RESPONSES TO THE MARIBOR PAPERS

Michael Devitt

In The Maribor Papers in Naturalized Semantics
Dunja Jutronic, ed.
Maribor: Pedagoska fakulteta Maribor (1997): 353-411

I set out to respond to all the papers in this diverse collection that directly discuss Coming to Our Senses (1996), and have done so in some detail. I did not set out to respond to several interesting other papers in the collection that were sent to me, but in the end could not resist doing so, albeit sometimes very briefly. I organized my responses around topics and discuss those topics roughly in the order in which they arise in Coming. In the heading for the section on a topic, I mention parenthetically any author whose paper is discussed in the section, even if only in a note. Some papers cover more than one topic and are discussed in more than one section. All references are to papers in this volume, or to Coming, unless the context makes it clear otherwise. The contents are as follows:

1. Analyticity (Yagisawa, Miscevic, Potrc)

Coming largely consists of the proposal of a methodology for semantics and the application of that methodology to a range of familiar issues. One of those issues is holism. The book starts, however, with a discussion that does not depend on the methodology: this is the discussion of the arguments for holism and hence against molecular localism. Probably the most influential of these arguments is the one from the rejection of analyticity. Responses to my criticism of this argument are the subject of the next section. But I start with two criticisms of my account of analyticity itself.

What is at issue is whether sentences can be analytic in the sense of true solely in virtue of meaning. I argue that they cannot (1.6). Takashi Yagisawa in "Knocked Out Senseless" and Nenad Miscevic in "Analytic Conceptual Truths are Vacuous" seek to vindicate the view that they can. Nothing much hinges on this so far as my defense of localism is concerned. Still, the issue is interesting in its own right

My molecular localist allows that a few of the inferential properties of a word may constitute its meaning (the atomistic localist thinks that no inferential properties constitute its meaning).[1] Suppose that \`bachelor' is such a word and that its meaning is partly constituted by its inferential links to \`unmarried'. I claim that (B), \`All bachelors are unmarried', would still not be true solely in virtue of this meaning link: it would be true partly in virtue of the truth of (U), \`All unmarrieds are unmarried'. Yagisawa wonders why I claim this.

The short answer that I had in mind was that, given the meaning link, (B) would not be true if (U) were not. All attempts to derive the truth of (B) from the meaning link rest, implicitly or explicitly, on the truth of (U).

Against this, Yagisawa makes several attempts to show that (B) is true solely in virtue of meaning. My short answer deals easily with one that he endorses (p. 21). That attempt rests partly on his unargued view that the "the meaning of the sentence form \`All...are...' assures the truth of every instance of th[e] schema": \`All Fs are G' is true if and only if \`G' applies to all to which \`F' applies to (p. 8). How could the meaning do that? Only if it could assure the truth of each instance of \`All Fs are F' - for example, (U) - which is what each instance of \`All Fs are G' on the LHS of the biconditional - for example, (B) - is (partly) "reduced to" by the semantic facts on the RHS of the biconditional. But the meaning of \`All...are...' can't assure this. The meaning of \`All Fs are F' assures that it means that all Fs are F. What makes it true, given that meaning, is that all Fs are F. And that is not a matter of meaning at all. The only analyticity we can hope for is reduction by definition to logical truth. (And if this were not so, Russell would not have struggled, without success, for seventy years to justify logic.)

[1]In "A Localist Metaphysical Semantics?" Matjaz Potrc takes localism to be the view that the meanings of words are constituted by some links to other words and hence to be inconsistent with atomism. This is not the molecular localism that is defined above and that Coming defends. That localism is strictly consistent with atomistic localism: it allows meanings to be constituted by inferential properties but does not require that they are. I expect that the meanings of many but not all words are constituted by inferential properties but do not argue for this.

Reflecting upon Yagisawa's attempts to establish that (B) is true solely in virtue of meaning, I think that I would have done better not to focus on the dependence of this sentence's truth on that of (U) in particular. Rather, I should have emphasized the following general points: first, the view that a statement can be true solely in virtue of meaning takes logic for granted; and, second, this is not appropriate (despite "What the Tortoise Said to Achilles") because logical truths are not true, and logical inferences are not valid, solely in virtue of meaning. Yagisawa's attempts do not count against these points, so far as I can see.

In virtue of what are logical truths true? A "substantialist" answer, in Miscevic's terminology, is that they are true partly in virtue of the way the world is. Miscevic claims that it is a mistake to <u>assume</u> this view in order to argue for the substantialist view of analytic truths, because that assumption is more radical and controversial than the view it is suppose to support (p. 5). He regrets the brevity of my discussion of this matter and clearly thinks that I come close to making the mistake.

Miscevic would like me to do something that I am not primarily concerned to do: argue for a view of analytic truths. In particular, I do not argue for a substantialist view of these truths. So I do not assume a substantialist view of logical truths in order to argue for this. My concern is rather with the view of analytic truths <u>that localism is committed to</u>. For, that commitment alone is alleged to be disastrous for the localism I am defending. I argue that localism alone is <u>not</u> committed to the view that analytic truths are true solely in virtue of meaning. For, their truth depends partly on logical truths and localism alone is not committed to those being true solely in virtue of meaning (p. 32). So my main point in the brief discussion that Miscevic is alluding to is the neutrality of localism on the matter of logical truths. (However, this is not the important point for the defense of localism: the important point is that localism is <u>epistemologically</u> neutral; sec. 2 below.) Miscevic is certainly right that I favor a substantialist view of them and recommend this view to the localist. But that is not my main point. That was one reason for the brevity he regrets. Another was that I did not have anything worthwhile to say on the subject.

However, Miscevic's interesting discussion stimulates me to go a little further. According to the substantialism that I recommend, (B) is made true partly by the worldly fact that all unmarrieds are unmarried. Miscevic claims that this fact is an instance of the general fact that "everything is what it is." From this he infers that my substantialism has the consequence that (B) does not depend on "anything concerning bachelors or unmarried men" but only on this fact "about the most general scaffolding of the world" (p. 6). He finds this unintuitive. I doubt that his inference is good. Furthermore, my substantialism goes along with the view that (B) is about bachelors and unmarrieds, as it appears to be. That view strikes me as much more intuitive than Miscevic's nonsubstantial alternative that (B) is about concepts and hence true solely in virtue of meaning.[2]

[2]Miscevic is influenced by his delightful "maloeba" example. But caution is appropriate because this example is a <u>fiction</u> and truths about fiction, particularly about impossible fictions, are known to require special treatment. In my view they involve fiction operators (1981a: 171-4).

2. The Argument for Holism from the Rejection of Analyticity (Levine, Yagisawa, Miscevic)

This argument aims to show that that there can be no principled basis for the molecular, or "moderate," localist's distinction between inferential properties that constitute a word's meaning and those that do not. Coming criticizes this argument at some length (1.5-1.13). Part of this criticism (1.11) is a reply to Levine's response (1993) to an earlier version of the criticism (1993). In this section I shall consider Levine's response, "Troubles with Moderate Localism," to Coming's reply. My aim in this ongoing discussion is to refute an argument for the no-principled-basis consideration and hence for holism. The localist also needs to show that there is a principled basis. I use my methodology to propose a basis. I shall consider Levine's criticism of this proposal in section 8.

The problem for localism that we are concerned with is alleged to come from Quine's arguments against the analytic/synthetic distinction ("a/s"). The parts of these arguments that Levine and I agree should bother us are aimed at various unacceptable epistemological views that have usually accompanied doctrines of analyticity: a priori knowledge, empirical unrevisability, and the like.

I am interested in defending the "inference version" of localism according to which a few of the inferential properties of a token may constitute its meaning. However, many are interested in a slightly different doctrine, the "belief version" of localism according to which a few of the beliefs associated with a token constitute its meaning. Levine thinks that "this distinction does some important work for Devitt" (p. 5). This is not so. I think that holist arguments against both versions are bad. I draw attention to the distinction in claiming that the appearance of commitment to a priori knowledge is much greater for the belief version than for the inference version (1.8-1.10). But perhaps this is wrong. Levine points out that "there is a perfectly legitimate sense of a priori that modifies `inference'" not knowledge (p. 5): he has in mind inferences that are not justified by experience. Perhaps the degree to which the inference version may appear to lead to a priori inferences, in this sense, is much the same as the degree to which the belief version may appear to lead to a priori knowledge. In any case, as Levine says, the naturalist should be equally opposed to both a prioris. I am. Whatever the appearance, I argue that the reality is that neither version of localism faces a problem with the a priori.

The crux of my argument is that localism is a purely semantic doctrine and so does not involve the a priori or any other unacceptable epistemological view. There is a saying, "Garbage in, garbage out!" Well I say, "Semantics in, semantics out!" How could you get a nasty epistemological conclusion without an epistemological premise? An important part of my argument is that the localist need not accept the Cartesianism view that linguistic/conceptual competence yields some "privileged access" to the truth about meanings. (Later I argue against Cartesianism; sec. 3 below.) For Cartesianism seemed to supply the epistemological premise: crudely, the implicit thought was that a/s + Cartesianism --> apriority. (I say "seemed" because even Cartesianism was not enough: knowledge of logic had to be assumed as well.) And so it was appropriate for Quine to attack a/s by attacking apriority.

Localism certainly has <u>some</u> consequences that might <u>appear</u> to be epistemologically objectionable. I argue at some length that these are not <u>in fact</u> so (pp. 30-6). Thus, suppose that a token's meaning **BACHELOR** is partly constituted by its inferential relations to ones that mean **UNMARRIED**. Then it cannot cease to be so related without ceasing to mean **BACHELOR**. But this unrevisability is of no more epistemological interest than the following. Suppose that a person's being a capitalist is partly constituted by her owning means of production. Then she cannot cease to own such means and still be a capitalist. The unrevisability in both cases is "metaphysical" not epistemic.

Of course, localism, unlike economics, is <u>about</u> things of epistemological interest: it is about inferences which are epistemic, and about meanings, which play a causal role in inferences. If we pursue this interest and wed localism to an epistemological premise, we will get an an epistemological consequence. If the premise is nasty - for example, Cartesianism - then the consequence will be nasty; if nice, then nice.

Levine's discussion of disconfirmation provides an ingenious example wedding localism to a <u>nice</u> epistemological premise. Suppose that tokens mean **CAT** partly in virtue of inferential links to ones that mean **ANIMAL**. (I go along with this example although I think it is implausible.) According to me, a token in Jane's belief box[3] that means **ALL CATS ARE ANIMALS**, and so is analytic, is nonetheless open to empirical discomfirmation. Levine wonders how this could work. His suggestion is that some experience prompts Jane to the following inference:

1. If cats are animals and Gracie is a cat, then Gracie is an animal.
2. Gracie is not an animal and Gracie is a cat.
3. So, not all cats are animals.

He uses the following <u>nice</u> epistemological premise to find fault with this suggestion: If this inference is to be valid, there must be no equivocation over the meaning of `cat'. Yet, according to my localism, the very act of accepting the conclusion changes the meaning of `cat'. (This is not quite right because on my preferred inference version Jane could accept 3 without changing her disposition to infer `x is an animal' from `x is a cat'. But let us not quibble.) So the inference is not valid (p. 8). I agree. <u>But what is nasty about that for localism?</u> Valid argument depends on not equivocating. If Levine's story goes through, Jane is equivocating and hence ruins the argument. So what? We need have no Cartesian access to meanings to prevent us equivocating. We have a conceptual competence - a cognitive skill - which mostly does prevent us equivocating. Perhaps it does not here and Levine's story goes through. (But perhaps it does here; see below on <u>psychological</u> impossibility.) So Jane is equivocating. When we wrongly identify two objects we often make a similar mistake; eg. my cases of Nana-Jemima and the two Liebknechts (pp. 227-8). A related example is provided by Saul Kripke's famous case of Peter: Peter has "contradictory" beliefs about

[3]At the end of his paper, Yagisawa stops being "overly charitable" [!] to chastise me for this metaphor. But it can be easily unpacked, as he himself shows, and I briefly indicate (p. 30). It is surely harmless.

Paderewski because he wrongly distinguishes Paderewski the musician from Paderewski the politician (pp. 229-40). In such cases a person may use valid rules of inference and so in that sense may be logical, but the good use of valid rules depends on getting meanings right in a way that the person does not.

Everyone would agree, of course, that Jane's analytic belief that all cats are animals is "revisable" in that it can be removed by violence or surgery. I claim that (setting psychological doubts aside) it can be removed as a result of the ordinary process of empirical disconfirmation. Because of the equivocation, the process that Levine describes is not rational in the normative sense, but it is rational in the descriptive sense, the sense in which we might say that the rational process underlying a person's behavior was the gambler's fallacy.[4] Given the presumed facts about meaning, Jane's belief is reducible to a logical truth. So, Jane could go through a normatively rational process of removing it only if her experiences called logic itself into question. Still, as I indicate (pp. 22, 25), molecular localism entails no view at all on the epistemic status of logic[5] and so can go along with the Quinean view that even logic can be called into question by experience. So Levine's example does not show any clash between this localism and Quinean epistemological purity.

Levine thinks that molecular semantics alone makes the empirical disconfirmation of analytic sentences impossible. My point is that it does not. This throws no light on the psychological likelihood of such a process. Perhaps the sort of equivocation involved makes the process psychologically impossible in the normal human. But, even if this is so, it is of no epistemological interest.[6]

3. Cartesianism and the A Priori (Miscevic, Bar-On, Jutronic)

Cartesianism about meanings is rife, yet almost entirely unargued. Why should we suppose that merely understanding the words `bachelor' and `unmarried' (merely having the concepts

---

[4]Cf Yagisawa: "According to Devitt, dropping of a token from the belief box accompanied by a reinterpretation of the token is a form of empirical disconfirmation" (p. 26). This ignores the following crucial words in the passage of mine that Yagisawa quotes: the dropping must be "in the face of empirical evidence" (p. 35). Empirical disconfirmation is a rational response to the evidence. Furthermore, in his subsequent criticism of this misrepresented view, he wrongly writes as if the reinterpretation in question were a conscious decision.

[5]So, my localist's defense against the analyticity argument does not depend on assuming that logic is empirical, contrary to what Miscevic seems to suggest (p. 6).

[6]In discussing the inference version, I make a similar point about the possibility of dropping an analytic belief without changing meaning (p. 29). Yagisawa is strangely baffled by this discussion (II.ii). My point is that although this dropping is metaphysically possible, it may not be psychologically so. But, even if it is not psychologically so, this is of no epistemological interest.

BACHELOR and **UNMARRIED**) yields propositional knowledge of their meanings, including the relations these meanings have to each other? Coming argues as follows:

> The meaning of a person's token is presumably constituted by relational properties of some sort: "internal" ones involving inferential relations among tokens and "external" ones involving certain direct causal relations to the world. Take one of those relations. Why suppose that, simply in virtue of the fact that her token has that relation, reflection must lead her to believe that it does? Even if reflection does, why suppose that, simply in virtue of the fact that the relation partly constitutes the meaning of her token, reflection must lead her to believe that it does? Most important of all, even if reflection did lead to these beliefs, why suppose that, simply in virtue of her competence, this process of belief formation justifies the beliefs, or gives them any special epistemic authority, and thus turns them into knowledge? Suppositions of this sort seem to be gratuitous. We need a plausible explanation of these allegedly nonempirical processes of belief formation and justification and some reasons for believing in them. (p. 53)

A competent speaker is of course privileged in her access to linguistic data, but we have no reason to suppose that she is privileged in her access to the truth about the data (pp. 75, 81).

The Cartesian view of competence is not only implausible, it is unnecessary: there is the more modest alternative that competence is an ability or skill, a piece of knowledge-how not knowledge-that (1.7-1.8). This alternative is central to my semantic methodology.

Miscevic rejects the alternative, insisting frequently that competence alone can yield knowledge of concepts. Apparently he, like many many others, finds this Cartesianism too obvious to need an argument.

