, that Stalnaker uses to solve what he calls "the problem of deduction," which is basically the problem of saying how, on his theory, one can fail to believe what's entailed by what one believes.

I don't really see how the fragmentation of belief states solves the problem of deduction, but I guess I'd better not go off on that issue.[4] What I want to comment on is the way in which Stalnaker's conception of rationality and propositions believed as properties of a belief state constrains what he is free to say about what makes an internal state a belief. For I

think that the nature of the constraint makes achieving the first subtask seem a lot harder than it would seem without the constraint. Without the constraint, Stalnaker must merely find some functional role such that a state is a belief if it has that functional role, and many people are sympathetic to the idea that a state is a belief by virtue of its functional role. With the constraint, however, Stalnaker must make sure that the functional role which makes a state a belief is one that in the ideally rational agent is possessed by just a *single one* of that agent's internal states. There is just one internal state of that agent which is a belief. The content of that state is given by a set of possible worlds, and the many propositions believed by the agent are just those that are true in each of the possible worlds in the set of possible worlds that represents his single belief state. It's true that in us, who are less than ideally rational, the functional role that makes a state a belief will be had by several distinct states; but even here it must follow from the nature of the functional role that the distinct belief states which have it are in the appropriate sense "separate centers of rationality," and thus have whatever properties are required to solve Stalnaker's problem of deduction. Needless to say, all this is very different from the standard view, according to which there are, roughly speaking, as many distinct internal states of belief as there are propositions believed. One would like to know more about the strategy that Stalnaker has in mind for achieving the first subtask; that is, for explicating what it is for a state to be a belief.

IV

I turn now to the second subtask in Stalnaker's strategy for solving the problem of intentionality. This is the task of specifying a materialistically adequate relation R such that:

x believes p iff x is in some belief state s such that R(x, s, p).

Now Stalnaker does have a proposal for how this subtask should be pursued, and it exploits what I'll call *the fuel gauge model of representation*, myself exploiting an example and mode of explanation used elsewhere by Fred Dretske.[5] Suppose I believe that it's raining. Then I'm in a belief state that *represents* that it's raining. So we're really looking for a theory of representation of which *mental* representation is a special case. And representation, we know, is not just a feature of mental states. After all, the position of a needle on the dashboard of your car can represent your gas tank as being 3/4 full. How is it that the position of the needle can have this representational power? Surely, because of the capacity of the fuel gauge to be a reliable indicator of the amount of fuel in the gas tank. More exactly, the position of the needle represents the gas tank as being 3/4 full

because under optimal conditions the needle is where it is—viz. pointing at '3/4'—*only when* and *because* the gas tank is 3/4 full of gasoline. This may be taken to suggest that representation *in general* derives from the capacity of a system to be a reliable indicator of its context, and that what accounts for the fuel gauge's representational features also accounts for those of a person's mental representations.

Stalnaker's development of this idea, which is directly indebted to work of Dennis Stampe's,[6] may be restated, in words very close to his own, thus:

[A] x believes p iff x is in a belief state that, under optimal conditions, x is in only if p is true, and under optimal conditions, x is in that belief state because p is true, or because some proposition which entails p is true.

(It is understood that 'p' ranges over propositions as defined by Stalnaker.)

The first thing we should notice is that [A] implies that, under optimal conditions, everyone is infallible: when conditions are optimal, no one can have any false beliefs. One wonders what these "optimality conditions" are; whatever they are, it's clear that they never obtain. But the most important point to be made about the optimality conditions is that *Stalnaker can make no legitimate appeal to them unless they are specifiable in wholly non-intentional terms*. If these optimal conditions can only be specified in terms that presuppose intentionality, then, while [A] may or may not be true, it can be no part of a strategy for solving the problem of intentionality. In view of this, the real suggestion encapsulated in [A] that bears on the problem of intentionality may be restated thus:

[B] There is some ("optimality") condition C which perhaps never obtains (though it could obtain) and which is specifiable in wholly non-intentional and materialistically adequate terms, such that:
x believes p iff x is in some belief state s such that: were C to obtain, then x would be in s only if p were true, and x would be in s because p was true, or because some proposition which entailed p was true.

This is better than [A] because it makes it clear that we're not being offered an explication of the belief relation that would solve the problem of intentionality, but are rather being offered a proposal as to what form such an explication must take. (Stalnaker did not think that in suggesting [A] he was offering a solution to the problem of intentionality; he was well aware of the need to cash in the notion of optimal conditions for notions that clearly begged no questions about intentionality, which is why he was careful not to claim to have offered any explication of the belief relation, but claimed only to have offered a *strategy* for explicating the belief relation.)

But why on earth think that [B] is true? Why think that a reasonable strategy for explicating belief is to start with [A] and then try to replace the reference to "optimal conditions" with a specification of suitably non-

intentional conditions? How could anyone have confidence that there was a condition of the sort [B] required without having some idea of what that condition was?

I can think of two things that Stalnaker might say in response to the challenge to say why [B] represents a promising strategy for solving the problem of intentionality.

First, he may say that he has such confidence in the fuel gauge model of representation, and in the Stampean approach to spelling out that model, that he's justified in thinking that something along the lines of [B] is correct. Second, he may feel that [A] seems close to being right when the notion of optimal conditions is taken at face value, and that because the righthand side of [A] is to that extent promising as an explication of the belief relation, we should expect the notion of optimal conditions to be cashable out in naturalistically adequate terms that don't themselves presuppose the intentionality being explicated.

In response to the first thing that Stalnaker might say, I should say, in the first place, that we aren't now able to give any clear account in naturalistic terms even of such simple representational systems as fuel gauges, and that we mustn't overlook the fact that it may be crucial to explaining how fuel gauges represent levels of gasoline in gas tanks that these devices were *designed* by human beings to perform the representational function in question. And in the second place I should say that the central importance of our own representational mental states makes it ludicrous to suppose that any approach to the nature of representation in general could be thought promising independently of its application to us. No one should deny that the representational features of a mental state have essentially *something* to do with the role of that state in guiding behavior, and with the causal relations that run from the head containing the state to the larger world; but such platitudes can hardly sustain the very strong philosophical claim that the representational features of mental states can be *defined* in causal terms, let alone defined in the Stampean terms proposed by Stalnaker.

In response to the second thing that Stalnaker might say, I should deny that [A] has any plausibility when considered just on its own, apart from any attempt to say what is meant by 'optimal conditions.' On the pretheoretic understanding of "optimal conditions" that we would naturally bring to a first reading of [A], we would certainly suppose that a person could have a false belief even under the best of conditions. In what sense must we conclude that conditions were less than optimal for Newton when he formed the false but then well-justified belief that his mechanical theory was true? Intuitively, then, [A] fails to state a necessary condition. There are other reasons for thinking that [A] fails to state a necessary condition. Suppose I'm looking right at Fido, and form the belief that he is standing before me. The belief state that I'm in is one that the presence of Fido caused me to be in. But it seems clear that that belief state would have been caused

to occur in me if some qualitative duplicate of Fido had been before my eyes instead of Fido. So it seems false that a certain state of mine is a belief that Fido is standing before me only if under optimal conditions I am in that state only if Fido is standing before me: even under optimal conditions I would be in that belief state if any creature sufficiently resembling Fido were before me.

There are also apparent counterexamples to the claim that [A] provides a sufficient condition. A philosopher who approached [A] in innocence of the strategies of possible worlds theorists like Stalnaker or Hintikka would surely protest that it could not be right, as it entails that to believe any proposition is to believe everything entailed by that proposition, which would include, of course, all necessary propositions. This does show, as noted earlier, that [A] can't be used to motivate the possible worlds framework; but such a "counterexample" should not be pressed now, for, as regards the prospects of solving the problem of intentionality from within a possible worlds framework, the more interesting question concerns the adequacy of [A] even assuming that that framework is correct. Yet even on this assumption there remain what look like counterexamples to the alleged sufficiency of [A]. Suppose I believe that I'm looking at something red, that conditions are optimal, and that the internal state that is my belief that I'm looking at something red is a state that I'm in only if and because I'm looking at something red. But putting that together with other stuff we know, we can also say that the belief state in question is one that I'm in only if and because the receptor cells in my eyes are being stimulated in such and such way. So it ought to follow that in addition to believing that I'm looking at something red I also believe that my receptor cells are being stimulated in a certain way. But I have no such belief. Here is another example that also seems to show that [A] doesn't provide a sufficient condition. Suppose I believe that Sally has the flu, and that the state which is that belief is one that, under optimal conditions, I'm in only if and because the proposition that Sally has the flu is true. Now suppose that the proposition Q provides the complete explanation of why Sally has the flu, and consequently that the proposition that Sally has the flu is true just in case Q is true. Given the connection between Q and the proposition that Sally has the flu, it would seem to follow that the state which is my belief about Sally is also one that, under optimal conditions, I would be in only if and because Q was true. Yet it seems that in believing that Sally has the fly I needn't have any beliefs about why she has the flu. If [A] provided a sufficient condition for believing a proposition, then people would believe many more things than they actually believe.

I conclude, then, that there is no obvious reason whatever for thinking that [B] is the least bit promising. I also think that there are reasons for thinking that [B] is false. Let me mention two of them.

(1) Propositions are enmeshed in a huge causal nexus. Any true empirical

proposition will have many other propositions whose truth is causally necessary or sufficient for its truth. Obviously it's possible to believe a proposition without also believing very many of these other causally connected propositions. The problem for [B] is that it's very hard to see how a candidate specification of the quantified condition C will enable us to select out of all the very many propositions that are directly or indirectly causally involved in the explanation of why one is in a given belief state just those propositions as properly belong to the belief's content. In other words, it's very hard to see how any candidate for condition C won't let in too much and thereby fail to provide a sufficient condition. As I pointed out in the flu example, this problem already infects the unregimented notion of optimal conditions as it enters into [A].

(2) [B] demands the specification in non-intentional terms of some "optimality" condition such that beliefs formed under it can only be true. This condition will be very complex, and will include a description, in non-intentional terms, of the ideal workings of our internal belief-forming mechanisms. But however well our cognitive mechanisms are functioning, that won't secure that we form only true beliefs. Quite a bit of cooperation will be required on the part of the larger world that includes us information-processing systems. So the sought after optimality condition will include very strong restrictions about the physical environment in which our belief-forming mechanisms are forced to operate. In effect what we are after is a completion of the schema,

Were x's belief-forming mechanisms to have property F and x's physical environment to have property G, then x would form only true beliefs.

where the substituends for 'F' and 'G' specify the properties non-intentionally. *Is* there such a pair of properties that satisfies this form? That seems very far from certain. But even if there were two non-intentionally specifiable properties that satisfied this schema, it seems pretty clear to me that no human could ever state them, and that the property of the environment would be such an unnatural and concocted affair, would have so little to do with how we worked as cognitive systems, that the "optimality" condition determined by these two properties would stand no chance of entering into a solution to the problem of intentionality.

V

Let me sum up the main point that I've been trying to make. Stalnaker claims that believing is a relation between a person and a proposition, this understood as any function from a non-empty domain of possible worlds into truth-values. This claim has credibility only if it is possible to give a naturalistic account of the belief relation thus construed. But Stalnaker has

not succeeded in giving any reason to suppose that such an account can be given.

Would the problem of intentionality prove more tractable if one took something other than Stalnakerian propositions to be the objects of propositional attitudes? Not really, I'm inclined to think. I think that Stalnaker's conception of propositional attitudes is very restrictive as to what can count as a solution to the problem of intentionality, and that other approaches may have greater resources available in trying to get a naturalistically adequate account of believing; but I don't think any of these approaches can be made to work, either. I think that the whole enterprise of trying to give a physicalistic reduction of mental states and intentionality is hopeless, and I don't think that the ontological or conceptual integrity of intentional phenomena depends on there being such a reduction. Nothing disastrous follows from the inability to show that belief facts, such as the fact that I believe that I own more than two pairs of socks, are identical to facts statable in sentences that are devoid of intentional and mentalistic idioms. Part of the reason that I think this is that I also think that the relational theory of propositional attitudes is false. Believing isn't a relation to anything. But this isn't the occasion for my story.

University of Arizona
Tucson, Arizona

NOTES

[1]Robert C. Stalnaker, *Inquiry* (Cambridge, Mass.: MIT Press, 1984), p. 32. All page references are to this work.

[2]See the editor's introduction.

[3]If he did accept it he would have to say that both 'that'-clauses referred to the same proposition, for on Stalnaker's theory there is only one impossible proposition. It is also a consequence of Stalnaker's account of belief states that no one can believe an impossible proposition.

[4]Field's Cantor example (p. *109* this issue) is a good example of the problem I have in mind.

[5]"Aspects of Cognitive Representation," in M. Brand and R. Harnish (eds.), *Problems in the Representation of Knowledge and Belief* (Tucson: University of Arizona Press, 1986).

[6]"Toward a Causal Theory of Linguistic Representation," in P. French, T. Wehling, Jr., and H. Wettstein (eds.), *Contemporary Perspectives in the Philosophy of Language* (Minneapolis: University of Minnesota Press, 1979).

STALNAKER ON INTENTIONALITY[1]

BY

HARTRY FIELD

STALNAKER offers us a theory of the deliberation and inquiry of intelligent agents that makes heavy use of believing and desiring, construed in a particular way. Let a *Stalnaker-proposition,* or *S-proposition,* be a function from some algebra of possible worlds (not necessarily comprising all the possible worlds) into truth values. (If we ignore the qualification that not all possible worlds need to be assigned truth values, we could just as well say that an S-proposition is simply a set of possible worlds.) Stalnaker wants to construe believing and desiring as relations between the cognitive states of the agent and S-propositions. Stalnaker thinks that the S-proposition assigned to an intentional mental state should be viewed as the *content* of that state, or the *object of* that state.

Stalnaker sees three (not entirely distinct) sorts of advantages to his view of the objects of intentional states over rivals: first, that it better fits a pragmatic picture of the purpose of postulating intentional states; second, that its account of the identity conditions for mental states has various benefits; and third, that only by conceiving of the objects of mental states in this way we can solve *the problem of intentionality,* that is, the problem of giving a naturalistic account of what gives mental states the content they have. I will be disagreeing with his views on all three points.

I should say at the outset that putting the issue in terms of "the" objects of intentional attitudes seems to me unfortunate. I do not know what it means to talk of "the correct view" as to what sort of entity should be used to "represent the content of," or "be the object of," an intentional state. Even talk of "the best view" seems dubious: it is not in the least obvious that there is an overall best view; maybe some sorts of entities are most useful as contents in some contexts, others in another. Indeed it isn't even obvious at the outset what there is to be gained by assigning contents to

Pacific Philosophical Quarterly 67 (1986) 98-112 0031-5621/86/0200-0098-$01.50
Copyright © 1986 by University of Southern California

states at all, in giving a theory of deliberation and inquiry. So I will not be claiming that Stalnaker's decision to use a notion of content in which S-propositions are contents is *incorrect;* nor will I claim that he is wrong not to employ any finer grained notion of content (perhaps in addition). I will be objecting, though, to a model of deliberation and inquiry that Stalnaker is attached to, and which motivates his putting his account of deliberation and inquiry in terms of S-propositions.

I. The Pragmatic Picture and the Linguistic Picture

Stalnaker says that there are two very different pictures that can govern people's conceptions of intentional mental states—he calls them *the pragmatic picture* and *the linguistic picture*. He advocates the pragmatic picture, and holds that adoption of this picture over the linguistic picture leads almost inevitably to

1. the view of objects of belief as "coarse-grained," in the sense that there can not be distinct but logically equivalent objects of belief; and
2. the view that there can be no philosophical point to adopting a language of thought hypothesis.