Dorit Bar-On concludes her paper, "`Natural Semantic Facts - Between Eliminativism and Hyper-Realism," by wondering briefly how linguistic competence could be simply a skill (n. 39). The short answer is that it could be a complicated set of dispositions; for example the disposition to infer `x is unmarried' from `x is a bachelor' (p. 27), and the disposition to think thoughts about Gail on hearing sentences that include `Gail' (p. 167). There is no reason to suppose that such dispositions must involve linguistic propositional knowledge.

Prior to this, Bar-On nicely describes the ways in which linguistic facts seem to be dependent on us; thus it seems impossible that English speakers always form passives in the wrong way or always misunderstand the word `chair'. This may encourage some to think that linguistic facts depend on speakers' judgments. I prefer another view, as she notes: the facts depend not on what we judge but on what we do. And the facts concern linguistic symbols not linguistic competence as supposed by Chomskian linguistics. She is worried that this "hyper-realism" may not have the resources to "stay clear of Quinean eliminativism" (p. 19). Eliminativism is certainly a threat and I do not pretend to have all the answers to it (5.3). In attempting to deal with it, however, I have more resources than Bar-On supposes, for she takes me to divorce linguistic facts "from the psychology of speakers" (p. 19). Nothing could be further from the truth. I give priority to the language of thought and suggest a basically Gricean explanation of the "public" linguistic facts in terms of thought (4.4). I am as mentalistic as can be. But I reject the identification of the linguistic task with the study of linguistic competence, and I reject the view that this mental state consists in knowledge-that.

Dunja Jutronic, in "Knowledge of Meaning and Knowledge of the World," argues against my view of knowledge of meanings. In assessing this argument it is important to distinguish three sorts of knowledge: (a) linguistic competence; (b) folk linguistic intuitions; (c) the scientific theory of language. My view of (a) is anti-Cartesian: the competence is a piece of knowledge-how not knowledge-that. Jutronic clearly disagrees, but her argument does not bear on this view. Her argument bears rather on views of (b), perhaps even of (c). In her discussion of folk identifications of meanings, she wrongly takes me to have a have a knowledge-how view of (b) also. Against this she insists that when the folk identify tokens that mean **CAT**, applying a t-clause of the form `that ...cat...' to them, the folk are evincing knowledge of the tokens, items of knowledge-that. Jutronic is pushing on an open door here. What the folk evince in these identifications are their most basic linguistic intuitions. These intuitions, and any richer ones that the folk may have about the tokens, are indeed propositional, just like any other intuitions (2.10 and 2.11). What I deny, and she maintains, is not this but rather that these intuitions are "the voice of competence." I argue that they are ordinary empirical responses to linguistic phenomena.

Bar-On is concerned with a doctrine, (SV), which claims that semantic facts "depend on speakers' judgments." That I certainly reject. But (SV) also claims that semantic facts "cannot be beyond the epistemic capacities of the speakers of the language" (p. 5). I doubt the "cannot" but I agree that the semantic facts are not beyond the epistemic capacities of speakers. (I am optimistic enough to doubt that any facts are beyond those capacities!) Indeed, I think that (b), folk linguistic intuitions, are pieces of protoscience that usually capture some of those facts. So, contrary to what Jutronic supposes (p. 2), I do not think that speakers are totally ignorant of meanings. Furthermore, I think that the speakers engaged in (c), the scientific theory of language, are slowly but surely capturing more facts about meanings.

In sum, what I maintain is not that linguistic facts are inaccessible to speakers but that linguistic competence does not give speakers any privileged access to those facts. (a) is knowledge-how. (b) and (c) are knowledge-that. (a) alone does not yield (b) or (c).

I argue that Cartesianism alone is insufficient for a priori knowledge of worldly facts like that all bachelors are unmarried: knowledge of logic is also required. Miscevic thinks that Cartesianism alone is sufficient for a priori knowledge of conceptual facts like that **UNMARRIED** is included in **BACHELOR**. I did not say this, but I should have. Miscevic endorses this a priori knowledge, claiming that it is unobjectionable because it is not knowledge of substantial worldly facts. It seems objectionable enough to me. In any case it is contrary to naturalism as I define it, as Bar-On points out.

4. Naturalism and the A Priori (Rey, Miscevic)

Naturalism is also central to my semantic methodology. The naturalism in question is an epistemological doctrine: that there is only one way of knowing, the empirical way that is the basis of science (whatever that way may be). So I reject a priori knowledge. I do not give a detailed argument for my rejection but I do give two reasons (2.2): Briefly, first, with the recognition of the holistic nature of confirmation, we lack a strong motivation for thinking that mathematics and logic are immune from empirical revision; and, second, the idea of a priori knowledge is deeply obscure, as the history of failed attempts to explain it show.[7]

Georges Rey has a different view of the a priori, presented first in "The Unavailability of What We Mean" (1993) and now in "Devitt's Naturalism: A Priori Resistence to the A Priori?". He rightly insists "that whether or not there is a priori knowledge is an empirical issue" (p. 1); on my view, <u>every</u> issue is an empirical one.[8] But he thinks that this issue is open and may well be settled <u>in favor</u> of a priori knowledge. And he is pleased with Levine's conclusion that my molecularism commits me to the a priori, the conclusion I have just rejected. In response to my first reason against the a priori - the lack of motivation - Rey is rather scornful. In response to my second reason - the obscurity - he appeals to reliablist theories of knowledge and claims to give at least a sketch of how we might have a priori knowledge of logic, mathematics, and analytic truths. He finds my comments on this sketch in <u>Coming</u> "bewildering" (p. 6) and "disconcerting" (p. 7). I shall discuss his responses in turn. But I start with some preliminary points.

(1) I am claiming that knowledge can be <u>justified</u> only by experience, that the <u>evidence</u> for it must be experiential. I doubt that there is any innate knowledge and so am inclined to think that experience must also be part of the <u>source</u> of a person's knowledge. But that is another matter and not my concern in <u>Coming</u>. Rey sometimes writes as if he thinks some knowledge may be innate. Suppose that some is. My second reason against the a priori still applies. If what is innate is indeed knowledge, it must be justified. We have <u>some</u> idea how we <u>might</u> establish the justification for innate empirical knowledge: experiences of the worldly facts that are the subject of the beliefs might play a role via adaptation in the production of the innate beliefs. But we have no idea how we might establish the justification for innate a priori knowledge.

(2) Although obviously not a naturalist in my sense, Rey claims to be one in some sense. It is not as clear as it should be in what sense. Two sharply different doctrines are often called "naturalism," one metaphysical and the other epistemological. Metaphysical naturalism is <u>physicalism</u>: the view, roughly, that all entities are physical entities and that the laws they obey are in some way dependent on physical laws. This is a reductive doctrine. It has <u>nothing</u> to say about ways of knowing except that they must be, like everything else, physicalistically acceptable: so it alone entails nothing one way or the other about a priori knowledge. Rey endorses this doctrine. Quine and I do too, but we call this doctrine "physicalism" not "naturalism." What we mean by "naturalism" is an epistemological doctrine that is <u>not</u> reductive and <u>is</u> opposed to a priori knowledge.

[7] So I do not simply <u>assume</u> that logic is empirical, as Miscevic implies (p. 6), I argue for it.

[8] So the answer to the question in Rey's title is: "No!"

Quine is expressing this doctrine in rejecting first philosophy and insisting that reality must be examined scientifically.[9] It is what he aims to capture in his vivid metaphor of the seamless web and I aim to capture in my claim that the only way of knowing is the empirical way. Epistemological naturalism applies to all knowledge. So it applies to knowledge of ways of knowing themselves, to epistemology. So Rey is right to claim that "it can be pressed further to embrace...[what] Quine calls `naturalized epistemology'" (p. 13). Or, more accurately, he would be right if it were clear that he was talking of epistemological and not metaphysical naturalism. However, Rey is misleading, at least, in implying that he is <u>endorsing</u> this application of epistemological naturalism. Rey cannot endorse this because he is arguing that epistemological naturalism may well be false! He is arguing that there may well <u>be</u> a priori knowledge and first philosophy. What he really endorses is the use of the empirical method to argue for the a priori. Now if one must argue this - and one mustn't and shouldn't - it is certainly better to do so empirically rather than a priori: better not to start in sin even if one ends up there. But using the empirical method from time to time doesn't make you an epistemological naturalist, else everyone would be one. What makes you an epistemological naturalist is a commitment to there being <u>no other</u> method.

(3) Rey's discussion of Quine's attack on a/s and the a priori/empirical distinction is subtle and highly illuminating. Yet in one respect it strikes me as rather obtuse: Rey trivializes Quine's revisability thesis by characterizing it as follows: "any belief can be reasonably revised in the light of experience."[10] He finds this indistinguishable from "banal fallibalism" (1993: 72): "people could be wrong about anything; they can make errors in reasoning, rely on experts that mislead them, or just reason themselves into strange corners" (p. 2). For example, Rey, in balancing his checkbook concludes, "16 + 17 + 18 + 23 + 100 = 174" but abandons this arithmetic truth on being told by the bank that he is wrong. He reasons that surely the bank is better at addition that he is (1993: 70-1). Rey points out that this fallibalism "<u>has nothing to do with the a priori</u>," for surely no rationalist ever denied that you could make errors in your a priori reasoning (p. 2). True enough, but this fallibalism also <u>has nothing to do with Quine's thesis</u>, for surely Quine and the empiricists were aware that the rationalists accepted this fallibalism. Banal fallibalism <u>can't</u> be the right way to understand Quine and the dispute over the a priori.

An empiricist ought to accept a distinction between two ways that further experiential evidence can and should lead a person to change her mind about a statement <u>p</u>. (i) On the one hand,

---

[9]See, for example, Quine 1981: 21, 67, 72. The distinction between his naturalism and his physicalism is implicit in a passage on p. 85. (In response to p. 72, Rey draws a red herring "about whether empirical science as a whole can be justified by some means external to it" (n. 1). Quine is pointing out that if epistemological naturalism is true then there is no such justification. Rey responds by rejecting the reverse conditional: that if epistemological naturalism is not true - if there is a priori knowledge - then there is such a justification. This is indeed "a quite different issue" but it is not one Quine has raised.)

[10]My thoughts on this matter were sharpened in extensive correspondence with Rey prior to his 1993.

the evidence might bear for or against p itself. (ii) On the other hand, the evidence might throw light on the goodness of her thinking about p. She is a fallible calculator. New evidence may throw no direct light on p but may suggest that she has made a mistake in her thinking about the relation of the evidence to p. Quine's revisability thesis is surely concerned only with (i): "no statement is immune to revision" (Quine 1953: 43) in that experiential evidence might directly bear against it. The thesis is simply concerned with the relation between evidence and statement not with the relation between evidence and the view that a particular person (or even a particular community) has thought well about the statement. In fact, the thesis is the epistemological naturalism that I have just described, captured by the metaphor of the seamless web. Understood in this way the thesis strikes at the heart of apriorism and is far from trivial.[11]

This having been said, Rey's discussion raises an interesting possibility. It is common to confront the apriorist with historical examples of allegedly a priori knowledge abandoned in the face of experience; for example, the Euclidean view of space. In the light of the above distinction, it is always open to the apriorist to respond that this experience was relevant in way (ii). The evidence does not count for or against the statement in question, it simply shows that our process of nonempirical justification was defective in this case. Of course, it remains to be argued that this is a plausible response in a particular case. In my view, it is rather clearly not in the case of Euclidean geometry.

In any case, the argument against apriorism and for the seamless web that we should take from Quine does not rest primarily on these historical examples. It rests primarily on the two reasons I gave. The first of these involves confirmation holism, but not quite in the way Rey seems to think (pp. 2-6; 1993: 78-81). The argument starts by pointing out how scientific laws that are uncontroversially empirical are holistically confirmed. Evidence for this is not to be found only in the discussions of Duhem and Quine: most of the evidence comes from the movement in the philosophy of science inspired by Thomas Kuhn and Paul Feyerabend. It is then plausible to extend this holism to all beliefs, even those of logic and mathematics; there is no motivation for a seam in the web.

I turn now to Rey's response to this reason. Rey has a lot of rhetorical fun mocking the remarks that Quine and I make about the empirical way of knowing and about the application of this way to logic and mathematics. He thinks that I am under "the illusion" that these remarks amount to "a serious theory" (pp. 4-5). I am not. I agree with Rey that "no one yet has an adequate theory of

[11]I made this criticism briefly before (1993: 53n). In response to it, and to criticisms by others, Rey discusses the following view: the Quinean "point is that someone could hold on or revise any statement and still be rational" (Rey 1993: 71). But this is not my point. Rey is right that the apriorist should accept that it may be rational for someone to stop holding any belief in the face of evidence of type (ii); for example, Rey at the bank. The Quinean point is that the content of the belief can be objectively disconfirmed by recalcitrant experience. The distinction between (i) and (ii) is roughly the same as the one Hartry Field draws between "primary" and "secondary" sorts of evidence in response to a point similar to Rey's (1996: 4).

our knowledge of much of anything" (pp. 3-4); as I say, "we do not have the rich details of the empirical way of knowing that we should like to have" (p. 50). In any case, Rey's mockery is largely beside the point. Since we do not have a serious theory that covers even the easiest examples of empirical knowledge, the fact that we do not have one that covers the really difficult examples from logic and mathematics hardly reflects on the claim that these are empirical knowledge too. We all agree that there is an empirical way of knowing. Beyond that, this part of the argument against the a priori needs only the claim that the empirical way is holistic. We have no reason to believe that a serious theory would show that, whereas empirical scientific laws are confirmed in the holistic empirical way, the laws of logic and mathematics are not; that it would show there is a principled basis for drawing a line between what can be known this way and what cannot.

I would be the first to concede that this part of the argument alone is far from conclusive, a long way from proving that the holism extends to logic and mathematics. That is why I put a lot of weight on the rest of the argument: my second reason, which is about the obscurity of a priori knowledge. In this part, we attempt to show that the alternative a priori explanation of our knowledge of logic and mathematics, indeed of anything, is very unpromising. If this is right, we have a nice abduction: the best explanation of that knowledge is that it is empirical.

The a priori way of knowing is typically characterized by what it is not: it is not empirical. But what we need if we are to take the a priori way seriously is some idea of what it is. We need a positive account, not just a negative one. Why? This question may seem particularly pressing since I have just agreed that we do not have a serious theory of the empirical way. However, there are two crucial differences in the epistemic status of the two ways. First, the existence of the empirical way is not in question: everyone believes in it and Rey is even urging us to use it to show that there is an a priori way. In contrast, the existence of the a priori way is very much in question. Second, even though we do not have a serious theory of the empirical way, we do have an intuitively clear and appealing general idea of this way, of "learning from experience." In contrast, we do not have the beginnings of an idea of what the a priori way might be; we lack not just a serious theory but any account at all.

Rey claims to provide just such an account, appealing to the idea that knowledge is true belief arrived at by a reliable process. He is pained by my unenthusiastic response. This response is certainly very brief (p. 51n; it draws on my 1993). I shall expand it here.

The objection to a priori knowledge is that we don't know what it would be for something to be known a priori. So a successful resurrection of a priori knowledge must describe a nonempirical way of knowing, a process for justifying a belief that does not give experience the role indicated above. The difficulty in meeting this demand is well-demonstrated by the failure of traditional attempts: on the one hand, these assumed that we have Cartesian access to meanings; on the other hand, they took knowledge of logic for granted. The trouble with Rey's ingenious proposal is that it does not meet the demand either. He has seriously underestimated what is required for the resurrection.

Rey proposes that we know logical truths a priori because they are produced by a sub-system of the brain that enjoys a reliability "completely independently of whatever input (i.e. experience) an agent may receive" (1993: 91); for example, Ellen realizes in her brain "a non-axiomatic system of natural deduction, relying entirely on the operation of standard rules like modus ponens, universal generalization, conditionalization. etc." An example of its output is:

(R) Nothing bites all and only those things that don't bite themselves. (p. 6).

This "would be knowledge because each of the rules she used were surely justified if anything is. As soundness proofs of first-order logic show, they are, indeed, absolutely reliable in this sense: it is impossible for them to produce a falsehood as a theorem" (p. 7). The system's insensitivity to sensory input is intended to show that this is not an account of the empirical way of knowing. My objection is that it is not an account of a way of knowing at all.

It has been clear, at least since Plato, that for a belief to count as knowledge, it must not only be true, it must be justified: there has to be something about the way it was produced that makes it epistemically nonaccidental. Thus, if I believe on no good basis that it will rain tomorrow, or that a certain number is a prime, and I should turn out "by accident" to be right, my belief does not count as knowledge. The problem with Rey's account is that it does not show how Ellen's belief in (R) is, in the appropriate way, epistemically nonaccidental.

Consider another belief of Ellen's: the clearly empirical belief, (E), that Tom is mortal. Suppose that Ellen knows that Tom is a man. Suppose further that (E) is the product of a sub-system that took Ellen's knowledge that Tom is a man as input and that always yields the output that x is mortal given the input that x is a man. Inspired by Rey, we then argue: "(E) is knowledge because the rule she used was surely justified if anything is. As a vast amount of empirical research has shown, this rule is reliable in this sense: given a truth of the form `x is a man' as input it will always produce a truth as output. And the rule took a known truth as input." Clearly something is wrong with this argument. More needs to be said to show that the production of truths like (E) by "the mortality sub-system" is epistemically nonaccidental. To help see this, suppose that the sub-system in Ellen was itself produced by a "random" process that also produced other sub-systems which frequently yield false outputs; for example, given the same known input, yield the output that Tom is wise, round, brave, etc.; or, given other inputs such as that x is a bachelor, yield the output that x is rich. Among these many sub-systems, one just happens to reliably produce truths, the mortality sub-system. We need to say more to rule out that the sub-system is thus accidental. We might say something about how the sub-system was produced: that experiences of a mortal-man world played an essential role in that production. But it is not necessary to say this.[12] It is necessary to show that the sub-system, however it was produced, is causally sensitive in an appropriate way to the fact that we live in a mortal-man world. The output of Ellen's sub-systems in the imagined situation are not knowledge because those sub-systems are impervious to the way the world is; impervious to whether it is a mortal-man, wise-man, round-man, brave-man, rich-bachelor world.

[12]In my 1993 I wrongly insisted that it was necessary (p. 55).

The problem with Rey's proposal can be put briefly: he does not say more. What he says about his logical sub-system is analogous to my Rey-inspired argument about the mortality sub-system. We have just seen that that argument fails. Rey's argument fails for the analogous reason. He needs to show that the production of truths like (R) is epistemically nonaccidental. Once again, we can bring out the problem by supposing that the logical sub-system in Ellen was produced by a "random" process that also produced other similar sub-systems which frequently yield false outputs. Perhaps these sub-systems include one that realizes the gambler's fallacy; one that fails to take proper account of the base in probabilistic reasoning; one that commits the fallacy of asserting the consequent. Rey needs to say something to rule out that it is a mere accident that one of Ellen's sub-systems yields truths. And, of course, he cannot say that it yields truths because it is causally sensitive in an appropriate way to the fact that we live in a logical world. For then the knowledge of (R) would be empirical.

It does no good to insist that the logical sub-system is reliable in that it regularly produces truths, for that is true also of the mortality sub-system which, as we have just seen, may not be producing knowledge at all. Being reliable in this respect that Rey emphasizes is simply not enough. It does no good to appeal to proof theory to demonstrate how sure we theorists are of the reliability, in that respect, of the logical sub-system. For the reliability of the sub-system in that respect is not in question; and, in any case, we theorists are also sure of the reliability of the mortality sub-system in that respect. We need to know something more about Ellen to establish that she arrived at (R) by a procedure that is, in a broader respect, epistemically reliable.

In saying this I am not insisting that Ellen can only know (R) if she knows that she knows - the KK principle - or if she has available to herself some justification of the rules of the logical sub-system (cf. Rey: 8; 1993: 92), nor am I simply refusing to accept a reliablist approach to knowledge.[13] I think that the reliablist idea for empirical knowledge, briefly indicated above, may well be along the right lines: Ellen knows (E) if the mortality sub-system that produced it is causally sensitive in an appropriate way to the fact that we live in a mortal-man world. This can be the case without Ellen knowing that it is, let alone knowing that it must be the case for her to know (E). I am insisting that there be a justification for Ellen's believing (E), not that she knows the justification. Putting this reliablist idea together with my naturalism, I think that it may well be the case that Ellen knows (R) because it was produced by a logical sub-system that is causally sensitive in an appropriate way to the fact that we live in a logical world. (We do not know the details, of course, but this does not count against (R) being empirical because we are hardly better off with (E).)

There is a difference between the logical and the mortality sub-systems which seems to play a role in Rey's thinking: the former unlike the latter yields necessary truths.

[13]Indeed, for all I know, my view may be consistent with the particular reliablist theories that Rey cites as the basis for his view (1993: 91n).

If empirical knowledge can be the result of a process that reliably...issues in true beliefs in relevant circumstances, why couldn't a priori knowledge be the result of a process that reliably issues in true beliefs in all possible circumstances? (p. 7; see also 1993: 92, 95)

The answer is: because that metaphysical difference in the output is not epistemically relevant. In the case of the mortality sub-system, we have seen that we have to say more than that it is reliable in Rey's respect if the contingent (E) that is its output is to count as knowledge. It would be strange indeed if we were relieved of the responsibility of saying more in the case of the logical sub-system by the fact that its output (R) is necessary; so the requirements on knowing a necessary truth would be less demanding! The fact that a statement is necessarily true can no more show that the process of arriving at it is epistemically nonaccidental than can the fact that it is true. That fact does not undermine my argument.[14]

Aside. How might the necessity of the output of the logical sub-system be accommodated by a reliablist account of empirical knowledge of the output? What could the appropriate causal sensitivity to a logical world amount to? We are tempted to say that Ellen's mortality sub-system is sensitive to a mortal-man world in that if the world were different she would not have that system. If we then apply this approach to Ellen's logical sub-system, we seem to have a problem: the world cannot be different from the logical world; it is necessarily logical. So, it is common to think, the logic of subjunctive conditionals makes it trivially the case that if the world were nonlogical she would not have had that logical sub-system. I doubt this view of subjuctive conditionals.[15] In any case, perhaps we were wrong to be tempted by this approach. It builds a certain sort of infallibility into knowledge: Ellen knows (E) only if she could not have the mortality sub-system unless she lived in a mortal-man world. This seems to overlook the message of Descartes' First Meditation and the underdetermination of theories by the evidence. Perhaps what we need to say is that Ellen would not have had the mortality sub-system unless her world appeared to be a mortal-man world, unless she had experiences appropriate to such a world. She might have had other experiences, perhaps caused by an Evil Demon, even though she does live in such a world. Similarly, perhaps we should say that she would not have had the logical sub-system unless she had experiences appropriate to living in a logical world. She might have had other experiences, perhaps caused by an Evil Demon, even though the world she lives in is necessarily logical.

The point that we need to say more than Rey does for (R) to be knowledge may be more obvious if we change the examples a little: replace talk of sub-systems with talk of general beliefs. Suppose that (E) was not produced by the mortality sub-system but is inferred from the general belief that all men are mortal. Clearly, (E) will count as knowledge only if the general belief does. So we

[14]Rey himself notes a sign of the irrelevance of this fact (p. 7). A person might have sub-systems that yield necessary truths, whatever the input, and yet those truths are indubitably empirical; for example, the truths that water is H2O and that Hesperus is Phosphorus. (I don't follow Rey's brief response to this problem.)

[15]As does Hartry Field (1996: 17).

need to say more to show that general belief does. Now suppose that (R) was not produced by a non-axiomatic system of natural deduction but is inferred from some general logical beliefs. Once again, the epistemic status of (R) depends on the status of the general beliefs. So we have to say more to show that they are knowledge. It is hard to see how the change from general beliefs to sub-systems of rules could remove the need to say more.

I wonder if the talk of reliability is confusing the issue. So, consider what Rey tells us about (R) that might constitute its <u>justification</u>. First, of course, it <u>is</u> true. Second, it is the product of a sub-system that regularly produces such truths. Obviously, we have no justification so far; think of the mortality sub-system. Third, the logical sub-system is insensitive to sensory input. But this insensitivity clearly cannot justify the system's output.

Note that it would do Rey no good to suppose that his logical sub-system is innate. What <u>nonempirical</u> story could possibly be told of its presence that would support the view that its output is knowledge? The evidence suggests that some "good" and some "bad" logical sub-systems may be innate. What could justify the "good" ones apart from some empirical story?

In sum, the comparison with the mortality sub-system shows that the respect in which, according to Rey, his logical sub-system is reliable is insufficient to establish that its output is knowledge. To characterize a way of knowing he needs to say more, showing that the system is, in a broader respect, epistemically reliable. There is no reason to suppose that if he were to say more he would characterize a way of knowing different from the empirical way.

<u>5. Methodology and "The Canberra Plan" (Bigelow)</u>

For the rest of my argument, including my argument for molecular localism, I need a semantic methodology. I start with the question: What are the semantic tasks? What should semantics be trying to do?

There seems to be a simple answer: the "basic" semantic task is to say what meanings <u>are</u>, to explain their <u>natures</u>. But there is a problem: it is far from clear what counts as a meaning that needs explaining (2.3). So I set about saying what should count by focusing on the purposes for which we ascribe meanings (or contents) using t-clauses (`that' clauses) in attitude ascriptions: in particular, the purposes of explaining intentional behavior and of using thoughts and utterances as guides to reality. I call these purposes "semantic." I say further that a property plays a "semantic" role if and only if it is a property of the sort specified by t-clauses, and, if it were the case that a token thought had the property, it would be in virtue of this fact that the token can explain the behavior of the thinker or be used as a guide to reality. We are then in the position to add the following explication to the statement of the basic task: A property is a meaning if and only if it plays a semantic role in that sense. And the basic task is to explain the nature of meanings in that sense (2.4-2.6).

John Bigelow's response to this, "Devitt's Double Standard," is provocative. My question about

> the theoretical utility of attributions of content, if taken in the way I believe that he does himself often take it, is the first step in an investigation conducted according to the Canberra Plan (pp. 5-6).

As a result I am alleged to traffic in "a priori analyticities" and "a `language-first' methodology." If I did, that would indeed be ironic because my methodology is built around opposition to these. Bigelow thinks that my heart is really in the right naturalistic place but that I keep lapsing into the sin of First Philosophy: "Devitt's real project is an innocent one, free of Canberra Plan. But too often he mis-speaks himself." He finds these lapses understandable given the company I keep and so urges charity. Indeed, he wants to help me with my marketing (p. 3). What a wag!

Where does Bigelow think that I have gone wrong? (i) He is unconvinced by my reasons for thinking that it is important to answer "the utility-question." (ii) He thinks that the uses I make of my answer to the question, uncharitably construed, follow the Canberra Plan: "The Canberra Plan runs exactly parallel to the way in which Devitt proceeds from his utility-question to his theory of content" (p. 7). This construal is not simply uncharitable, it is <u>perverse</u>. Yet, Bigelow tells me with glee, the construal is gaining popularity. I am naturally concerned to put a stop to this. So I wish that Bigelow had said more to show <u>how</u> my procedure can be seen as parallel to the Canberra Plan. I remain largely in the dark about this. I shall discuss the two points in order.

(i) Bigelow mentions some sciences where the utility question does not arise. Why then should we worry about it in semantics? I think that semantics is different, as he notes. I claim that the basic semantic task of explaining what meanings <u>are</u> is

> analogous to such tasks as saying what genes, atoms, acids, echidnas, or pains <u>are</u> but not, we should note, to such tasks as saying what genes and so forth <u>do</u>, stating the laws that advert to them. (The difference is illustrated by the difference between molecular and Mendelian genetics.) However, we start the semantic task in rather worse shape than we do its analogues. With them, the subject matter of investigation is already identified relatively uncontroversially. This reflects the fact that we have clear and familiar theoretical or practical purposes <u>for which</u> we identify the subject matter. Semantics does not start out like that. It is far from clear what counts as a meaning that needs explaining. Indeed, the intractable nature of semantic disputes largely stems from differing opinions about what counts. (pp. 54-5)

I cite evidence that others have trouble with talk of "meaning." Bigelow does not say what he finds unconvincing about all this, but the following claim makes me wonder if he has misunderstood:

> I don't accept that, as contrasted with other kinds of theorising, semantics is deeply problematic in such a way as to require guidance from metatheoretic principles about what theoretical purposes are served by the activity of semantic theorising. (p. 5)

The utility-question in not about the purposes of <u>semantic theorizing</u>, it is about the purposes of <u>ascribing meanings</u>. I argue that this question should not only be kept in mind in semantic theorizing, it should <u>precede</u> that theorizing: "We start semantics in the unusual position of having to specify a subject matter" (p. 55). Once a consideration of our purposes has enabled us to <u>identify the</u>

subject matter that should concern semantic theorizing, the purpose of that theorizing is, like that of our theorizing about the nature of genes, too obvious to be worth mentioning.

Any token thought or utterance has indefinitely many properties, constituted by its relations to the world and to other tokens. We might study the natures of any of these properties. Yet, clearly, in semantics we should be concerned only with a special set of these properties. Which set? What is the principled basis for choosing it? These questions are insufficiently addressed. The implicit or explicit answers people give to them are often different, leading to cross purposes.

I confront these questions in arguing against holism, eliminativism and revisionism, and in discussing direct-reference, two-factor, and verificationist theories. Here is one illustration of the importance of this confrontation. Direct-reference philosophers maintain that there is nothing more to the meaning of a name than its role of referring to its bearer. They claim that the difference between `a = a' and `a = b' is not a difference of meaning but rather a pragmatic or cognitive difference that is irrelevant to semantics. This strikes many as implausible. Surely the fact that `a' and `b' differ in their modes of reference - in my view, they differ in their causal modes - is a difference of meaning and hence so is the difference between the two identity statements. At this point the disagreement reduces to an unsatisfactory clash of intuitions about what counts as a difference of meaning. How can we move beyond this "intuition mongering"? I claim that there is only one way: we must consider our purposes in ascribing meanings. I focus on two purposes: explaining people's behavior and using others as guides to reality. I argue that those purposes are served by ascribing modes of reference (4.2-4.3). "The very same considerations that lead us to ascribe meanings at all lead us to ascribe modes" (p. 183). Of course, one might respond to this with some other view of our semantic purposes. But direct-reference philosophers never confront this methodological issue at all. As a result, their denial that the different modes of reference of `a' and `b' constitute a difference in meaning "is theoretically arbitrary and ad hoc, nothing but a verbal maneuver" (p. 186). (See sec. 11 below for more on direct reference.)

Notice how powerless Bigelow's rival strategy is to deal with this disagreement. He urges that we "ostensively identify examples of content by appeal to particular cases" (p. 8). The ostensions of direct-reference philosophers identify the same content (meaning) in the two identity statements; the ostensions of their opponents, different contents (meanings). What does Bigelow do then?

Bigelow (pp. 9-10) likes the following points that I make against the holist: the localistic properties that we ordinarily ascribe do sometimes explain behavior and guide us to reality; holistic properties could not do this. He wants me to leave it at that: no need to tie our ascriptions of meanings to certain theoretical purposes. But then how does he deal with holists who agree with those points but insist nevertheless that meanings are holistic? I deal with them by saying that ascribing the localistic properties that I am calling "meanings" serves serious theoretical purposes whereas ascribing the holistic properties that they are calling "meanings" could serve no serious purpose at all (pp. 124-6).