Apparently he holds, then, that advocates of a language of thought, or of finer-grained content, deny the pragmatic picture of intentional states.

It seems to me that his discussion of these matters is quite misleading. The pragmatic picture of belief and desire is described in part as follows:

> . . . [O]ur conceptions of belief and of attitude pro and con are conceptions of states which explain why a rational agent does what he does. . . . Linguistic action, according to this picture, has no special status. Speech is just one kind of action which is to be explained and evaluated to the same pattern. (p. 4)

The pragmatic picture described here is most compelling. Stalnaker says that philosophers who have advocated thinking of belief and desire in terms of a "language of thought" or an "internal system of representation" have had a different picture, but that does not seem to me to be true: except for the restriction in the above quotation to *rational* action, the quotation seems to precisely describe the attitude toward belief and desire of Fodor's book *The Language of Thought* or of Harman's book *Thought* or of my article "Mental Representation". Indeed, although Stalnaker begins his characterization of the supposedly contrasting "linguistic picture" by saying that on this picture "rational creatures are essentially speakers" (p. 6), he ends up saying that any view that conceives the intentionality of intentional mental states on analogy with the intentionality of linguistic expressions counts as

a version of the linguistic picture (p. 7). But then the pictures do not contrast: the whole point of much of the argumentation in the three works cited was that there was reason to postulate structure resembling linguistic structure in intentional mental states *in order to explain non-linguistic as well as linguistic behavior.* I'm not here saying that this argumentation is right—in fact, I now think that my claims on this point in "Mental Representation" are rather overblown and are stronger than I needed for the points I was primarily concerned to make in that paper—but I think that Stalnaker's long discussion of language of thought views is seriously warped by his suggestion that the point of a language of thought theory is to do something other than explain behavior.

Turning now to the issue of "coarse-grainedness" of content: an almost universal first reaction to a possible worlds view of belief and desire like Stalnaker's is that it can't be adequate, since people can clearly believe that P and disbelieve that Q when P and Q are logically equivalent. Now, Stalnaker has an ingenious reply to this line of objection, which I will consider at the end of the paper—but for now I just want to emphasize that this line of objection against Stalnaker is an initially natural one *from the point of view of the pragmatic picture*. That is, even when one's task is to explain *non-linguistic* behavior, there is often a *prima facie* need to attribute a belief in one proposition and a disbelief in an equivalent proposition (i.e., there is a *prima facie* need for a conception of content more fine-grained than S-content). For instance, if I offer someone who doesn't know much mathematics $1000 for an example of a plane map that requires more than four colors to color (according to the usual coloring conventions), he will behave very differently than he would if I had offered $1000 for a trisection of a Euclidean 60 degree angle by straight edge and compass; to explain this, I need to attribute different beliefs and desires to him in the different cases, and it is *prima facie* difficult to see how I can do this in a relevant way if the desire to do one impossible task is identified with the desire to do any other impossible task.

Stalnaker thinks that the pragmatic picture leads inevitably or almost inevitably to the coarse-grained conception of the objects of intentional states. Why, given the obviousness of examples like the one above? The argument appears to be that on a pragmatic picture we ought to think of the deliberation that is involved in determining behavior as consisting of *the weighing of possibilities* (i.e., weighing which possible states of the world are more likely, what the effects of various courses of actions would be in these different states of the world, and how desirable each of these possible effects might be). If deliberation did indeed have to literally consist of the weighing of logical possibilities, Stalnaker's conclusion might well follow. But why not say instead that deliberation consists of the weighing of states of affairs which are *epistemically possible for the agent*—or more accurately, the weighing of state of affair *representations* which are *treated*

by the agent as consistent? One can presumably spell out what it is to *treat a representation as consistent* in terms of the employment of that representation in thought. This proposal, unlike Stalnaker's, handles the $1000 offer case; in addition, it has the advantage of not relying on intentional notions like *representing a possibility.* In his book (p. 25) Stalnaker objects to a version of the proposal I have just made, by saying that it uses an intentional notion that can't be cashed out. I am saying that his objection is exactly backward, in that if *either* theory is to be rejected on this score it is the one that he favors.

II. The Problem of Intentionality, Part 1

Much of the first two chapters of *Inquiry* is a discussion of *the problem of intentionality*—the problem of giving a naturalistic explanation of intentional relations like believing and desiring that apparently relate people to the objects of belief. Stalnaker does not attempt a detailed solution to this problem, but he does make some sketchy remarks about how one might develop a solution that was based on the pragmatic picture. And he argues at some length for the superiority of this sort of solution over an alternative solution-sketch that I suggested in "Mental Representation" and an earlier paper.

Before discussing this, I need to make a distinction between using "intrinsically representational" entities as objects of belief and using entities that are not "intrinsically representational." Suppose that we think of intentional mental states as follows: most (perhaps all) of them have-content, and among those that do there is a relation of sameness-of-content (applicable at least between states of the same agent), explicable in terms of functional role, and which is an equivalence relation. Then in one reasonable sense of 'objects of belief,' we could use the equivalence classes of a person's states under the sameness-of-content relation as objects of belief; this would be an example of what I have in mind as a conception of "objects of belief" according to which they are *not* "intrinsically representational." By contrast, if we view the object of a state of believing that Caesar crossed the Rubicon as an ordered triple whose components are Caesar, the Rubicon, and the relation of crossing, that is an example of an intrinsically representational object of belief: it is intrinsically representational because it can be construed as representing the world as such that Caesar crossed the Rubicon. Similarly, if we take as the object of the belief-state the set of possible worlds in which Caesar crossed the Rubicon (or, what comes to the same thing, the S-proposition that assigns 'true' to worlds where Caesar crossed the Rubicon and 'false' to the others), that too is a conception of objects of belief as intrinsically representational.

It seems to me that the most central philosophical question about intentionality is whether there are any good reasons to utilize intrinsically rep-

to give a place to something roughly on the order of "rules of logical reasoning"; and that involves enlarging on the Boolean structure.

Indeed, this extra structure is important not only in giving an idea of how the Boolean structure might be represented. It is also of direct importance in psychological explanations (indeed, *more* important than the Boolean notion of implying): think of the $1000 offer discussed earlier. I do not claim that the enriched structure need be *linguistic* structure in any very obvious sense (and indeed, the bounds of what counts as linguistic structure and what doesn't are quite vague); my point is only that it is important to develop psychological models that contain that extra structure, and that "language of thought" models at least have the virtue of doing this.

Of course, Stalnaker recognizes that there is a distinction between a person whose beliefs we would describe with one sentence and a person whose beliefs we would describe with a logically equivalent (but *unobviously* equivalent) one: he suggests in one place (p. 23) that the *objects* of these two people's beliefs are the same, but that the *forms* by which the common object is represented in the two cases differs. Presumably, then, he could recognize "rules of logical reasoning" (or some such thing): they would have to be stated in terms of the *forms* by which beliefs are represented, not in terms of the *contents* of those beliefs. I don't see that we can call this insistence on excluding form from the objects of belief *wrong:* he's free to use the term 'object of belief' as he likes. But Stalnaker seems to regard it as a substantive difference between his account and "the linguistic picture" that the former recognizes the form/content distinction while the latter doesn't; he says that "the conceptual separation between form and content is . . . the central feature which distinguishes the conception of thought implicit in the pragmatic picture from the one implicit in the linguistic picture" (p. 23). But the substance of this "central issue" eludes me. As far as I can see the serious substantive issues are whether one needs to go beyond Boolean structure in giving psychological explanation, and if so what the nature of the extra structure postulated should be.

III. The Problem of Intentionality—Part 2

Now let us turn to phase two of the problem of intentionality: the special problem that arises when one takes one's objects of belief to be "intrinsically representing" entities.

Half of Stalnaker's Chapter 2 is devoted to a criticism of a strategy for solving this part of the problem of intentionality which I suggested in my "Mental Representation" (relying heavily on my "Tarski's Theory of Truth").

The strategy of my "Mental Representation" was to assume that mental states of believing, desiring, and so forth could be given a roughly sentential structure, and then to impose on these states a componential semantics

which specified truth-conditions for sentences with analogous structures. That is, a person explicitly believes a proposition if he is in a state of believing that has that proposition as its content; where the association of contents with states of believing goes via a compositional semantics that utilizes the sentential structure of the states. (This was to hold for explicit belief rather than for belief generally. Belief generally was viewed roughly as a disposition to explicitly believe.) The main ideas of this picture could be easily generalized so as to apply to lots of non-sententially structured states, e.g. states with pictorial structure. But the picture doesn't apply to states with only Boolean structure, and that of course is one of the reasons that Stalnaker is unsympathetic to the picture. Aside from that general reason however, he suggests several specific criticisms of the strategy that I presented, some of which I will mention after describing the strategy in more detail.

Underlying the strategy in question is a specific view of componential semantics: that the point of a componential semantics is to explain the semantic features of sentences in terms of the semantic features of their component parts. For instance, we might explain the truth-conditions of 'It is not the case that Caesar crossed the Rubicon' in terms of the fact that 'Caesar' stands for Caesar, 'the Rubicon' for the Rubicon, and 'crossed' for the relation of crossing, and the fact that 'it is not the case that' is a symbol of negation. Of course, a componential semantics says nothing about *in virtue of what* one of the non-logical components (such as 'Caesar' or 'crossed') stands for what it does; but I expressed hope that something vaguely like a causal theory of reference, applied to the structural components of a belief state (the "morphemes of the language of thought") might serve to fill the lacuna. I skirted the issue of logical connectives (as Stalnaker points out—this is one of his criticisms), but it seems to me pretty clear what I should have said about them: the facts in virtue of which 'it is not the case that' means negation are facts about its conceptual role; more specifically, the standard truth-tables for 'not,' 'and,' etc. are the only truth-tables that make standard inferences involving these words truth-preserving.

Stalnaker seems to have two main objections to the sort of picture just sketched. The more interesting one is this: Stalnaker says that the approach that I advocated is too *atomistic* in that it claims that the most basic kind of representation holds between words as opposed to sentences, or "morphemes in the language of thought" as opposed to whole states. I have some feeling that there may be a legitimate complaint here, but I am unsure exactly what it is. Two points on this:

1. Stalnaker characterizes my view as holding that "the name-object and predicate-property relations *come first;* the sentence-proposition relation is derivative" (p. 34). Is this supposed to mean that first people invented names and predicates, and then

some genius thought of putting them together to form sentences? Obviously I never held that view (nor the analogous view at the mental level). Rather, the view was that our *goal* is to explain the truth-conditions of sentences and belief-states; name-object and predicate-property relations (and their analogues for components of belief states) are theoretical relations needed in characterizing those truth-conditions; and *an independent characterization of those theoretical relations is then needed* (just as an independent characterization of valence in terms of chemical structure is needed, even though the *goal* of talk of valence is the holistic one of explaining chemical combination).

2. A theory which is formally atomistic in the sense required above may accommodate a great deal of interaction among the "atoms". Thus a theory of reference for names might well have the form "reference is that relation R which is the first member of a pair <R,X> such that if you assign to each name the R-related object and to each predicate the X-related property and calculate truth-conditions in the usual way, then the resulting truth-conditions are such that blah-blah-blah." (Some sort of restriction, e.g. to causal relations, would of course be needed to eliminate some of the indeterminacy.)

I don't know if these points are enough to make "atomicity" seem unobjectionable. It does seem to me however that Stalnaker's own proposal, to which I now turn, is unworkable largely because it is *not* atomistic in the sense that my proposal was.

Stalnaker emphasizes that there are unproblematic naturalistic relations between objects and S-propositions; among them is the relation of *indication*, which Stalnaker takes to be particularly relevant to the problem of intentionality. He says,

> Consider an object which has intrinsic states that tend, under normal or optimal conditions, to correlate with its environment in some systematic way, and where the object tends to be in the state it is in *because* the environment is the way it is. For example, the length of a column of mercury in a thermometer tends to vary systematically with the temperature of the surrounding air. . . . (p. 12)

Under these conditions we say that the length of the column of mercury *indicates* the temperature of the surrounding air. The suggestion that emerges later is that "belief is a version of the propositional relation I called indication" (p. 18).

I think there can be little doubt that if there is to be any hope of solving the problem of intentionality on the supposition of intrinsically representing

objects of mental states, then something like indication must play a role in the account. The crucial question is, what exactly is the role it plays? In both my Tarski paper and "Mental Representation," I argued that the motivation for introducing intrinsically representing objects (or equivalently, a correspondence theory of truth) into semantics and philosophy of mind was to enable us to formulate such indication-relations as exist between belief-states (and utterance-states) and the environment. Semantic notions such as 'truth-conditions' were thus to be viewed as theoretical terms in a "reliability theory," i.e. a theory of indication. Semantic notions such as 'refers' were viewed as still more theoretical notions of the theory; they were needed to formulate truth-conditions which were in turn needed to formulate indication relations. An attractive feature of this way of viewing things is that many of the features which have been argued to be essential to a plausible theory of reference are features that clearly have to enter in to any explanation of why my states or utterances are reliable indicators of the external world. For instance, to the extent that my utterances containing the phrase 'the Lyons silk-weavers strike' are reliable indicators of facts about the Lyons silk-weavers strike, that reliability is obviously totally dependent upon the existence of "experts" whose reliability I inherit; and the reliability of the experts (and hence, indirectly, of me) obviously depends on a causal network of some sort emanating form the silk weavers strike to their (and my) belief states and utterances.

That's one view of the relation between indication and belief. It would however be a bit misleading to sum up *that* view by saying that belief is a species of indication (since on that view there is no reason to suppose that belief states in general reliably indicate the states of affairs that we would regard as their contents); and in any case, since that view is atomistic, it seems clear that Stalnaker's view is different.

But just what is Stalnaker's view? The obvious problem with making belief literally a species of indication is that most people are unreliable about a great many things. Are we to say that the unreliability of the average person's beliefs about science or about their own motivations or about the reasons politicians act as they do is due to the fact that the average person is not under "normal or optimal conditions"? It is hard to see how the notion of "normal or optimal conditions" could be spelled out so as to give this result, *especially* since we would be unable to use intentional or semantical notions in the spelling (those being the notions that Stalnaker is trying to explicate). It seems initially more attractive to try to loosen the connection between belief and indication, in such a way as to be compatible with the possibility that even in quite normal and optimal conditions our beliefs be unreliable. The view I advocated in the papers Stalnaker criticizes was an attempt to do this, but it relied on a formally "atomistic" conception of representation. Can we get the desired result without the "atomism"?

Stalnaker sometimes seems to suggest that there are more resources available than just the indication relation. For after discussing indication, Stalnaker goes on to say this:

> Belief and desire, the [pragmatic] strategy suggests, are correlative dispositional states of a potentially rational agent. To desire that P is to be disposed to act in ways that would tend to bring it about that P in a world in which one's beliefs, whatever they are, were true. To believe that P is to be disposed to act in ways that would tend to satisfy one's desires, whatever they are, in a world in which P (together with one's other beliefs) were true. (p. 15)

That these facts hold of belief and desire is of course totally uncontroversial; the only question at issue is whether they can be used to give a non-circular explanation of the belief relation and the desire relation. The obvious way to try to use them for that purpose would be as follows:

> Consider pairs (B,D), where B is a relation that maps an agent's belief-states into propositions and D maps the agent's desire-states into propositions. Then *the* belief relation is the first component, and *the* desire relation the second component, of the unique pair <B,D> such that the agent is disposed to act in ways that would tend to bring about the propositions in the range of D given the truth of the propositions in the range of B.