Finally, this issue about the purposes of ascribing meanings should not be confused with people's motives in ascribing them. Bigelow is surely speaking for us all in saying: "In some contexts I have no interest in the theoretical utility of attributions of content" (p. 5). People may make the attributions because they are curious, bored, showing off, or whatever. This is beside the point. The issue that I am raising concerns the theoretically interesting purposes for which we ascribe meaning. For, the properties that serve such purposes have natures worthy of study in semantics. Very often, of course, people ascribe meanings to serve these theoretically interesting purposes. It is of no significance that they may often have other motives.

(ii) On the face of it, my procedure is nothing like the Canberra Plan. That Plan, as described by Bigelow (p. 6), would call for the following approach to the semantic task. We should start by asking about the meaning [!] of semantic words like `mean', `refer', and `true'. We should answer by constructing a Ramsey sentence from the conjunction of all the important assertions involving those words. The Plan would then take the words to "refer to whatever, if anything, will make the Ramsey sentence true if taken as the values of the variables which have replaced the words you were interested in" (p. 6). Coming has no discussion that comes close to this; indeed there is hardly any discussion of the meaning of semantic words at all. Rather, Coming follows the procedure recommended by my methodology.

We start by settling on the "basic" task of explaining the natures of properties that play a certain role in explaining behavior and guiding us to reality. Our ordinary thought ascriptions ascribe properties for those purposes. We examine the nature of those properties. That is the "descriptive" task. We need to know about not only the properties that we do ascribe but also about the properties that we ought to ascribe for those purposes: we need to accomplish the "normative" task. For, the properties we ought to ascribe are the ones that concern the basic task. The descriptive task provides the main evidence for the normative/basic task: given the apparent success of our ordinary thought ascriptions, it is likely that the properties they ascribe really do explain behavior and guide us to reality.

How could this procedure be read uncharitably as the Canberra Plan? The main clue Bigelow provides is in his talk of "stipulation." He thinks that my start - settling on a task - can be seen as a stipulation of meaning differing only trivially from the Plan's start - a definition of meaning using a Ramsey sentence. Thus I can be seen as "trafficing in a priori analyticities" and "a `language-first' methodology" (p. 3).

I am mystified. Everything is similar in some respects to everything else and so doubtless there are parallels between my methodology and the Canberra Plan. But Bigelow needs to say a lot more about what these parallels are and why they have the significance he supposes. Meanwhile, here are a few words in response.

(a) My start is a "stipulation" only in the most gentle sense. This is brought out in the paragraph that ends my discussion of "other views of the semantic tasks" (2.7):

> In conclusion, I emphasize that I am not claiming that the properties that I have called "meanings" are the only real meanings nor that the tasks I have called "semantic" are the only proper tasks for semantics. I doubt that there is any interesting matter of fact about such claims. I do claim that those properties are worth investigating and that those tasks are worth performing. I claim further that those tasks are appropriately fundamental. Perhaps all this is true of other tasks, but that always needs to be demonstrated. (p. 67)

Thus my criticism of direct-reference philosophers is not merely that my view of the tasks shows them to be wrong; it is also that they offer no rival view of the tasks that shows them to be right. And my criticism of holism does not depend on my linking of the tasks to the purposes of explaining behavior and guiding us to reality: any view of the purposes would do (p. 92).

(b) The quoted paragraph also brings out that what I am primarily "stipulating" is a task: the study of certain properties. Now it is true that in calling those properties "meanings" I can be seen as "stipulating" a meaning of `meaning' (but, note, there is no "stipulation" about `refer', `true', etc.). But this manifestly is not a piece of language-first philosophy. (c) I do not claim that this "stipulation" yields an analyticity, although perhaps it does. (d) I certainly deny that it has the status of a priori knowledge that the Plan claims for its definition. (e) Indeed the differences between the status and role of my "stipulation" and the Plan's definition are striking. The role that the Plan gives to the definition takes the subject matter of semantics to be whatever fits ordinary assertions using semantic words like `mean'. My discussion of the semantic tasks starts from the assumption that the ordinary use of `mean' is too vague to select a suitable subject matter. I select one by looking for properties that are theoretically interesting. For convenience, and out of respect for tradition, I call the selected properties "meanings." But nothing hinges on what we call them: they are worth investigating whatever they are called. The semantic program I urge would be unchanged if the properties it explains were never called "meanings."

6. Methodology and Consensus (Slater)

I emphasize that if semantics is to be a science it must be concerned with properties worthy of study (pp. 55, 63). Carol Slater, in "Semantics as Immature Science," rightly points out that the properties I identify as meanings, indeed the properties anyone identifies as meanings, are worthy of study only if they are instantiated. If there were angels on the point of a pin, or anywhere else, they would doubtless be worthy of study. But eliminativism is right here and so angels are not worthy of study. If eliminativism were right about meanings, if nothing played that role in explaining behavior and guiding us to the world, then meanings would not be worthy of study.

I think that there is a lack of consensus about the subject matter of semantics, partly concealed by the common use of the vague term `meaning'. My suggestion for achieving consensus is to consider what purposes are served by ascribing meanings. Slater remarks that "consensus alone does not make a domain apt for scientific study." (p. 4) In my view, consensus has nothing to do with making it apt for scientific study. This aptness is just the aforementioned worthiness. What makes meanings worthy of study, on my view of them, is the roles they play in explaining behavior and guiding us to reality. The first role is a straightforwardly scientific one in psychology. The second is unusual, even unique: meanings guide us to reality not only in science but throughout ordinary life. Properties that have these two roles, and indeed others, are clearly worthy of study. Slater rightly points out that there is no guarantee that the one sort of property will play both roles and so form one "natural kind." I argue, however, that the one sort of property - a "Representational" property - does in fact play both roles (4.3).

Finally, I do <u>not</u> think that "a communally shared research framework is necessary for the conduct of scientific inquiry" (p. 10) I think that you can conduct it all on your own. But having some company is nice.

<u>7. Methodology and Intuitions (Yagisawa)</u>

The methodology I urge is naturalistic. Yagisawa comments:

> I find little concrete evidence that Devitt practices what he preaches. I fail to find his "scientific methodology for semantics" [4.8] in action anywhere in the book. What empirical scientific evidence does Devitt offer for his important claims? I see no data collection, no empirical experimentation, no statistical analysis. (p. 1)

The latter practices are certainly common in science, but I do not preach that semantics, as presently practiced, is a science. I preach that it is empirical. Most empirical claims are not part of science, they are part of "folk theory." Semantics ought to be a science but at this stage it is more of a "protoscience." What makes semantics empirical is that such support as it has and might have is empirical and not a priori.

Yagisawa cites a number of my claims as evidence of my unnatural practices. He rightly sees these claims as intuitive judgments but implies that they are <u>therefore</u> not empirical. A central conclusion of my methodology is that a claim can be both intuitive and empirical:

> So, we have found a place for intuitions. It is important not to exaggerate that place. At best the intuitions are likely to be seriously incomplete, reflecting only part of the theory we need. And they may be wrong: They are empirical responses to the phenomena. ...even the most basic intuitions, expressed in identification experiments, are subject to revision in the face of scientific examination. <u>Intuitions are often needed to identify the subject matter for the descriptive task, and may be otherwise helpful, but nothing ultimately rests on them</u>. (p. 74)

Naturalists do not deny a place for the intuitions so beloved by the defenders of armchair philosophy. We simply insist that the place is empirical not a priori. Philosophers bring with them to their armchairs a whole lot of hard-won empirical information, the wisdom of the ages.

It is clear that Yagisawa finds <u>Coming</u> not simply wrong but distasteful (even finding its use of `meaning' "singularly devious"! (p. 4)). Why? He has a taste for the unnatural practices that Bigelow wants to save me from.

8. The Argument for Molecular Localism (Levine, Demopoulos)

We have earlier (sec. 2) considered Levine's response to my criticisms of an argument against molecular localism, against the view that a few of the inferential properties of a token may constitute its meaning. In this section we shall consider his response to my argument for molecular localism (Coming: ch. 3). This response raises a number of interesting issues.

Levine's presentation of my discussion is a bit misleading in two respects. First, he writes as if my concern is with just one principled basis issue but I actually distinguish the following two: (A) the basis for supposing that meanings are localistic not holistic; and (B) the basis for supposing that meanings are certain localistic properties not other localistic ones. The issues are closely related but they are different: in arguing about (A), the molecular localist rejects holism; in arguing about (B), the molecular localist tries to sustain a particular localism (not merely some localism or other). Clearly both issues are important to the defense of localism. Second, Levine claims that I offer two arguments for my localism (pp. 10, 14). In fact I offer three on issue (A) and two on issue (B). The ones that Levine criticizes directly are the two on (B). So Levine does not directly confront any of my three arguments on (A), any of my arguments against holism. The reason for this, he tells me in conversation, is that he has no problem with those arguments, or at least no problem with their conclusion: he has in mind that we should be localists but atomistic ones like Jerry Fodor, holding that none of a token's inferential properties constitute its meaning. Still, if my arguments on (A) are good, it would be odd indeed if the related arguments on (B) were not pretty much on the right track.

My methodological framework yields a perspective on the two principled basis issues that is the point of departure for all five of these arguments. The perspective is that the principled basis for treating one property of a token and not another as a meaning is that the one and not the other plays a semantic role; hence it serves our semantic purposes to ascribe the one and not the other. So the three arguments on (A) are that holistic properties do not serve our semantic purposes, and the two arguments on (B) are that certain localistic properties and not others do serve those purposes.

One of the three arguments on (A) that Levine does not discuss is the argument from our interest in generality. A holism-localism issue can come up anywhere. So we can get the issue in biology, astronomy, and so on. Consideration of a range of examples from various areas show that, outside semantics, we ascribe properties that are localistic, not holistic (3.6). This descriptive view supports the normative one that we ought to ascribe localistic not holistic properties, for just as the descriptive supports the normative in semantics, so it does elsewhere. Normative localism has further support. I argue that our liking for the local is not surprising: Ascribing localistic properties meets our explanatory and practical purposes, particularly our interest in generality (3.7). I apply these conclusions about the holism-localism issue elsewhere to semantics to yield the least theory-laden arument against holism: an interest in generality accompanies any semantic purpose we might have, and so we should always ascribe localistic properties (p. 124).

Another of these arguments on (A) against holism is the argument from the success of our ascriptions. This is a straightforward application of my second methodological proposal: that we should argue for normative/basic doctrines about meaning by looking at the descriptive issue (2.9).

So first I argue that the properties that we do, as a matter of fact, ascribe to token thoughts and utterances for semantic purposes, are all localistic (3.8-3.9). The apparent success of these ascriptions in serving our semantic purposes is very good evidence that we ought to ascribe these localistic properties for those purposes and hence that they are meanings.

Levine does not address this argument directly and, so far as I can see, nothing he says bears on it. But he does discuss a closely related argument on issue (B). I pose the issue as follows:

> What is the basis for our choice among the many localistic properties of a token? We could ascribe properties constituted by many different small sets of inferential properties, each of which would be localistic and many of which might satisfy our desire for generality. How do we choose? (p. 126)

I answer the question by appealing to the success of our current practices again.

> We ascribe those localistic properties that we think serve our semantic purposes in explaining behavior and teaching us about the world. Our ascriptions are remarkably successful, good evidence that we are right in thinking this. So we have shown that we should choose to ascribe the particular localistic properties that we do ascribe. (p. 126)

Levine is not convinced because he has "serious doubts" about whether "it's in virtue of the properties attributed" that the ascriptions are successful and he rightly thinks that this is crucial (p. 12). Now, of course, there is no guarantee that this sort of argument will yield a true conclusion. Still, I have argued, and Levine does not dispute, that people have been, day in and day out for centuries, ascribing to each other certain attitudes to tokens with certain localistic properties, and those ascriptions have been, by and large, successful at serving our semantic purposes. I take it that this is very strong evidence that people really are in those relations to tokens with those properties and that those properties really play semantic roles and so are meanings. I could be wrong - the epistemic life is hard - but in the absence of a powerful alternative explanation of the success of these ascriptions, I do not think it is reasonable to think that I am wrong. I do not suppose that Levine would claim to have offered such an alternative but, in any case, I do not think that he has.

Levine supports his doubts as follows:

> How does the fact that Jane's symbol <cat> maintains an inferential connection with <animal> help us understand her behavior? If anything, it's the connection between <my cat> and <I care about it> that does the explanatory work here. (pp. 12-13)

I think that this embodies an interesting and important mistake. We are trying to identify the <u>meanings</u> of tokens of <cat> without which it can do no explanatory work anywhere. According to his example (an implausible one, as I said earlier) the molecular localist's view is that <cat>'s property of being inferentially related to <animal> partly constitutes one of its meanings. For this to be supported by my success argument we must <u>first</u> have reached a descriptive conclusion: that this inferential property is part of the nature of the putative meaning **CAT** that we do, as a matter of fact, ascribe in t-clauses to Jane and to millions of others. The success argument concludes that this putative meaning is a meaning and so, it follows, that this inferential property is meaning constituting. We do not have to find any further explanatory work for it to do. The meaning of Jane's symbol <cat> does explanatory work because the belief containing the symbol does explanatory work in virtue of the fact that the symbol has that meaning. So any inferential property that is part of that meaning thereby does explanatory work. So, if we are right in our descriptive conclusion, the inferential property of being linked to <animal> thereby does explanatory work. And if we discovered in our descriptive task, as we obviously would, that the property of being inferentially linked to tokens that mean **CARED ABOUT BY JANE** is not part of the putative meaning **CAT** that we ascribe to millions of token thoughts and utterances produced by people who have never heard of Jane, then that inferential property does not thereby do explanatory work.

Of course, we might be mistaken in our descriptive conclusions but that is another matter on which nothing has been said. And, of course, it might be the case that the putative meaning that, according to the descriptive conclusion, we ascribe for semantic purposes does not really play the semantic role and that some other property constituted by different inferential properties does. Revisionism is possible. But Levine is not arguing for this.

None of this is to deny that the inferential property of being related to a token that means **CARED ABOUT BY JANE** is explanatory in Levine's example. Any inferential property of a token is likely to have explanatory significance in explaining some behavior or other, independent of whether it, or any other inferential property of the token, constitutes its meaning. That is the nature of an inferential property. But any inferential property of a token can have an explanatory significance at all only because the token has a meaning of some sort. An inferential property of a token can do explanatory work either by constituting a meaning that does explanatory work, or by taking us from that token to a token with another meaning that does explanatory work.

The third of my arguments on issue (A) against holism is the argument from Representationalism. This third argument is very theory-laden because it assumes Representationalism: the view that meanings are entirely constituted by representational properties, particularly referential properties. If this theory is correct, there are two reasons for thinking that meanings cannot be holistic. First, our propensity for ignorance and error about the world requires that at most a few inferential properties of a token could be reference determining (unless we are prepared to adopt the sadly popular, but nonetheless bizarre, metaphysical view that we all make our own worlds and so aren't really ignorant and wrong after all). Second, the localistic nature of the properties that we ascribe to the world requires that the meanings that represent those properties be localistic too (3.11).

Levine does not address this argument directly and, so far as I can see, once again, nothing he says bears on it. Once again, he does discuss a related argument on issue (B). This argument uses Representationalism to argue that what makes one localistic property and not another a meaning of a token is that it plays a role in determining reference (3.12). I think that this argument raises some interesting questions, but I want to emphasize that I do not need the argument to defend my localism on issue (B): I can rest with the success argument discussed above. This is just as well, because the argument from Representationalism is so theory-laden.