The problem, of course, is that talk of "the unique pair" is totally inappropriate: indeed, given *any* relation B that is a candidate for the belief relation (in the sense of being the first component of such an ordered pair) and *any function at all,* call it h, *that maps the set of possible worlds 1-1 onto itself,* then the relation B_h defined by composition of B and h in the obvious way is also a candidate for the belief relation. [B_h is defined as follows: state s stands in the relation B_h to the set of worlds P if and only if s stands in the relation B to $\{h(w) \mid w \in P\}$.] Similarly for desire. This leaves the contents of belief and desire *totally undetermined.*

I don't claim that this is news to Stalnaker: he makes what is close to the same point, and sums it up by saying that "the *content* of belief and desire cancels out on the pragmatic analysis. Even if that analysis does give us an account of the structure of explanations of rational action, it gives us no account at all of how beliefs represent the world" (p. 18). This suggests that as far as the pragmatic picture goes, there is no need for intrinsically representing contents; only the Boolean structure of mental states (or some more complicated computational structure of the sort I discussed earlier) is of interest to psychology. But to accept that conclusion would of course be to give up on thinking of belief as a kind of indication.

Stalnaker does not accept that conclusion; instead, he says, a pragmatic analysis gives only part of the picture (p. 18). The part of the picture that it gives us is (i) that of determining which states are belief states and which states are desire states; and (ii) that of helping to determine the content of desire states *once the contents of belief states have been independently determined.* But as far as determining the content of belief states goes, Stalnaker ends up forgetting about the pragmatic picture and putting the whole burden on the relation of indication. And so the obvious problem is left totally unresolved: how are we to bring in indication in our account of belief-content without flying in the face of the obvious fact that most people's beliefs about a great many things are extremely unreliable? Above I sketched one (admittedly quite programmatic) strategy for solving this problem, the strategy that I advocated in the papers that Stalnaker is criticizing; but that is precisely the strategy to which Stalnaker is trying to find an alternative. It is quite unclear to me what the alternative strategy is to be.

There does seem to me to be one suggestion in Stalnaker's discussion that is worth pursuing. He distinguishes *backward-looking relations* like indication, in which "a state of an object is defined in terms of what tends to cause it," and *forward looking relations,* in which "a state of an object is defined in terms of what it tends to cause" (p. 14). In the main text, the only role that forward-looking considerations are given in solving the problem of intentionality is in determining what states are belief-states and what states are desire-states, and in determining the content of desire states by their role in action once the content of belief-states have already been determined. So the main burden of determining content still rests upon the backward looking (roughly, causal) considerations. Still, a natural suggestion (and one that Stalnaker makes in a footnote) would be to give the forward looking considerations a more prominent role in determining content. This is an attractive idea. The idea is, of course, neutral to the issue of whether a theory of representation be "atomistic" in Stalnaker's sense, in that an atomistic theory of representation could just as well invoke forward looking considerations in its theory of reference as could a non-atomistic theory. I suppose it is conceivable that the invocation of forward looking relations in addition to indication could be used to help solve the problem for non-atomistic theories that I have been discussing; but I have only the foggiest of ideas as to how this would go.

IV. Failure of Logical Omniscience

I now want to return to the difficulties that are apparently raised for Stalnaker's view by the fact that people's beliefs appear to be neither consistent nor closed under logical consequence. That there is a *prima facie* problem here is clear: Stalnaker wants to describe the beliefs of not only the "ideally rational believer," but of ordinary believers as well, as relations to sets of worlds that are (at least) *logically* possible. How then can inconsistency and failure of closure under consequence be represented? I will confine my discussion here to the problem raised by inconsistency. [*Logical* inconsistency: perhaps in a broader sense of 'inconsistent,' 'Hesperus is a

planet' and 'Phosphorus is not a planet' are "inconsistent," on the grounds that there is no "metaphysically possible" world in which both are true; but I don't think that "inconsistency" in this expanded sense need cause any worry to Stalnaker. For the fact that the beliefs are logically consistent guarantees that there is at least a *logically* possible world in which both hold; and I do not see any reason why he should require anything more. The discussion to follow will focus, then, on purely logical inconsistency.]

Stalnaker has a discussion in Chapter 5 which appears to me to plausibly handle many cases of inconsistency. Suppose that, in the course of theorizing about the philosophy of mind, I come to accept the existence of some sort of entity (perhaps propositions or properties or sets) which I reject while doing ontology. If I notice the conflict, then presumably I will take an attitude of less than belief toward at least one of the conflicting existence claims. But if I *don't* notice the conflict, it is quite conceivable that I will believe in the entities in question in one context while disbelieving in them in another. Is this sort of inconsistency among beliefs a problem for one who wants to describe beliefs as relations to S-propositions or sets of possible worlds? Stalnaker points out that it is not. For the natural way to describe what is going on is that I have *two* belief states involving the existence of these entities—one belief-state according to which the entities exist, the other according to which they don't. In one sort of context, one of these belief-states enters into the explanation of my thought and action, in other contexts the other does. If I noticed the conflict, I would presumably try to integrate these states; but prior to my doing so, the description of me in terms of these two conflicting belief states seems intuitively correct. And there is no problem with describing the content of either of the two belief states in terms of S-propositions (or sets of logically possible worlds).

I have no doubt that the above provides an attractive solution to the problem raised by this sort of example of inconsistency of belief. However, there is another sort of example, which Stalnaker does not discuss. Consider the belief-states of Cantor while he was developing set theory. Cantor assumed a "naive comprehension schema" for set theory; one instance of this proved logically inconsistent (Russell's paradox). Surely there is no plausible way to explain this along the lines above: rather, it would appear that one must attribute logical inconsistency to the content of a single belief state.

Stalnaker cannot accept the idea that a single belief-state has inconsistent content. The reason is that on Stalnaker's view of content, two things hold. First, any belief-state with inconsistent content has the same content as a state of believing that snow is both white and not white would have; and second, belief-states with the same content have to enter into the explanation of behavior in essentially the same way. If these assumptions are both correct, then we have to find some way of describing Cantor's belief so that it has a consistent content. If we can not do this, that is good reason

to reject at least one of the two assumptions. (Which one to give up is largely a matter of convention as to the use of the word 'content'.)

Stalnaker discusses one other method which might be employed in an attempt to describe Cantor's belief state as having a consistent content: the method of going metalinguistic. That is, instead of ascribing to Cantor a belief-state whose content is the infinite conjunction of the comprehension axioms of naive set theory (or some such thing), we ascribe to him a belief-state whose content is that these sentences *express truths*. This maneuver strikes me as having little to intuitively recommend it, but I will not press that point (except to note that it is an accident of the example that the *prima facie* subject matter of the belief is mathematical; the same sort of inconsistent characterization could arise in describing a subject matter with an uncontroversially non-linguistic content). A more basic doubt is whether the metalinguistic move will work. The move does succeed in giving Cantor's belief-states *consistent* content, since it is consistent to hold that the sentences that in fact characterize set theory characterize the decline of the Roman empire instead. The trouble is that it appears to be of little use to Stalnaker to ascribe to Cantor's belief state a content that is *merely* logically consistent; the content needs to be consistent with other facts that Cantor perfectly well knew. And among the facts that he knew, we may assume, were the fact that 'ϵ' means set membership, that '&' means and, and so forth; these known facts imply that the axioms of set theory mean just what they do in fact mean. If we now consider the metalinguistic belief *that the naive comprehension axioms express truths and that they mean so and so* (or, *that the naive comprehension axioms express truths and that their meaning is determined by such and such rules*), then that metalinguistic belief is itself inconsistent. So the ascent from object-level propositions to metalinguistic propositions appears to have gained nothing.

A number of people made essentially this objection to an earlier presentation of Stalnaker's views. In *Inquiry,* Stalnaker considers the objection (in a more abstract setting, not tied to the Cantor example). His reply is to grant that even the metalinguistic beliefs ascribed to, say, Cantor form an inconsistent set; but then to introduce the discussion (which I've summarized above) of how having an inconsistent set of beliefs is not ruled out by the possible worlds analysis, since the belief-states that compose this inconsistent set may themselves be consistent. It seems to me however that this is plainly inadequate: Cantor's metalinguistic belief that the comprehension axioms express a truth and his other metalinguistic belief about the meaning of these axioms are not at all like a belief in properties while doing philosophy of mind and a disbelief in them while doing ontology; it isn't as if Cantor had two distinct belief-states that could not both be operative at once. Rather, the two belief-states were involved together in explaining his actions, and the discussion in Chapter 5 of "compartmentalized" belief just seems irrelevant to the case at hand.

If, as I have been suggesting, we cannot find a reasonable way to describe belief states such as Cantor's as having consistent content, then we must give up Stalnaker's fundamental idea that the explanation of behavior is to be done in terms of mental states described solely in terms of the real possibilities that they represent. And if we give that up, there is little point in insisting on using the term 'content' in such a way that only the real possibilities that a state represents can be part of the state's content. Again, the basic difficulty with Stalnaker's views is not his use of the term 'content,' but his picture of rational action. But once that picture of rational action is undercut there is no longer any motivation for insisting that the word 'content' be used in such a way that only the real possibilities that a state represents can be part of its content.

University of Southern California
Los Angeles, California

NOTES

[1]This is the paper I presented at the symposium on Stalnaker's *Inquiry* at the American Philosophical Association Central Division meetings, May 3, 1986. It is a shortened version of my Critical Notice of that book, appearing in *Philosophy of Science* 53: (Sept. 1986). I thank the editor of *Philosophy of Science* for allowing me to use the material again for the publication of this symposium.

REPLIES TO SCHIFFER AND FIELD[1]

BY

ROBERT STALNAKER

THE PROBLEM OF intentionality is a little like the tar baby. Once you start to tangle with it, it is hard to get free. My discussion of this problem in the book was supposed to be a preliminary discussion—I wanted just to justify a framework for treating some issues in epistemology, and then to move on to those issues. But the problem dominates the book, and it is what both Hartry Field and Stephen Schiffer have focused on. It is a hard problem to which I have, at best offered a very partial and preliminary response. The difficulties that Field and Schiffer raise for this response are, I think, serious challenges that cut to the heart of the issue. I am not completely confident that these challenges can be satisfactorily answered, but I will try to clarify the project and make it seem a little more plausible.

I will begin with some general remarks about strategy and method: about reduction of the intentional to the natural, and about the contrast between the pragmatic and the linguistic pictures. Then I will comment on the abstract argument with which I tried to connect the causal-pragmatic account of intentionality with the possible worlds analysis of proposition. Third, I will try to defend the causal-informational account of representation against some of the criticisms offered by both Field and Schiffer. Finally, I will try to clarify my strategy for solving the problem of deduction and defend it against the skepticism expressed by both of my critics.

Stephen Schiffer describes my program as "the enterprise of trying to give a physicalistic reduction of mental states and intentionality," or of trying "to show that belief facts, such as the fact that I believe that I'm in St. Louis, are identical to facts statable in sentences that are devoid of intentional and mentalistic idioms." He is skeptical about the possibility of such a reduction, and I confess that I am too. I may, in careless moments,

Pacific Philosophical Quarterly 67 (1986) 113-123 0031-5621/86/0200-0113-$01.10
Copyright © 1986 by University of Southern California

have described my aims in terms of reduction, but I would be content with something more modest. I am mainly concerned with one specific kind of fact about belief and other mental states, the fact, or at least what I take to be a fact, that such states stand in intentional or representational relations. What I want to explain in naturalistic terms is the nature of such relations. Such an explanation need not take the form of a reduction. My main strategy, at least initially, was to give examples of states and relations that are much simpler than belief, but that exhibit the features that are the reason philosophers have found belief and other intentional mental states problematic. The claims, then, are first that such simple states and relations are unproblematic from a naturalistic point of view, and second, that belief is just a much more complicated instance of the kind of state exemplified. Both claims may be questioned: one may question the first claim—that the simple states and relations are unproblematic—by arguing that we understand them by analogy with, or in terms of, full blown intentional states. And one may question the second claim—that the differences between the simple representational states and genuine beliefs are just matters of complexity—by identifying problematic features of genuine belief states that are not shared by the simpler states. A successful physicalistic reduction would provide definitive answers to such challenges, but such a reduction is not necessary. I would be satisfied if I could make a convincing case that belief is a species of a kind of state all instances of which are representational in the way belief is, and some instances of which are clear examples of purely physical states. I would be satisfied with this even if I could not give a clear account of what distinguished the states of this kind that are properly called "beliefs" from those that are not. (Although I think this latter question is also an interesting one, and one that one might be able to say a lot about even if one could not give a reductive definition.)

Second, let me respond to Hartry Field's comments about my contrast between the linguistic and the pragmatic picture. He thinks that my use of that contrast is misleading, and that it has seriously warped my discussion of the language of thought hypothesis. I did paint the linguistic picture with pretty broad strokes, and he may be right that the way I used this contrast to categorize theses and projects tended to distort them. Let me try to be a little clearer. First, my initial aim was just to make explicit a cluster of metaphors and analogies that I think have motivated and influenced many discussions of intentionality. I called it the linguistic *picture* because I was talking not about a thesis, a doctrine, or an analysis, but a way of looking at the issues. My characterization of the linguistic picture was something of a caricature, and I did not intend it to represent a philosophical position, or even a clear set of assumptions about the relation between thought and speech. I did then go on to identify a more specific issue, a question about the strategy for solving the problem of intentionality: should one explain the intentionality of thought in terms of the intentionality of language, or

the intentionality of language in terms of the intentionality of thought? This contrast needs to be developed and explained, and there are several different ways of taking it, but the issue has been discussed explicitly in these terms (for example in the famous published correspondence between Wilfrid Sellars and Roderick Chisholm[2]), and I think the contrast between the strategy Field defended in "Mental Representation" and the one I wanted to defend can usefully be described in this way. But I should have more clearly distinguished the picture from the philosophical question. I did not mean to saddle Field with all the baggage suggested by my caricature.

Second, it is clear that most of the philosophers I associated with the linguistic picture, including Field, Harman, Fodor, and Davidson, were guided in their discussions of intentionality by the pragmatic as well as the linguistic picture. All explicitly defended a functionalist or dispositional account of belief and desire. My caricatures emphasized the contrasts and exaggerated the opposition between the two pictures. Real philosophical positions are more complicated, consistently combining elements from both sides.

Third, my focus on the contrast between the linguistic and the pragmatic picture may have tended to obscure other contrasts that cut across this one. The two projects that I discussed in detail under the heading of the linguistic picture—Field's and one modeled on Davidson's theory of interpretation—are in some ways at opposite extremes. Field's is atomistic while Davidson's is holistic (Though I agree with Field that much more needs to be said about just what this contrast comes to); Field's project was individualistic, and focused on internal states, while Davidson's project regards intentionality as essentially social, and focused on external relations.

Despite the ways in which my discussion of the linguistic picture may have been misleading, I think it is useful to focus on the ways that theorists have relied, implicitly and explicitly, on the analogy between language and thought, and to look critically at some of the assumptions that are motivated by this analogy. The analogy may be fruitful in some contexts, but I think it has had a pernicious influence, and that it is a useful corrective to consider non-linguistic models and explanations of representation.

Both Field and Schiffer say they don't see how the pragmatic-causal account of representation motivates the possible worlds analysis of proposition. Let me quickly sketch the main argument for this, which is not as clear in the book as it should have been, and then comment on its status. The argument begins with the assumption that the fact that some state has the content that P is to be explained in terms of modal, dispositional and counterfactual facts expressed in terms of the proposition that P, or perhaps in terms of propositions that entail P. In such statements, substitutivity of necessary equivalents preserves truth. So necessarily equivalent propositions will be functionally equivalent for any propositional relation defined in these terms, and functionally equivalent objects should be identified.