The "basic principle" of Representationalism for word meanings is not quite what Levine says (p. 14): it is "that the meanings of words are entirely constituted by properties that go into determining their references" (Coming: 3).[16] This principle provides an immediate answer to the question that concerns us: What is the principled basis for distinguishing one inferential property of a token from another as meaning-constituting? The answer is that one but not the other determines the reference of the token. Levine raises another question: Why does one inferential property of a token and not another determine reference? He responds: "The obvious answer, the only one I can imagine," is that the one and not the other is meaning-constituting (p. 15). This answer, he thinks, generates a circularity. But it does not, because it is simply another consequence of the "basic principle" of Representationalism. Representationalism entails both that only reference-determining inferential properties are candidates for constituting meanings,[17] thus providing the principled base we want; and that only meaning-constituting inferential properties determine reference, thus answering Levine's question.

Now, one might respond that Representationalism is clearly a very powerful doctrine, needing a lot of independent support. I agree and spend the rest of Coming trying to provide that support.[18] The argument from Representationalism is, as I said, "very theory-laden." Nevertheless, Representationalism is prima facie plausible, as its historical popularity shows, and it is interesting to see how decisively it alone counts against holism and for localism.

I had another interest in this argument. First, we need a principled basis for counting an inferential property as meaning-constituting only because we need one for counting any property as meaning-constituting.[19] Yet, second, it is noteworthy that many who insist on that need with regard to the inferential properties are not similarly insistant with regard to any other property. Fodor is a

---

[16]Levine distinguishes the causal sense of `determines' from what he calls a "satisfactual" sense. I mean `determines' in the sense in which something that constitutes or realizes a property determines that property. So I certainly do not mean it in the causal sense.

[17]I say "candidates" because it is compatible with Representationalism that some properties that go into determining reference should not constitute meanings.

[18]Applied to sentences, Representationalism is the view that the meanings of sentences are entirely constituted by the properties that go into determining their truth conditions. So, in some sense, I agree with the Dummettian conclusion explored in William Demopoulos' "The Centrality of Truth to the Theory of Meaning." Still, as he notes, the disagreements are vast. Dummett does not accept any of the three assumptions of my semantic methodology that lead me to that conclusion: anti-Cartesianism, naturalism, and realism about the external world.

[19]This point plays a role in the criticism of direct reference mentioned in sec. 11 below.

striking example. On the one hand, he is as vigorous as anyone could be in demanding a principled basis for molecular localism. On the other hand, he proposes an atomistic localism without even mentioning its need for a principled basis. According to his atomism, the meaning of a word, say `horse', is constituted by certain direct causal links to reality - those to horses - and not by others - those to cows, tables, and so on. Why? The question is never explicitly addressed. My third point, and reason for raising this matter here, concerns Fodor's implicit answer. It is easy to see what this answer is because Fodor talks of those causal links to horses in proposing a theory of reference. So the implicit answer is that the links to horses constitute the meaning of `horse' because they are reference-determining. My point is that insofar as this Representationalist answer is good - perhaps not far, given how theory-laden it is - so also is the molecularist's answer to the analogous question about inferential links.

Since Levine does not like the answer it is no surprise to find him leaping to Fodor's defense, in effect denying the first of my three points. He claims that the atomist is in a different position from the molecularist: "the question how you determine which causal links are meaning-constituting never arises"; the atomist is "never in the position of having to identify which are the right causal links by appeal to their being the ones that determine reference" (pp. 16-7). Why on earth not? And if the atomist were really never in that position why is the molecularist in the position of having to identify which are the right inferential links? Why does the molecularist face this terrible problem of selecting an inferential link as meaning-constituting and yet the atomist strangely faces no problem of selecting a causal link? Levine continues: "Rather, if anything, one determines the reference by appeal to the already determined causal links, their identity having been given by one's general theory of meaning" (p. 17). To the extent that this is true of causal links, it is equally so of inferential links: one determines reference by appeal to already determined links, whether causal or inferential, in the sense that one tells what a term refers to by applying one's theory of reference. But this epistemic matter is quite beside the metaphysical point: we need to say in virtue of what the links, whether causal or inferential, are meaning-constituting. The contrast implied by Levine's "rather" is mistaken.

The point could be put like this. `Horse' stands in a referential relation to horses that Fodor thinks is explained by a certain causal relation that I shall call "C." `Horse' also stands in relations to many other things including cows. Call its relation to cows "reference*" and suppose that this is explained by some causal relation C*. Fodor and Levine simply take it for granted that meaning involves C not C*? Why then do they not allow the molecularist to take it for granted that meaning involves one inferential link and not another? Of course, they take it for granted that meaning involves C not C* because they take it for granted that meaning involves reference not reference*.[20] But if they thus take Representationalism for granted, why are they so critical of the molecularist for doing likewise? "And why beholdest thou the mote that is in thy brother's eye, but considereth not the beam that is in thine own eye" (Matthew, ch. 7, v. 3).

[20]In my 1997a (pp. 330-8), I claim that causal-theory responses to Putnam's model-theoretic argument for the indeterminacy of reference still leave us with the problem of explaining the significance of the difference between reference and reference*. I argue that the explanation is to be found by appealing to the purposes for which we ascribe meanings.

In conclusion, Coming offers three arguments that there is a principled basis for supposing that meanings are localistic not holistic. Levine does not address these arguments. Coming also offers two arguments that there is a principled basis for supposing that meanings are certain localistic properties and not others, an argument from the success of our ascriptions and an argument from Representationalism. I place most weight on the first of these because the second is so theory-laden. I have argued that Levine's criticisms of these arguments fail.

9. "Interpretationism" (Davies, Potrc)

The application of Coming's methodology leads me to a robustly "realist" view of meanings. Token thoughts and utterances really have meanings and they have them independently of our practice of ascribing them. An ascription is true only if there is a token with the meaning ascribed. In "Naturalised Semantics and Content-Ascription," David Davies favors a Davidsonian "interpretationist" view. This is a largely anti-realist view of meanings. Meanings (or contents) are not so much discovered as imposed by us in our practices of interpreting each other using principles of charity and the like. On this view, content ascriptions are not, for the most part, descriptive of an independent reality; semantics is, in a sense, ascientific.

My discussion of interpretationism in Coming is short (pp. 66-7), but I have criticized the view in some detail elsewhere (1981a: 90-126; 1997a: 186-99). My line, briefly, is that no good reason has been produced to treat semantics in a special way different from any other science. We ascribe meanings to thoughts and utterances. For the most part these ascriptions appear to be successful in explaining behavior and guiding us to reality. So we have good reason to suppose that the thoughts and utterances really have those meanings, independently of our ascribing them. Interpretationism's opposition to this highly mentalistic view rests on an unargued behaviorism: "Meaning is entirely determined by observable behavior, even readily observable behavior" (Davidson 1990: 314). Why should we not regard this behaviorism as just another dogma of empiricism?

Davies argues that my view[21] faces difficulties arising from a holistic conception of content ascription. He rightly distinguishes two different holistic conceptions. On the first, what content -

[21] Davies is right in thinking that I will not like his calling my view, "teleological realism." That label is stimulated by my frequent talk of our "purposes" in ascribing meanings, a talk which Davies takes over and expands for his interpretationism. But he wrongly claims that I define meanings as properties we ought to ascribe for semantic purposes (p. 5). My "definition" is actually as follows: "a property is a meaning if and only if it plays a semantic role" (p. 61). I consider our purposes in making ascriptions in order to identify semantic roles, in particular the roles of explaining behavior and guiding us to reality (pp. 55-60). And I argue that meanings, so understood, are what we ought to ascribe to serve our semantic purposes (p. 62). But I do not build those purposes into the nature of being a meaning. And, contrary to Potrc's suggestion (p. 8), there is surely nothing unnaturalistic in my teleological talk.

or, as I prefer to call it, "meaning" - a given ascription ascribes in a particular context of utterance depends on many features of that context. This is the conception that I have in mind in emphasizing the distinction between holism about meaning ascriptions and holism about the meanings ascribed; and in arguing that the former does not support the latter but rather some form of meaning localism (pp. 118-20). Later I argue that this holistic view of meaning ascriptions is false anyway (4.11; see sec. 13 below). However, it is the second conception that Davies favors and thinks leads to difficulties for my view. According to this "methodological holism," we should take many features of the context into account in making a content ascription. He likens this holism to Putnam's method of "discounting differences." This method

> requires that we take account of all the inferential properties of a token in determining its correct intentional characterisation, but it also requires that differences in the inferential properties of two tokens not be taken as a sufficient condition for denying that those tokens share intentional content. (p. 3)

I did not discuss this view but, insofar as it is a strictly epistemological view about the evidence we should take into account in making ascriptions, I certainly agree with it: for, construed in this way, it is an instance of confirmation holism.

Davies' first reason for thinking that this methodological holism brings difficulties for my view is a bit limp. From this holistic perspective, he claims,

> it is questionable whether the tokens to which a given localistic property is correctly ascribed, given the norms that govern our ascriptive practice, share any set of "basic properties" in virtue of which this "higher order" property is correctly ascribed to them. (p. 10)

The "basic" properties might be inferential properties or referential properties and the "higher order" ones are putative meanings constituted out of those properties. Now it is, of course, "questionable" whether tokens have the putative meanings ascribed to them, and some eliminativists <u>do</u> question it. The issue is: <u>Can the realist answer the question</u>? I argue that she can (ch. 5). Davies' interpretationism surely leads him to doubt this, for he thinks that there is little reality to meanings beyond the holistic practice of ascribing them according to epistemological norms. But, of course, that is not an <u>argument</u> against the realist view that what makes an ascription correct is that a token really has the meaning ascribed.

Davies' second reason is, as he says, more significant.

> Consider, for comparison, the putative property of being "davesfave", possessed by that painting in a given room in a gallery that, on a given occasion, I judge to be most attractive. Suppose my judgments always reflect a weighing of different criteria relating to "objective" properties of paintings...we would surely resist saying that a given painting possesses the property of being "davesfave" independently of my appreciative engagement with it. (p. 11)

Davies thinks that methodological holism yields a similar "weighing and weighting" of various criteria in ascribing meanings and hence that meanings are no more objective and independent than being davesfave.

Davies' analogy is not a good one. Being davesfave is constituted by the property of being judged by Davies to be the most attractive painting on a given occasion: the dependence on Davies is built into the property's very nature. In contrast, according to the realist, the meaning of a token is constituted by, say, certain of its inferential properties: these are objective properties of the token, not dependent on any ascribers. Aside from that, the weighing and weighting of criteria is beside the metaphysical point of what constitutes being davesfave and of what constitutes that meaning. In the case of being davesfave, the weighing and weighting is an epistemological matter of how Davies tells that a painting is davesfave (perhaps also, partly a matter of how being davesfave is realized on that occasion). In the case of the meaning, the weighing and weighting is an epistemological matter of how we tell that a token has a meaning. To suppose that this epistemological matter is constitutive of a meaning is to beg the question against the realist.

In sum, it seems to me that the difficulties that Davies finds with my view take for granted something that I reject: interpretationism,

10. The Identity Problem and Cognitive Significance (Bertolet)

In Coming, chapter 4, I argue for a certain sort of Representationalism. A surprising aspect of this doctrine is the claim that a word token has several referential meanings. One of these is simply the property of referring to its referent but the others are properties of referring under modes that vary in their "fineness of grain." These modes may often be "descriptive," being constituted by inferential links to other words; that is my molecularism. However, I argue for the somewhat shocking claim that the modes for some words must be "causal," being constituted by direct noninferential relations to reality; proper names are very likely examples (4.5). Drawing on earlier works (1974, 1981a, 1981b, 1989b), I describe an historical-causal theory of names and other "referential" singular terms ("IT") to illustrate this idea (4.6).

According to the appealing Millian theory of names, recently resurrected by direct-reference philosophers, a name's meaning is simply its property of referring to its bearer (more usually, the meaning is simply the bearer itself; but this is an unimportant verbal difference). Various problems drove most philosophers out of the Millian paradise long ago. The most prominent of those problems is the Identity Problem. My position on this problem is the main concern of Rod Bertolet's paper, "Meaning, Cognitive Significance, and the Causal Theory."[22]

Bertolet works very hard at getting a position right before criticizing it. This was not easy for him in this case, firstly because I have changed my position, and secondly because the details of each

[22] He also briefly criticizes my treatment of the Opacity Problem. I shall not discuss these criticisms but my discussion can by easily adapted to apply to them.

of my positions are rather complicated. Still, he has managed it. He is rightly very critical of my earlier position on cognitive significance - particularly as set out in "Against Direct Reference" (1989b) - although, I shall argue, he does not draw quite the right conclusion about it. He is also a bit critical of my position on cognitive significance in Coming. Here I think that he is dead right, and I have now modified my position accordingly. He concludes by briefly airing some further disagreements. Here I think that he is dead wrong and I shall indicate why in sections 11 and 16.

I start with a summary of Coming's view of the Identity Problem (pp. 171-7). Properly conceived, the problem is simply that `a = a' and `a = b' differ in meaning. This difference is intuitively obvious, but philosophers have rightly felt the need to argue for it. The most popular argument applies an epistemic principle along the following lines:

`S1' and `S2' mean the same only if they are alike in informativeness and cognitive significance to all competent speakers.

So, `a = a' and `a = b' must differ in meaning because the latter is informative whereas the former is not. The popularity of this argument provides dramatic evidence of the pervasiveness of Cartesianism in semantics. For, what have epistemic issues about informativeness got to do with semantic issues about meaning? Underlying the argument are assumptions like the following:

(A) `S1' and `S2' mean the same only if all competent speakers know that they do.[23]

(B) A competent speaker knows that `S1' and `S2' mean the same only if they are alike in informativeness and cognitive significance to the speaker.

The Cartesianism of (A) is blatant. There is no good reason to suppose that a person who is competent with a sentence -- who has the ability to use it with a certain meaning -- must thereby have any propositional knowledge about what constitutes its meaning. Hence there is no good reason to suppose that she must know whether another sentence she understands is similarly constituted. (See sec. 3 above.)

Consider IT's picture of competence with a name, for example. Briefly, a person has this competence if she is appropriately linked into the causal network for the name and is disposed to assign inputs of the name to the network and to produce outputs of the name that are causally based in the network. The competence does not involve any knowledge about these matters.

If the argument from cognitive significance fails to establish that `a = a' and `a = b' differ in meaning, what does? An application of the methodology described earlier (sec. 5). We start with the descriptive task. When we look at what is common and peculiar to the tokens to which the folk

[23] Consider the following, for example: "It is an undeniable feature of the notion of meaning . . . that meaning is transparent in the sense that, if someone attaches a meaning to each of two words, he must know whether these meanings are the same" (Dummett 1978: 131).

ascribe the property **A = B** in thought ascriptions, the evidence is overwhelming that `a = a' and `a = b' differ in putative meaning. The folk are normally prepared to assert "Flora believes that a = b" on the basis that Flora has a belief she would express using `a = b', but not simply on the basis that she has a belief she would express using `a = a' (even if the folk know that a = b). The latter belief does not have what is common and peculiar to beliefs that the folk would call "that a = b." One way of putting this is that the folk normally construe ascriptions of identity beliefs opaquely.

We now turn to the normative task. Not only do we normally ascribe different properties to `a = a' and `a = b' for semantic purposes, but the success of our ascriptions suggests that we ought to. An examination of our purposes adds to this case. I conclude that `a = a' and `a = b' really do differ in meaning.

A theory of meaning must then explain this by assigning different meanings to `a' and `b'. So the Millian theory fails. The Fregean alternative solved the problem by identifying a meaning with a property of referring by a certain descriptive mode. But some words must have causal modes and it is very plausible that names are among them. Those modes will solve the problem just as well as descriptive ones. Thus, according to IT, `a' and `b' have different meanings because underlying them are d-chains of different types in virtue of being parts of different networks. A network is constituted by d-chains involving tokens of certain physical types - the conventional forms of the name in speech, writing, and so on - linked together by a certain sort of mental processing in the speech community.

That is *Coming*'s view of the Identity Problem. In earlier works (1974, 1981a, 1989b), however, I implicitly accepted the argument from cognitive significance. Furthermore, I accepted the related requirement on a theory of meaning for names: the theory must *explain* the different cognitive significance of `a = a' and `a = b'. This requirement rests on the equally Cartesian "reverse" of the above epistemic principle: if `S1' and `S2' differ in meaning then they differ in cognitive significance. I accepted the argument and requirement, even though I was anti-Cartesian, because I failed to note that they presuppose Cartesianism. As a result of this mistake, about which I shall say more later, I tried to show how a non-Cartesian theory like IT, positing nondescriptive causal meanings, explained the different cognitive significance of `a = a' and `a = b'. Bertolet's paper is mostly a criticism of my attempt at this task.

Before considering that criticism, I want to emphasize something that Bertolet notes: the criticism is *not* aimed at my earlier attempts to solve the Identity Problem *as Coming claims the problem should be conceived*, i.e. as simply the problem of the different meanings of `a = a' and `a = b'. The IT solution summarized above is a modified version of those earlier attempts (particularly 1989b: 218-9). Aside from the views briefly aired at the end of his paper, nothing Bertolet says counts against these attempts to solve *that* problem. His arguments are aimed at my attempts to use a causal theory like IT to solve a different problem, that of the different cognitive significance of `a = a' and `a = b'. My earlier discussions conflated the problem of explaining this difference with the problem of explaining the difference in meaning.

These causal theories explain reference partly in terms of mental processing, what Bertolet nicely calls "file management." My earlier attempts to explain the difference in cognitive significance appealed to this processing (e.g., 1989b: 228). Bertolet argues convincingly that such an explanation is "indifferent to whether it is the causal theory or the direct reference theory or some Fregean theory hawking descriptive meanings that gives the correct account of the determination of the reference of names" (p. 3). This is a nice point, but I doubt the conclusion he draws from it: "If this is right, then Devitt's proposals about the nondescriptive determination of the reference of names make no contribution to his solution to these puzzles" (p. 3). The task I had set myself was to show how a non-Cartesian causal theory that could explain the difference in meanings could also explain the difference in cognitive significance. That the latter explanation is available to other theories does not count against it. The other theories are ruled out on independent grounds: the Fregean theory, for example, because names do not have descriptive meanings; the Millian theory, because `a' and `b' differ in meaning. My task was to show that the meager resources of the causal theory were sufficient to explain the cognitive difference not that they were necessary to. It is also worth noting that no Cartesian theory could accept my explanation because it is not built on the view that competent speakers know about meanings.

I think that a better conclusion for Bertolet to draw from his discussion of my earlier attempts to explain the cognitive difference is a conclusion he does draw from his discussion of my later attempt in Coming: "if one is anti-Cartesian about meaning, one is not entitled to say that two names seem different because they have different meanings" (p. 15); "whether there is a difference in meaning is...irrelevant" (p. 16). Indeed, from an anti-Cartesian perspective, my task of showing how a causal theory of names could explain the differing cognitive significance of `a = a' and `a = b' is simply a mistake: theories of names, whether causal or descriptive, are irrelevant to this explanation. Of course, a theory of names has to be consistent with the right explanation of this cognitive difference- it has to be consistent with the right explanation of anything! - but it has no responsibility for it.

I came close to getting this right in Coming but I made a few remarks implying, as Bertolet points out, that a causal theory has "a minor role in the explanation of different cognitive values" (p. 15). The theory has, and should have, no role in the explanation.

Still, the different cognitive values do have to be explained somehow. In Coming I moved away from my earlier explanation in terms of mental processing to a much simpler one, the core of which is as follows (pp. 177-8). An instance of the law of identity has two occurrences of the one name (used unequivocally). Anything that appears to be such an instance will seem uninformative. Tokens of `a = a' will (usually) seem to be such an instance, but tokens of `a = b' clearly will not; for, tokens of `a' and `b' differ in crude physical characteristics like sound and shape.

Bertolet rightly insists on the irrelevance of the causal theory to this explanation. The difference in cognitive significance arises from the nature of identity and crude physical differences in the names. Any theory of names can accept the explanation (provided that the theory is consistent with the explanation), whether the theory assigns the same or different meanings to `a' and `b'. So my causal theory has no advantage over direct reference so far as the cognitive-significance problem is concerned, as Bertolet points out (p. 14). But, harking back to what I emphasized before, this is not to say that it has no advantage over direct reference so far as the Identity Problem, properly conceived, is concerned. Indeed, it has every advantage because it offers a solution to the Identity Problem.

11. Direct Reference (Taylor, Bertolet)

Direct-reference philosophers think that a name's meaning is simply its property of referring to its bearer. I agree that this "coarse-grained" property is one meaning of a name but argue that a "fine-grained" property of referring by a certain mode of reference is also a meaning. (Actually, I argue that several such "nested" properties are meanings (4.14-4.18), but we can ignore that here.) My position rests on the argument that it is usually in virtue of having the fine-grained property that a name plays its role in explaining behavior. Given my previous discussion of methodology (see sec. 5), this is sufficient to make this property of a name a meaning.

Bertolet seems to embrace direct reference in a short discussion at the end of this paper, and Kenneth Taylor certainly does in his "The Psychology of Direct Reference." My criticism of direct reference is, briefly, that it lacks a principled basis (4.8, 4.18). As we have already noted (sec. 8), it is appropriate, and common, to demand that the molecular localist provide a principled basis for counting as a meaning certain inferential properties but not others. It is just as appropriate, yet strangely uncommon, to demand that the direct-reference philosopher provide a principled basis for counting as a meaning the property of referring to an object but not the property of referring to it by a certain mode.

In locating the failure of direct-reference philosophers to meet this demand it is helpful to consider their answers to two questions about an apparently opaque thought ascription:

1. Does it ascribe to a name only the coarse-grained property of referring to a certain object?

2. Does what it ascribes explain behavior?

To answer the first question "Yes" is to claim that the ascription is not really opaque after all. Since almost everyone thinks that it is opaque, this response bites the bullet. Nathan Salmon bites it with awesome vigour (1986). Passages like the following suggest that Taylor may bite it too: "only if concepts are coarsely individuated is it plausible that concepts are specified by opaquely construed that clauses" (p. 9)[24] In any case, Taylor's answer to 2 is a clear "No." Indeed, because of cognitive diversity he thinks that the coarse-grained meanings that we ascribe do not come close to explaining behavior: only properties much finer grained even than my ones involving modes will do the job (pp.

---

[24]See also pp. 10-13. His argument here, and in his 1997 which is critical of Coming, does not seem to me to count against my view of what t-clauses specify; see also my 1997c, sec. 3. Taylor's rejection of the view that "opaque attributions are context sensitive" - "the hidden indexical theory" (p. 12) - is along very similar lines to mine (4.11 and sec. 13 below).

7-11). Since the received view is that thought ascriptions do explain behavior, this is revisionism, which I will discuss in section 16. My concern now is with the view that the coarse-grained property but not the very fine-grained one is a meaning (content, concept, etc.). If there is to be any theoretical interest in this view it must arise from the different roles of the two properties. On my view, meanings are the properties that play certain roles in explaining behavior and guiding us to reality. This cannot be Taylor's view because it would entail that the very fine grained property was a meaning. He proposes no other view. Like other direct-reference philosophers, he makes no attempt to provide a principled basis for counting something a meaning.

Bertolet's short discussion is focussed on beliefs rather than their ascription. Still, I take it that his answer to the two questions is the opposite of Taylor's. He answers "No" to 1 because he thinks that what a thought ascription ascribes to a name is not simply its coarse-grained meaning but also some contents of its "file." These contents are associations that the name has with non-reference-determining descriptions. And those contents, but not the meaning, are what explain behavior. So he answers 2: "Yes." So the coarse-grained property that he calls a meaning is part of what is ascribed but is not the part that explains behavior. What then makes that property, rather than the fine-grained contents of the file, the name's meaning? Once again, we lack a principled basis. The view of meaning seems to be an unmotivated stipulation.[25]

Aside from this, both Taylor's and Bertolet's view of thought ascriptions need an argument that they do not get in these papers. I claim that the t-clause of an opaque ascription ascribes to a name the property of referring by a certain mode, almost certainly a causal mode. I argue for this by seeing what is common and peculiar to the tokens to which the t-clause is taken to apply (4.2-4.6). Taylor argues convincingly that the very fined-grained property that he, but not I, thinks explains behavior is not common and peculiar to them. But he does not show that my less fine-grained property involving a causal mode is not common and peculiar.[26] And Bertolet does not show that non-reference-determining descriptions are part of what is ascribed, something that I argue against (4.9-4.10).

12. Definite Descriptions (Neale)

The theory IT, mentioned in section 10, concerns not only names but also definite descriptions, pronouns, and demonstratives. The part that concerns definite descriptions includes what has become known as "the ambiguity theory": a distinction between "referential" and

---

[25]See also my brief discussion (pp. 182-3n) of the similar views of Adams, Stecker, and Fuller 1993, Braun 1991, and Fodor 1990.

[26]Taylor claims that views like mine require a taxonomy of thoughts that "must follow from facts about our shared communicative competence" (p. 12). I think that we taxonomize here, as elsewhere in science, by explanatory significance; for example, significance in explaining behavior. It appears to be a consequence of Taylor's view that we could not taxonomise the thoughts of foreigners, babies, and dogs.

"attributive" descriptions, based on Donnellan's famous criticism of Russell's theory of descriptions (1966). (IT actually makes a cross-the-board distinction between referential and attributive singular terms.) Coming contains little argument for IT since it uses the theory for illustrative purposes only. However, I have developed and defended the theory elsewhere. Part of that defense was a paper (1981b) responding to Kripke's criticism (1979) of Donnellan. My paper abstracts from this criticism four arguments against the ambiguity theory and claims that each of them fails. However, my claim about one of these arguments is decidedly tentative. This argument concerns anaphora: in a dialogue, B uses a pronoun `he' that clearly refers to a woman's husband and yet is anaphoric on A's earlier use of `her husband' which, according to the ambiguity theory, is a referential description referring to the woman's lover.[27] So how can the theory be right? I propose that

> B's use of `he' might be a pronoun of laziness for `her husband' taken attributively, even though A's use of `her husband' is referential....However,...we need to know more about anaphora and more about how much laziness is generally acceptable. Meanwhile, I think it must be allowed that Kripke has pointed to something that may be an anomaly for Donnellan's account. (p. 522)

Stephen Neale does know an awful lot more about anaphora and in "Speaker's Reference and Anaphora" uses this knowledge to provide some evidence that my tentative claim may be right.

Neale's contribution is particularly welcome to the friends of the ambiguity theory because he is a well-known enemy of it. Indeed, he thinks that Russell's theory of descriptions not only does what it was intended to do but much more besides, and has written a terrific book to prove it (1990).

Neale constructs a criticism of my proposal based on a note in Kripke's paper: "such a proposal involves treating B's remark incorrectly as involving a play on words" (p. 12). One option Neale offers me in the face of this criticism is to accept that my treatment involves a play on words but, using his Dialogue V as a model, to argue that this is not incorrect. This seems unpromising to me and, I think, to him. Another option he offers is to deny that the treatment involves a play on words using his Dialogues IV and VI as models. This seems very promising to me, and I thank him for it.

In his book, Neale rejects many arguments for the ambiguity theory. I shall finish this section, ungraciously, by indicating two arguments that he does not consider and that seem powerful to me.[28]

---

[27]This is a more extreme version of the ambiguity theory than I endorse. I tend to the view that the lover's failure to be not only the unique husband of the woman but any husband of her at all prevents `her husband' referring to him (p. 319).

[28]I don't know of any explicit presentation of the first of these arguments but one might generously see it as implicit in my 1981a, pp. 50-4 and 1981b, 516-20. The second is in my 1981a, pp. 37-47, 50-2, and 1981b, 514-18.

Kripke and Neale do not deny that a description can be used referentially to refer to a particular object the speaker has in mind. They rightly emphasize that this fact about "speaker meaning" does not show that the object in mind is the literal or conventional referent. First argument: (i) Not only can we use descriptions referentially, we regularly do so. (ii) This regularity suggests that there is a convention of so using descriptions. (iii) This convention seems to be semantic - and not, say, pragmatic - as semantic as that for the Russellian attributive descriptions.[29] The second argument is summarized as follows in Coming: "(i) The description `the F' in its [referential] use is just like the deictic demonstrative `that F'. (ii) The plausible theory of these demonstratives is not a Russellian description theory but a historical-causal theory" (p. 164n).[30]

13. The Context-Dependency of Attitude Ascriptions (Bezuidenhout, Goble, Bach)

According to Coming's methodology, we should look to the descriptive semantic task for evidence on the basic task (2.9). The descriptive task is to explain the nature of the properties that we do, as a matter of fact, ascribe for semantic purposes, using attitude ascriptions. In tackling this task, I draw on the classic discussion of attitude ascriptions generated by Quine (4.2). According to my Quinean view, an ascription like

(a) Ralph believes that Ortcutt is a spy

is mildly context dependent: it is ambiguous, having both a transparent and an opaque reading. Some philosophers think that these ascriptions are much more radically context-dependent. I later consider and reject such views, including the "Hidden-Indexical" theory (4.11).

In "How Context-Dependent are Attitude Ascriptions?", Anne Bezuidenhout has the good idea of contrasting my view with a more context-dependent view that draws on work in pragmatics. She is critical of my view and skeptical of my criticisms of "HIT," the Hidden-Indexical theory. However, her grasp of my views is sadly infirm.[31]

I start with her account of my view of the meanings ascribed by attitude ascriptions (rather than of the meanings of attitude ascriptions). My initial discussion of sentences like (a) and

[29]But the distinction between what is pragmatic and semantic is subtle; see my 1997b: 124-8 and sec. 13 below.

[30]Although this argument is most persuasively run using "incomplete" descriptions like `the table' as examples, it is quite different from the argument that incomplete descriptions are not Russellian (an argument that Neale does consider). That argument is a negative one against Russell's theory. The present one is a positive one for the ambiguity theory.

[31]Is her suggestion (p. 1, n*) that she has "knowingly misrepresented" my views a Freudian slip?!

(b) Ortcutt is such that Ralph believes him to be a spy

leads me to conclude that the the folk ascribe at least three different sorts of putative meaning to a definite singular term (p. 148). First, there is the property of (purportedly) referring to a specified object under a specified mode; thus, the opaquely construed (a) ascribes the property of (purportedly) referring to Ortcutt under the mode of `Ortcutt'. Bezuidenhout describes this meaning as: "(1) the property of referring under a mode to an object" (p. 2). Note that this description does not make explicit that the reference must be under a specific mode, like that of `Ortcutt', and not simply under some mode or other. It soon emerges that she does not have a specific mode in mind and hence has misunderstood. Second, there is the property of referring to a specified object by any mode at all provided it is en rapport; thus, the "rapport-transparent" (b) ascribes the property of referring en rapport to Ortcutt. This is captured by Bezuidenhout's description (2). Third, there is the property of referring to a specified object by any mode at all; thus, the "simply-transparently" construed (a) ascribes the property of referring to Ortcutt. This is captured by Bezuidenhout's description (3). I later argue that although opaquely ascribed meanings - those of type (1) - may often involve descriptive modes, they must sometimes involve causal modes (4.5). Later still, to illustrate how such causal meanings are possible, I describe IT (see secs 10 and 12 above), pointing out in passing that IT can explain en rapport meanings, those of type (2) (p. 166). They are properties of referring by causal modes of the sort IT describes.

These later developments, together with her misunderstanding, lead Bezuidenhout to the following mistake:

> property (1) can be identified with the more specific property of referring under a descriptive mode to an object, whereas property (2) would be identified with the more specific property of referring under a causal mode to an object. (p. 3)

Meanings of type (1) always involve a specific mode of referring. This might be a particular name like `Ortcutt', a particular description like `Quine's favorite dean', or some demonstrative or other (see below). In contrast, meanings of type (2) do not involve a mode that is in this sense specific: (b) would be true if Ralph believes spyhood of Ralph under any en rapport mode at all; it does not require that the mode involves `Ortcutt' rather than `Bernard', `the man seen at the beach', a demonstrative, or whatever. According to the later developments, meanings of type (1) are sometimes descriptive and sometimes causal and the causal ones, like the meanings of type (2), can be explained by a theory like IT. But the causal meanings of type (1) still differ from those of type (2) in that they involve specific modes of reference; for example, they may involve a name like `Ortcutt'. Causal meanings of type (1) are a species of meanings of type (2).