Now this is a very abstract argument. Its conclusion clashes with lots of obvious *prima facie* counterexamples, and as Field brings out, the counterexamples include some that seem very natural from the point of view of the pragmatic picture. That is, there are counterexamples even among beliefs that are invoked to explain non-linguistic behavior, the kind of beliefs that fit most naturally with the pragmatic account of belief. So there are counterexamples, and the argument does not, by itself, point to a way of avoiding or explaining away the counterexamples. The argument leaves us with a puzzle: from an abstract point of view, it seems that a naturalistic explanation of intentional mental states must meet certain apparently very weak conditions. But any explanation that meets those conditions has consequences that apparently clash with the phenomena. We must, it seems, either deny one of the premises of the argument, find a fallacy in it, or explain away the counterexamples.

Both Schiffer and Field express some skepticism about the adequacy of the kind of naturalistic account of representation that I endorse: the Wisconsin-style causal-informational account developed and defended by Dennis Stampe and Fred Dretske, among others. Let me sketch the general idea of the account, and then try to reply to some of the problems they raise for it.

A representational system, according to this account, is a system that is capable of being in a range of alternative states that tend to correlate with, and be caused by, a corresponding set of alternative states of the world. The causal dependencies that hold in such a system could be expressed in a set of nested counterfactuals of the following form: if conditions are normal in the relevant sense, then the system is in state S_i only if the world is in the corresponding state W_i. This, as Schiffer emphasizes, is an abstract schema. To explain how a particular system is a representational system one must say what the relevant states are, and say more about what it is for conditions to be normal or optimal in the relevant sense. The claim is that belief states are states of a system of this kind, and that the explanation of how beliefs have content should be given in terms of this kind of relationship between belief states and the world.

Before looking at some of the problems, let me make a side remark: Field agrees that something like this notion of representation must play a role in an account of belief, at least if we assume that mental states have intrinsically representational contents. His sketch of an explanation of why his beliefs expressed with the phrase 'the Lyons silk-weavers strike' are about the Lyons silk-weavers strike seems to me exactly right. He says that this explanation must be different from the one I would give since it is atomistic, while I reject atomism. I am no longer very clear about what my atomism objection was, or whether it applies to the view Field was defending in "Mental Representation." I claimed that one could not explain the representational properties of names except in terms of the role names

play in the expression of sentences with certain representational properties. An explanation of reference in terms of reliability is not atomistic in any sense I object to, since it is the statements involving names, or the beliefs expressed with them, that are reliable.

Now for the problems: if beliefs are examples of representations of this kind, then, as Schiffer says, "when conditions are optimal, no one can have any false beliefs." Both he and Field express doubt that optimality conditions relative to which believers are infallible can be spelled out in non-intentional terms. Neither think that it is at all plausible to assume that beliefs tend, under normal conditions (in any reasonable sense of this phrase) to be true. We are all unreliable about a great many things. Schiffer, in fact, goes so far as to say that whatever these optimality conditions are, "it is clear that they *never* obtain."

This is a central problem for my account, but I think more can be said, and less needs to be said, than Schiffer and Field suggest. First, the assumption that beliefs are true under normal conditions (in some reasonable sense of normal) is compatible with lots of beliefs being false, and with believers being unreliable about a lot of things. Normal conditions, in at least one normal sense of that expression, need not be conditions that hold in some large proportion of cases. A reasonable concept of normal conditions requires only the assumption that deviations from the norm require explanation. It does seem to me that there is a presumption that beliefs are true—that false belief needs an explanation in terms of some limitation or abnormality, either in the belief forming mechanisms, or in the information available in the believer's environment. The presumption may be defeated in a great many cases, but it still operates. It would be suspicious if one had to say that the presumption was *always* defeated. But I don't know why Schiffer says that it is clear that the kind of normal conditions the account requires *never* obtain. It seems to me that conditions for the forming of beliefs are often normal in the relevant sense. Take any ordinary case of perceptual belief: the light and my visual system are normal and I am awake and attending to what is before me, which is an ordinary scene devoid of natural or contrived illusions. No preoccupations, preconceptions or emotional hangups prevent me from seeing, and coming to believe, that a dog is sleeping on an oriental rug in the hallway. Now I know I haven't told you, in non-intentional terms, what the normal conditions are that obtain. I am only making the point that in such ordinary situations, the relevant conditions *do* obtain.

Furthermore, even though the presumption that beliefs are *normally* true in some reasonable sense does not imply that beliefs are *usually* true, it also seems to me right to say that believers are on the whole pretty reliable—that most of what we believe is true (though please don't ask me to make this quantitative claim precise). Our attention is naturally focussed on beliefs that are more speculative or controversial, and so less reliable, but there is

a vast amount of information available to us that is in some sense unremarkable, but that is still important in permitting us to find our way about in the world. We have general tacit beliefs that determine our expectations and constrain our intentions, such as my beliefs that the floor will support me as I walk across it, and that I can't get to the other side of the wall by walking through it. We also have lots of specific beliefs about the configuration of our immediate environment, beliefs that are constantly changing as we move about in it. Beliefs about quantum mechanics, interest rates, the Philippine government, or God are just the unreliable tip of a very reliable iceberg.

Schiffer asks why should we have any confidence in this general account of representation, which he calls the fuel gauge model. We cannot, he says, give a clear naturalistic account even of such simple representational systems as fuel gauges. Now I agree that it would be suspicious if all the simple models of this account of representation were artifacts that were designed to represent. The suspicion is that their representational properties derive from the representational properties of the intentions of the designer. But there are simple natural examples: the shape of the tree is represented by its shadow, the number of rings on the cross section of the tree represents its age. There are also more interesting cases: it seems possible and fruitful to see some biological mechanisms, for example immunological, genetic and perceptual mechanisms, in these terms. The kind of pattern of counterfactual dependency described in the causal-informational account seems to be present in such systems, and seems to be necessary and sufficient for the appropriateness of descriptions of them in terms of information and representation. The conceptual order is somewhat tangled: it is clear that theories that describe mechanisms in terms of information and representation begin with an analogy with mental and linguistic representation, and one may question whether it is legitimate to turn around and use such theories to help support a naturalistic account of mental states. But if we can make explicit the basis for the analogy, and if we can see the structure that is claimed to be necessary and sufficient for representation more clearly in the simple cases, then I think this way of defending a naturalistic account is unobjectionable.

How much can be said, and how much needs to be said, about the normal, optimal or fidelity conditions that are relevant to understanding belief as a kind of indication? I think very little of substance can be said, or will ever be able to be said, on the basis simply of an analysis of the concept of belief. The reference to normal conditions is a kind of promissory note, but I think it is a note to be cashed in future empirical inquiry, not future conceptual analysis. If you want to know, in non-circular and non-intentional terms, what the relevant normal conditions are, you need to learn how the actual mechanisms by which we form beliefs work. They don't all work the same way, of course, and I don't think it will ever be possible to

give a *global* account of fidelity conditions. But it does not seem to me implausible as we come to understand the way specific belief forming mechanisms, for example perceptual mechanisms, work, we will be able to say something substantive about the range of conditions under which such mechanisms work normally. A defense of the plausibility of the hypothesis that belief forming mechanisms are capable of being understood in this way is a long way from a reductive definition of belief in naturalistic terms, and it is not a trivial matter to give even this more modest defense, but if we can, this may be enough to reconcile intentionality with the natural order.

> Schiffer says that what we need is "a completion of the schema, were x's belief-forming mechanisms to have property F and x's physical environment to have property G, then x would form only true beliefs,

where the substituends for 'F' and 'G' specify properties non-intentionally." I agree that the prospect of specifying such properties in a general definition of belief, or even the hypothesis that such properties exist, seems wildly implausible, but I don't think we need to fill this schema in, once and for all, to solve the problem of intentionality.