As just indicated, my initial Quinean position is that (deictic) demonstratives in opaque t-clauses ascribe the property of referring by a demonstrative but not necessarily by the particular demonstrative in the t-clause. I call a mode of referring by some demonstrative or other, "a general demonstrative mode," and by a particular demonstrative, "a particular demonstrative mode." Mark Richard's steamroller case shows that the property of referring by the general demonstrative mode is not adequate to explain behavior in certain circumstances: we need the finer grained property of referring by a particular demonstrative mode. So, according to my methodology, that finer grained property is a meaning. Yet, according to my initial position, our standard opaque ascriptions do not ascribe this finer grained meaning. So we need a nonstandard ascription to ascribe this meaning. This leads me to contemplate, and prefer, a minor modification of my initial position: in situations like Richard's, the context determines which particular finer grained meaning is ascribed by the standard ascription. This modification is a move toward the radical context dependency of HIT, but it still falls far short of it (4.14).

After describing this discussion, Bezuidenhout makes the following claim that she later says is the basis for her "central disagreement" with me (p. 20): "Devitt's distinction between general and specific demonstrative modes of reference seems to be a special case of Francois Recanati's (1993, chap. 4) distinction between linguistic and psychological modes of presentation" (p. 12). It seems to be nothing of the sort! Recanati's distinction is between the modes of presention of linguistic expressions and the modes of presentation of thoughts. This distinction is inconsistent with my position! I think that utterances and thoughts have the same modes, hence meanings.[32] Thus, on the modified view that I favour, C's utterance to B,`The man on the telephone believes that you are not in danger', in the context of Richard's example, ascribes to A's belief the property of referring to B by the particular demonstrative mode associated with `you'. And if C were to utter, `The man on the telephone said that you are not in danger', he would ascribe exactly the same property to A's utterance. I think my discussion in chapter 4 provides the basis for an argument against Recanati's distinction, but this is not the place to show that.

So much for the issue of the meanings ascribed by attitude ascriptions. Turn now to the issue of the meanings of the ascriptions themselves; in particular, to the issue of the role of the context in determining those meanings. (Differences of opinion on the latter issue may not have any implications for the former issue, as I indicate in discussing HIT; pp. 202-3.)

Bezuidenhout has doubts about my criticisms of HIT. She claims that my "principle objection is that HIT has no account to give of the way in which conventional meaning plus context enables a listener to recover the mode being ascribed" (p. 7). This is certainly one of my principle objections, but there is another that she does not mention. This is that HIT lacks motivation: "What we need are some examples that favor the extreme context dependency of the theory over the mild context dependency of the Quinean theory" (p. 203). I claim that the literature does not provide such examples and that I have been unable to imagine any. I shall return to this issue of motivation in a moment.

Bezuidenhout's discussion of what she takes to be my principle objection starts as follows: "Sometimes Devitt's complaint seems to be that HIT gives no role to the conventional meanings of the terms used in an attitude ascription" (p. 7). Oh dear! According to HIT the conventional

[32] Thus my preliminary remarks in chapter 4 include: "My main examples of meaningful sentence tokens will be beliefs, but I take the discussion to generalize to other thoughts and to utterances" (p. 140, emphasis added).

meanings of those terms, together sometimes with the context, determine the referent involved in the property ascribed. This is acknowledged in my discussion (p. 201) and is obvious anyway. So of course I don't make the alleged complaint. What those conventional meanings do not determine according to HIT, but do according to my Quinean view, is the mode of referring to that referent involved in the property ascribed. This is my "complaint." So Bezuidenhout's criticism using the analogy of `It is raining' is beside the point.

Next, Bezuidenhout, faults my argument against HIT for "assimilating the phenomenon of implicit meaning" and "semantic underdetermination to ellipsis" (pp. 8-9). This assimilation demands that HIT treat the dependence of attitude ascriptions on context like that of the elliptical `It is raining'. Some who urge HIT encourage this assimilation by making the comparison with `It is raining' (e.g., Crimmins and Perry 1989: 699-700) and so I spend some time arguing against the comparison (p. 205). However, I explicitly do not make the assimilation. I allow that an attitude ascription's dependence on context "could be sui generis" and argue against that (pp. 205-6). Strangely enough, Bezuidenhout herself summarizes this argument earlier (p. 1).

Finally, we must consider Bezuidenhout's criticism of my Quinean view. This criticism draws on work in pragmatics. I agree with much of this work and this is not the place to argue my disagreements. I shall make only one point, a point arising from my discussion of methodology. If claims that such and such is a matter of "pragmatics" not "semantics," and that so and so is part of "the semantic content," "the proposition expressed," or "what is said" rather than an "implicature," are to have substance and not be mere verbal stipulation, they need to be accompanied by an adequate account of the principled basis for the distinctions. I do not say that such an account cannot be given but I do say that it is hard to give and mostly is not given (cf my comments on direct reference in sec. 11).

Bezuidenhout argues that utterance interpretation in general is highly context-dependent whilst allowing that conventional meanings constrain those interpretations. This is fairly uncontroversial. The interesting issue is the particular contributions of context and convention in each case. To support her disagreement with my Quinean view of attitude ascriptions she needs to produce examples with three features: (i) the examples are attitude ascriptions, not other sentences; (ii) the examples show the role of the context in determining the mode of reference ascribed, not something else; (iii) they show that this role is larger than my view allows. All of this is necessary to motivate her more context-dependent view of attitude ascriptions, just as we have already indicated that it is to motivate HIT. Yet her first example (pp. 18-19), showing that the context can determine that a pronoun is used attributively, lacks feature (i): it is not an attitude ascription (and is congenial anyway; cf sec. 12 above). And her second example (p. 19), showing that we use the context to disambiguate anaphoric reference, lacks feature (ii): it is about the determination of reference not modes of reference (and is also congenial). So her case comes to rest on her third example (p. 19), which certainly has features (i) and (ii) and may seem to have (iii):

(16) Superman is such a master of disguise that even Lois Lane believes he is just a reporter.

On my Quinean view, this is a standard example of rapport-transparent ascription. So, the context has no special role to play: for (16) to be true Lois must have the reporter-belief under some en rapport mode of referring to Superman, which of course Lois does have. Bezuidenhout thinks that (16) requires Lois to have her belief under a more specific mode determined by the context. I think that she has a point, but her suggestion, "some such mode as nerdy newspaper reporter known as `Clark Kent'" (p. 19), is surely wrong. What (16)'s talk of disguise "implies" is that the mode is demonstrative: Lois is disposed to say of the disguised Superman, "He is a reporter" and the like. Should I then modify my Quinean view, accepting more context dependency? This is a subtle matter but, for reasons I have presented elsewhere (1997b: 124-8) in discussing a somewhat similar example of Richard's (1997: 101-5), I am inclined to think not. I am inclined to hold to my view of the literal meaning of (16) in this context, taking the more specific mode to be a conversational implicature. This distinction demands just the sort of principled base talked about in the last paragraph. Whence the subtlety.

Lou Goble, in "Translucent Belief Ascriptions," is also critical of my Quinean view of belief ascriptions, in particular of my view that an attitude ascription like

(1) Jones believes that Mr. Clemens is dull

is ambiguous, having both a transparent and an opaque reading. He argues that it has neither of these readings but rather a "translucent" one which has different truth conditions in different contexts. I don't think that the argument works.

Consider Goble's argument against the transparent reading. Goble builds on an example I took from Stephen Schiffer. There seems to be a way of interpreting

(3) Ralph believes that Smith's murderer is insane

such that it licenses

(4) Ralph believes that Big Felix is insane.

I argue that it licenses a range of belief ascriptions of this form obtained by substituting for `Smith's muderer' terms that refers to Big Felix. Goble agrees, but not with my conclusion that it licenses any such substitution. To support his disagreement he continues the story. Ralph is to have dinner with the mayor next Tuesday. Unbeknownst to Ralph, Big Felix is the mayor. Discussing this, the moll says,

(5) Ralph does not believe the man he's to have dinner with next Tuesday is insane,

which seems true. So, Goble argues, (3), (4) and the others in the range "cannot be entirely transparent for that would license the substitution of `the man Ralph is to have dinner with next Tuesday' for `Big Felix' or the other singular terms, which would render (5) false" (p. 8).

Goble has overlooked one of the morals of Quine's famous story of Ortcutt (1966: 185-9). On the basis of Ralph's beliefs about a man in a brown hat, it is true to say, transparently,

Ralph believes that Ortcutt is a spy.

But on the basis of his beliefs about a man seen at the beach, it is also true to say, transparently,

Ralph believes that Ortcutt is not a spy.

The moral we should draw is that a person can have two apparently contradictory beliefs about the one person under two different modes of reference, without having made any logical error. Similarly, in Goble's example, Ralph has two apparently contradictory beliefs about Big Felix, one under the mode, `the murdererer of Smith', and the other under the mode, `the man he's to have dinner with next Tuesday'. He has two distinct mental states brought about by two distinct causal processes. (3) and (4), construed transparently, are accurate descriptions of the state brought about by the murder scene. (5), construed transparently, is an accurate description of the state brought about by socializing with the mayor. The Quinean view survives.

Next, consider Goble's argument against the opaque reading. Goble has two cases here. First, he claims that there may be a setting in which it would not be admissable to substitute `Mark Twain' for `Mr Clemens' in

(6) Jones believes that Mr. Clemens is dull

but it would be to substitute `Samuel Clemens', or even `Sam' even though Jones may be ignorant of Mr. Clemens first name. He concludes that "although in this setting (6) is opaque and resists some substitutions of co-referring terms, it is not closed to them all" (p. 9). Second, he claims that "we can also imagine there being other terms, e.g. definite descriptions, such as `Jones's next door neighbor', that might be perfectly substitutable in that same context" even though Jones may again be ignorant of the relevant identity (p. 10).[33]

Goble does not give any details but, in the first case at least, I think he is right. Cases like this and, more obviously, the ascription of thoughts and utterances to foreigners, require some modification of my initial position on the meaning we opaquely ascribe to a name. The intuitive idea of this initial position is that the meaning is the property of referring by that very name. But what does this talk of "that very name" amount to? It amounts mostly to this: the meaning involves certain

[33]Kent Bach, in "Do Belief Reports Report Beliefs?", claims that substitutions are in order where the relevant identities are "presumed" (pp. 15-16). The idea is, I take it, that the substitutions are in order if the speaker presumes that the subject knows the relevant identities. The substitution is uncontroversial, of course, if the presumption is right, but what if it is wrong? What if the subject does not know the presumed identity? Bach says nothing to show that the substitution would still be in order. My argument is that it would not be.

physical types which are the various forms of the name - sounds in speech, shapes in writing, and so on through the other media - linked together by a certain sort of mental processing in the speech community (pp. 166-9).[34] But then cases about `London'-`Londres', `Ruth Barcan'-`Ruth Marcus', /<u>don</u>-al-n/-/don-<u>nell</u>-n/ lead me to modify this initial position: the name in a t-clause normally ascribes a "disjunctive mode"; for example, one involving the conventional physical forms of `Ruth Barcan' or `Ruth Marcus'; or one involving `London' or `Londres' (or other "translations" of the `London') (pp. 232-3). Applying this to Goble's case, `Mr. Clemens' in (6) normally ascribes a disjunctive mode involving `Mr. Clemens' or `Samuel Clemens' or perhaps even `Sam' (though this one seems a bit dubious to me). Furthermore, `Samuel Clemens' and perhaps even `Sam' would normally ascribe the same disjunctive mode and so could be substituted for `Mr. Clemens' in (6) and be guaranteed to save truth, as Goble claims. I think that I have accommodated such subtleties within my Quinean framework and I don't see that Goble has shown that I am wrong. In particular, I don't see that he has undermined the Quinean view that ascriptions like (6) have an opaque reading.

Consideration of Kripke's puzzle and of the ascription of identity beliefs indicate more subtleties in the role of names in t-clauses: sometimes those names ascribe finer grained <u>non</u>disjunctive modes involving, for example, `London' but not `Londres;' perhaps there is even some mild context relativity like that proposed in the modified position on demonstratives mentioned above. I argue that even these subtleties can be accommodated within the Quinean framework (pp. 234-40 and sec. 15 below).

I don't know how Goble would fill out the details in his second case to make it plausible that the description `Jones' next-door neighbor' could be substituted for the name `Mr. Clemens' in (6). He needs to describe a special stage setting that makes the mode involving the name salient - for that is the mode under which Jones has his belief - even though the speaker has used the description. He could probably do this. If he did, however, I predict that the case would be amenable to the sort of treatment I have just suggested for Bezuidenhout's (16) and given elsewhere to the example of Richard's: the sentence would literally ascribe the property of referring by the description but would conversationally imply the property of referring by the name.

<u>14. "Putting Metaphysics First" and Propositions (Potrc, King, Mills, Bach)</u>

Potrc' discussion of the bearing of metaphysics on the localism issue in semantics, "A Localist Metaphysical Semantics?", may reflect a misunderstanding of my position on this. So here are two points of clarification:

(i) My fourth methodological proposal is captured by the slogan "Put metaphysics first": that our semantics should be guided by general metaphysical constraints (pp. 83-4). I make only one

[34]So Goble has me wrong in attributing to me the view that the meaning ascribed involves the same physical type as that of the name in the ascribing t-clause (p. 5 and n. 9). The name token that ascribes the meaning must, of course, have <u>one</u> of the physical forms of the name, but a name that has the ascribed meaning might have <u>any</u> of those forms.

explicit application of this: to argue against positing Platonic entities like propositions (4.12). I shall say more about this in a moment.

(ii) Implicitly, however, I apply the proposal in giving a role to the metaphysical doctrine, realism about the external world, which I have argued for elsewhere (1997a). I think that realism is an over-arching metaphysical doctrine that should be a background assumption of all science including semantics. I argue that realism does count fairly decisively against one sort of holism, a holism about meanings constituted solely by referential properties. This sort of holism faces a choice between general reference failure and a bizarre, though sadly popular, constructivist anti-realism (3.11), as briefly indicated above (sec. 8). However, the more respectable holisms are not of this sort. Beyond this rather small role, realism does not play any direct role in my argument for localism. In particular, I do not think that the principled basis for localism can be established by appealing to metaphysics, a view that Potrc contemplates and rejects.[35]

Jeffrey King starts his paper, "Propositions Even a Naturalist Can Believe In," by raising the good question: What has naturalism to do with an opposition to propositions? Naturalism does support putting metaphysics first. First, according to naturalism, any semantic theory has no privileged a priori status: it is just one empirical theory among many of the world we live in. Second, semantics is in much worse shape than almost all empirical sciences. So semantics is a very inappropriate starting place for a metaphysics (so also is epistemology). Rather, we should use an empirical metaphysics based on our best science as a starting place for semantics (and epistemology). Thus naturalism reverses the order of procedure followed by the twentieth century's "linguistic turn" (and that followed by the epistemology-based philosophy of earlier centuries). We come to semantics, or should, with a well-based realist metaphysics (1997a: 232-3, 284-5).