Along with his general implausibility arguments, Schiffer has some examples that suggest that the causal-informational account of belief may not fit very well with the phenomena. Consider Fido: I believe that *Fido* is standing before me, but wouldn't I be in exactly the same state if some qualitative duplicate of Fido were standing before me instead, and mightn't conditions still be normal if this alternative situation obtained? Information, according to the kind of account I want to defend, is defined relative to a range of alternative possibilities which need not be an exhaustive set of possibilities. Information is essentially contrastive: to say that I know or believe that P is to say that my internal state distinguishes certain situations in which P is true from relevant alternatives in which P is false. In a normal context in which the belief that Fido is present is attributed, we will be concerned with the contrast between situations in which Fido is present, and situations that might plausibly obtain if he were not. Situations in which a qualitative duplicate of Fido is before me are not generally relevant alternatives, but one could easily make them relevant by raising the question, "why do you think it is *Fido* and not some other dog that looks exactly like Fido?" My answer to this question will not be that I can just see that it is Fido; *this* belief, if I have it, will not be a perceptual one, and so the relevant normal or optimal conditions will not be just conditions governing the perceptual situation. If someone had manipulated the situation by replacing Fido with his double, the conditions would surely be, in some reasonable sense, abnormal.

Or consider Sally with the flu. Proposition Q is a complete explanation for the fact that Sally has the flu; Q is causally necessary and sufficient, in the circumstances, for Sally to have the flu. So, it would seem, on the causal-informational account, I will believe Sally has the flu if and only if I believe Q. But of course I might believe she has the flu without knowing why. This is an instance of a very general problem: as Schiffer says, "Propositions are enmeshed in a huge causal nexus." For any proposition, there will be other propositions causally necessary and sufficient for its truth, but we won't necessarily believe such equivalent propositions together.

This is a very serious problem for the account I am defending. It is bad enough that I am stuck with the conclusion that necessarily equivalent propositions are identical; the kind of examples and arguments that Schiffer offers show that the causal-informational account of representation will have trouble distinguishing even *causally* equivalent propositions. I don't have a worked out solution to this problem, but I have some rough ideas about some of the resources available for a solution. First, I would emphasize again the context-dependence of belief attributions—the dependence on a presupposed set of contrasting alternatives. An attribution of content not only determines a set of possible situations, it also suggests or constrains a set of alternatives relative to which the attribution is to be understood. Causally, or even necessarily equivalent, propositional expressions may suggest different contrasts. One may believe that Sally has the flu (rather than having a cold, or being healthy) without being able to distinguish situations in which the disease we call "the flu" is caused by a virus from a situation in which it is caused by a bacterium. Second, I would note that the strategies available for solving the problem of deduction, or of logical omniscience, may also be applied to the problem of distinguishing causally equivalent propositions. I am a long way from making a convincing case that those strategies will solve the problem of deduction, but if they can be made to work at all, I think they can be applied more widely. Let me now make a few concluding remarks about that problem, and those strategies.

There are two parts to the strategy I suggested for solving the problem of deduction: the first is what is sometimes called the metalinguistic move, the claim that some beliefs, even if expressed in the material mode, are really beliefs about the relation between a sentence or other representation and a proposition it expresses. The second part of the strategy is the recognition that beliefs may be fragmented or compartmentalized—that we may be in different non-integrated belief states. I want to emphasize that these are complementary strategies; neither will get anywhere at all without the other. The explanation of examples of inconsistent belief of the kind that Hartry Field gives will involve both moves. Let me say a little more about each strategy, and about how they interact. I hope I can make it at least a little more plausible that they might contribute to an explanation of deductive ignorance and inquiry.

The first strategy is more general than the word "metalinguistic" suggests. First, it may apply when the subject matter of the belief is not in any natural sense linguistic. Consider a belief that Hesperus is not Phosphorus, that the flu is not caused by a virus, or that I'm a paragon, while that guy [pointing to what I fail to realize is a photo of myself] is a moral degenerate. In each of these cases we may ask, what does the belief attribution seem to imply about the way the world is according to the believer. Worlds compatible with the believer's beliefs are obviously not worlds in which Hesperus is distinct from Phosphorus—that is, from itself, or in which the actual disease, influenza, is not caused by a virus, or in which I am both a paragon and a moral degenerate. They are instead worlds in which distinct planets are named by the expressions, "Hesperus" and "Phosphorus," a different disease is called "the flu," and the picture I demonstrate is a picture of some person other than me. We can express the relevant proposition, in each case, by referring to the relation between a sentence and a proposition, but we need not say that the belief is about a linguistic or semantic fact. The relevant facts are astronomical, medical, and historical. The strategy can be applied to mental events that do not involve language at all. I might wonder as I look more closely at the picture, whether that is me without putting my question into words. The strategy, generally, is to take a proposition to be about a relation between a representation, which could be something other than a linguistic representation, and its content. If a representational system has different levels and components, it might be that some of my belief states are about other of my belief states. I might believe that P, relative to belief state one, while not knowing, relative to belief state two that this belief is a belief that P.

Field's example of fragmented belief—I make certain ontological commitments while theorizing about the mind which I deny when doing ontology—is only one very limited kind of case of what I have in mind in applying the second part of the strategy. Integrating separate belief states is not always simply a matter of paying attention, in one context, to what you are doing in another. It may be a non-trivial task to put one's beliefs together, even when one attends to them together. Let me use Field's lightbulb model to try to illustrate my point. (Field says that I would disavow any such model, but in fact I rather like it.) Suppose I have *two* arrays of lightbulbs each of which is wired up to the world so that which lightbulbs are lit up tends to carry certain information about the world. Each bulb of each array corresponds to a possible situation. The bulbs lit are the ones corresponding to the possible situations compatible with the information received through that array's input channel. In such a case, one array might carry the information that P and the other the information that Q without there being any available representation of the information that P&Q. (The system as a whole does represent, or carry the information that P&Q implicitly in the joint state, but that information may be unavailable to guide

behavior.) Even if the two arrays have access to each other in some way, or even if there is a third array that is able to take in information about which bulbs are lit in the other two, the information in the two arrays cannot be integrated unless we can match up the bulbs of the two arrays. The information needed to do this will be information about what information is carried by the primary arrays.

Let me consider a simpler and more specific example of fragmented or distributed representation. Suppose I have a device that has two components, each of which takes an integer as input and registers whether it is odd or even. Each bit of information might then be used in some way to control separate outputs of the device. Suppose the device needs to know whether the sum of the two numbers is odd or even. The device as a whole contains this information, but something needs to be added to make that information accessible. Perhaps we can feed our two one bit components into an exclusive or gate, yielding one result if the sum is even, another if it is odd. This would be a way of integrating the two separate informational states.

Now suppose we wire up our device so the same integer is fed into both of the one bit memories. Now we have redundant information. What information is carried, or tends to be carried, by the state of the component that integrates the two separate belief states? This component will register a positive signal only when something has gone wrong. We might think of it, in this case, as an inconsistent representation—that twice the input integer is odd—but it would be better to think of it as a representation of other parts of the system: a negative signal represents the system as functioning normally, while a positive one represents it as malfunctioning.

We don't need a very complex system to have levels and layers of representation. Both the so-called metalinguistic and the distributed belief strategies come into play when we try to describe representational systems that have such levels and layers. If we have a lot of structure of this kind—different representational components wired together in various ways—then we have processes going on that look a lot like computation; we have the extra structure and "rules of logical reasoning" that Field thinks we need, and we get them by trying to understand representation in terms of the causal-informational model. We, of course, are very complex systems. It does not seem to me at all implausible to suppose that the structures by which we represent the world in our attitudes exhibit these features, and that the complexities of belief and belief attribution can be accounted for in terms of them.

Cornell University
Ithaca, New York

NOTES

[1]This paper was my contribution to a symposium on my book, *Inquiry*, at the meeting of the Central Division of the APA in St. Louis in May, 1986. I want to thank Stephen Schiffer and Hartry Field for their incisive and challenging comments in their papers, and in discussion.

[2]"Correspondence on Intentionality," in H. Feigl, M. Scriven, and G. Maxwell, eds., *Minnesota Studies in the Philosophy of Science, II: Concepts, Theories and the Mind-Body Problem* (Minneapolis, University of Minnesota Press, 1958), 521–539.

PACIFIC
PHILOSOPHICAL
Quarterly

Vol. 67, No. 2
April 1986