This alone does not rule out propositions. But, guided by the slogan, I offer four reasons for resisting the positing of propositions in semantics (p. 210). First, in semantics, as in everything else, we should follow Occam and Quine in positing only such objects as are needed to explain the phenomena. I argue that we have no need for propositions (4.12). Second, propositions are posited primarily to give meanings to t-clauses. This smacks of positing a golden mountain to give meaning to `the golden mountain'. Part of what discredited this Meinongian procedure was that we had no nonsemantic reason for believing in golden mountains. In general, we should expect that the meaning of a sentence will be explained in terms of its relations to a reality that we already believe in for reasons independent of our theory of that meaning.[36] And, in the case of t-clauses, that reality

[35]Potrc also says: "Anti-realism claims that some or all of semantic properties depend on the mind" (p. 4). So, we might infer, realism denies this mind dependency of meanings. In any case, I do not deny it: as already noted (sec. 3), my view of meanings is as mentalist as could be. The realism I have argued for is committed to the mind-independent existence of the commonsense and scientific physical world not to the mind independence of meanings.

[36]Bach seems to disagree, and for a curious reason: "philosophy of language should not let metaphysical considerations so easily trump semantic ones, if only because a language might have bad metaphysics built into it" (p. 4). It is easy to see how people who use a language can have a bad

consists in people having psychological states that play certain causal representational roles; it does not consist in propositions. Third, the nature of Platonic objects like propositions is very mysterious. Fourth, because propositions and the like can play no causal role in mind and language, we have the best of reasons for thinking that they are not part of mental and linguistic reality.

King is fairly unmoved by the Quinean and nominalist scruples that these four reasons reflect. However, as he notes (p. 14 and n. 6), the difference between us may not be great. First, despite my nominalist scruples Coming is full of talk of the properties and relations of concrete entities (tokens). The nominalist hope must be that all such talk can be paraphrased away when the ontological chips are down. Second, King's propositions are not the Platonic entities I find mysterious. They are not "transcendent," separate from concrete particulars, and outside space and time. So it may well be that King's talk of propositions can be paraphrased into the talk of properties and relations that I already indulge in. Indeed, our views often seem quite similar:

1. My view that an utterance token has a meaning property with a complex structure (pp. 12, 56-7) is similar to his view that the utterance expresses a proposition with a complex structure (p. 6).

2. My view that an utterance (normally) has a complex meaning-determining structure (4.4) is like his view that it has an "SI" (p. 3).

3. I think that a (normal) token utterance/thought has its structured meaning partly in virtue of the token's structure. This mirroring of structure seems to be what he has in mind in claiming "the structure of a proposition is identical to the structure of the SI expressing it" (p. 5).

In contrast, my differences with Eugene Mills, indicated by his "Devitt on the Nature of Belief," are clearly great. I hold to the popular "representational theory of the mind" ("RTM") according to which the psychological reality underlying what seems to be a true ascription,

(1) Alan believes that pigs fly,

is as follows: Alan stands in a certain relation to a token mental representation that means that pigs fly. Mills rejects RTM and wants to find a place for propositions.

RTM raises the question: What are we to say about the semantics of (1)? In answering this we need to consider whether the underlying psychological reality makes (1) true. My preferred position is that it does, as Mills indicates. It is then natural to take (1) as asserting that this psychological reality obtains. We might then call the relation involved in that reality, a relation

---

metaphysics - some quantify over propositions, for example - but how can a language have a bad metaphysics? It can include terms like `Pegasus' and `unicorn' but Bach surely does not think that this requires the philosophy of language to commit to Pegasus and unicorns. Instead we seek theories of reference for names and kind terms that allows for the possibility of reference failure. Thus metaphysics easily trumps semantics.

between Alan and a mental token, "the believing relation" or, perhaps, "the belief-making relation" or even "BEL." Mills names this position "SS." Alternatively, one might take (1) as asserting that Alan stands in the believing relation to a proposition, a relation that obtains in virtue of his standing in, say, the "belief-making relation" to the mental token. Mills names this position "CS." Implicitly, at least, Coming favors the former construal of the preferred position. Finally, we might suppose that (1) ascribes a relation to a proposition and is therefore false; Alan stands in a relation of a mental token that does not make (1) true. This is my fall-back position that Mills names "DS." I write as if Alan stands in a "believing" relation to his mental token (p. 213), but nothing much hinges on this way of talking.

Mills makes very heavy weather of interpreting my talk of Alan standing in a believing, or belief-making relation, to a mental token (sec. II). His approach reflects a misunderstanding of RTM and a liking for the methodology of "ordinary language philosophy."

According to RTM, Alan stands in a certain functional relation to a mental token that means that pigs fly. Coming has a lot to say about the meanings of mental tokens but nothing about what it is to stand in this functional relation to one. From a naturalistic perspective, we need to give a scientific explanation of this relation. I do not give one because I do not have one to give. In any case, the matter is rather beside the point of Coming. Mills does not seek a scientific explanation of the relation. Rather, he seeks a construal of "the ordinary meaning of the phrase" \`believing a sentence-token' (p. 7). This way of proceding - trying to throw light on the metaphysics of belief by probing folk semantics - is a nice example of what the naturalism-based slogan, "Put metaphysics first," is against. (The futility of the way in this case is demonstrated by the unimportance of whether we call the relation in question "believing," "belief-making," or "BEL.") Suffice it to say that I do not subscribe to any of Mills' construals.

Mills thinks (pp. 4-5, particularly n. 6) that I must deny the "truism":

S believes x iff S believes that x is true.

\`x' stands in for a t-clause. Let the t-clause be \`that p'. Let us call the relation to a token posited by RTM, "BEL." Then according to RTM, on my preferred view, the truism holds because

S BEL a that-p token iff S BEL a that-it-is-true-that-p token.

Mills' view that RTM must deny the truism seems to arise from mistakenly thinking that, according to RTM, the truism holds because

S BEL a that-p token iff S BEL a that-\`p'-is-true token.

The latter is false, as Mills demonstrates.

This mistake may partly explain Mills' weird "Main Counter-example." The counter-example begins with Beth considering "two of her mental tokens, T1 and T2, without knowing what these tokens mean" and then later, still apparently ignorant of their meaning, moving them into her belief box (pp. 7-8). But this is nonsense! Where a token is in Beth's belief box - where Beth stands in the believing relation to the token - there is no <u>further</u> relation of understanding that Beth might or might not stand in to the token: the functional relation of believing brings with it all the "understanding" that is needed. It rather looks as if Mills does not understand RTM.

<u>15. The Meanings of Attitude Ascriptions (Pietroski, Taschek, Bach)</u>

What do attitude ascriptions mean? This is a fascinating and very complicated question that often seems to dominate semantics. Paul Pietroski's "Specifying Senses Innocently," William Taschek's "Putting Pierre and Peter in Context: On Ascribing Beliefs," and Kent Bach's "Do Belief Reports Report Beliefs?" are good examples of the interesting literature the question has spawned. This is not the place for a detailed discussion of these papers, but I shall make a few comments that relate the papers to my own discussion in chapter four.

That discussion pays a lot of attention to attitude ascriptions. However, its focus is unusual in not being on the meanings of those ascriptions, what I call "second-level" meanings, but rather on what the folk use of those ascriptions shows about the meanings of simpler sentences, "first-level" meanings. This focus is dictated by my semantic methodology. In using an attitude ascription, a person applies a certain property to a token thought or utterance for semantic purposes; she applies a putative meaning. So, by considering such uses we can discover what tokens the folk apply the putative meaning to. By seeing what is common and peculiar to those tokens, we get evidence of the nature of that putative meaning. (I call this the "ultimate" method; p. 72.) And that putative meaning is a real meaning provided that, contra eliminativism, it does indeed play a semantic role.

So, the methodology demands that we consider the uses of attitude ascriptions in investigating first-level meanings. It also demands, via its slogan "Put metaphysics first," that this investigation precede the investigation of the meanings of the ascriptions themselves, second-level meanings. In general, the slogan leads us to expect that the semantics of a sentence will be explained in terms of its relations to a reality that we already believe in for reasons independent of our theory of that meaning. That "prior" reality for attitude ascriptions consists of thoughts and utterances partly constituted by first-level meanings (sec. 14 above). So we must find out about those meanings first.

This having been said, it is impossible to avoid making some claims about second-level meanings in the course of investigating first-level ones. One such claim that I make has already been discussed: my Quinean view that attitude ascriptions are usually mildly context-dependent (sec. 13). Another is the "intimate link" - assumed by almost everyone - between the meaning ascribed by an attitude ascription and the normal meaning of its content sentence. The link often seems to be identity (4.11).

Pietroski subscribes to the doctrine of "semantic innocence": the meaning that the content sentence contributes to the meaning of the attitude ascription is the content sentence's normal meaning. This is different from the intimate link. Where the intimate link relates the normal meaning of \`<u>p</u>' to <u>the meaning ascribed by \`that p'</u>, semantic innocence relates it to <u>the meaning that</u>

`p' has in `that p'. Still semantic innocence leads naturally to the inimate link. Everyone agrees that, in an attitude ascription, `that p' ascribes a meaning (or content). What determines what meaning it ascribes? Clearly `p' in `that p' must play a determining role. The natural hypothesis is that `that p' ascribes a meaning that is closely related to the meaning of `p' in `that p'. So according to semantic innocence it ascribes a meaning that is closely related to the normal meaning of `p': the intimate link. I always thought of myself as accepting innocence in Coming, and say as much briefly in passing (p. 198n), but I had no clear idea of how to accept it. Pietroski's ingenious proposal may show me how.

Pietroski proposes a theory that is not only semantically innocent but also Fregean: the t-clause refers to the normal sense of the content sentence. He suggests that "the proposal is one that Devitt should be able to accept" (p. 1).[37] Give or take a few qualifications, I think I can.

First, my Quinean scruples about abstract, particularly Platonic, entities (sec. 14 above) apply to the Fregean senses that Pietroski is committed to. Still, if his proposal would work if we ignored the scruples, it should work if we heed them. Pietroski's "simplest hypothesis is that the Bedeutung of the complementizer `that' (relative to context C) just is the sense of the sentence introduced [by the complementizer] (in C)." The hypothesis identifies "the Bedeutung of a `that'-clause with the Bedeutung of the complementizer that heads it" thus capturing "the Fregean idea that a `that'-clause serves as a device for referring to the sense of its embedded sentence" (pp. 6-7). Heeding the scruples, I seek to capture the idea, roughly, that a `that'-clause applies to tokens with the meaning of its embedded sentence.[38] A Pietroskian proposal for achieving this identifies the role of applying to the tokens with the role of the complementizer `that'.

Second, as already indicated, I argue that the usual attitude ascriptions are ambiguous (sec. 13) and that singular term tokens have several meanings (sec. 10). So the idea I seek to capture is not quite the rough one above. It is that a `that'-clause applies to tokens with a meaning of its embedded sentence. Where the ascription is opaque, the relevant meaning for a singular term in the embedded sentence is its property of referring to a certain object under a certain mode. Where the ascription is transparent, the relevant meaning is simply its property of referring to a certain object. A Pietroskian

[37]Later he raises the worry that his proposal "trivializes at least one aspect of Devitt's thesis: the centrality Devitt gives to `that'-clauses imposes no constraints on semantic theory" (p. 17). He rightly sets the worry aside. In trying to clarify the vague task of "explaining meanings" I take meanings to be properties that explain behavior and guide us to reality (sec. 5 above). But that is not clarification enough: properties that are, intuitively, nothing to do with semantics might play those roles. So I go further: meanings are "properties of sort specified by t-clauses" (pp. 60-61). That is the centrality I give to t-clauses. The nature of these meanings must be discovered empirically, as Pietroski points out. This approach has the advantage of making it fairly clear what eliminativism eliminates (5.1).

[38]I take it that, in Bach's terminology, the Fregean idea is that an attitude ascription "specifies" a belief whereas my idea is that it "describes" a belief. If so, my idea is in accord with Bach's view that a belief report "does not specify what the person believes but merely describes it" (p. 1).

proposal would be that the role of the complementizer `that' is sometimes (partly) to apply to singular term tokens with the one meaning and sometimes, to singular term tokens with the other.

Third, various phenomena, particularly Mates-type sentences, drive Pietroski away from the simplest hypothesis to the much more complicated hypothesis that the Bedeutung of the complementizer is an "interpreted logical form" consisting of the sense together with the "linguistic form" (p. 7). In discussing the well-known puzzles of attitude ascriptions, Coming argues that any token thought or utterance has several meanings, including very fine-grained ones (4.14-4.18): once we see what warrants calling a property a "meaning" in the first place (sec. 5 above), we see that many properties of a token warrant the title. With the help of these many meanings, I would want to argue that Pietroski's phenomena can be accommodated by the simplest hypothesis.

Taschek is skeptical that the puzzles can be solved by the Fregean strategy of supposing that coreferential tokens can have different senses or meanings. Of course that strategy traditionally presupposed that a token has only one meaning that might differ from that of another token. Seeing that this presupposition is baseless should help to remove the skepticism about the strategy.

Taschek's alternative strategy starts with "the Logic Requirement (LR)": we standardly require that the content sentences of belief reports themselves possess, on the occasions of their use, logical properties corresponding to those of what the subjects' believe (p. 1). The logical properties of a sentence or belief are, of course, one part of what it means. So where the intimate link proposes a close relation between the whole meaning of the content sentence and the meaning ascribed, Taschek's LR proposes one between one part of that meaning and the meaning ascribed. My only problem with Tashek's proposal is that he takes LR to be more deep and general - "a fundamental semantic constraint" (p. 2) - than it really is: in my view, for example, the content sentences of transparent reports do not have logical properties corresponding to all those of the beliefs that make them true.

Next, Taschek shows how the "logical potential" of a sentence depends on its "global structure" a structure that is sensitive "to what expressions occur in other sentences." He urges "the limited autonomy of global logical structure": "two sentences can, when used, differ in their logical properties - specifically, in their global logical structure - even though they possess the same grammatical structure - indeed, even the same local logical structure - and the counterpart subsentential expressions each possess the same semantic content." This has "the liberating effect" that "we are no longer obliged to individuate the semantic content of [coreferring] names in the incredibly fine grained way" required by the Fregean strategy (pp. 5-7).

The puzzles arise from our attempts to capture fine distinctions in beliefs using ordinary belief reports. Tashek's strategy is to show that apparently similar content sentences of appropriate belief reports differ in global logical structure and thus, because of LR, ascribe different beliefs.

I do not think that Tashek's autonomy principle really does liberate us from fine-grained meanings. Consider any of the differences in global logical structure that Tashek indicates, and ask the question: In virtue of what do the two sentences so differ? For example, consider:

(5) Superman flies,
(6) Clark Kent flies.

The answer to our question is obvious: (5) and (6) differ in global logical structure because they contain different names. All the differences that Tashek finds in global structure, except the last one considered below, can be thus explained. Moral: the difference in names is explanatorily deeper than the difference in global structure. This raises another question: What is the explanatorily significant difference in the names? Tashek embraces direct reference and so thinks that the difference is not in "semantic content" - what I call "meaning." But then that is just what Fregeans think that the difference is in. Before we start celebrating our liberation we need a reason for supposing that the Fregeans are wrong. Tashek does not offer one. Once again, direct reference is assumed without a principled basis (sec. 11 above).[39] I argue that the explanatorily significant differences are in meanings, often in very fine-grained ones. So Tashek's differences in global structure are explained by differences in meaning and there is no liberation.

Tashek puts his strategy to work first on Kripke's case of Pierre's beliefs about London. Because Pierre's beliefs are not logically contradictory, LR requires that the content sentences of our reports of those beliefs have global logical structures that make them also not contradictory. So, the following pair of reports,

(8) Pierre believes that London is pretty,
(11) Pierre believes that London is not pretty,

are not in order, but (11) and

(16) Pierre believes that Londres is pretty

are in order (pp. 8-13). I have quite a lot of sympathy for Tashek's discussion although I think (pp. 232-5) that the matter is less clear-cut than he does. But the discussion illustrates my point. The content sentences of (11) and (16) differ in global structure because `London' differs in a fine-grained meaning from `Londres'. That difference in meaning is the deep difference.

Tashek turns finally to Kripke's case of Peter's beliefs about Paderewski. Here it seems to me that Tashek's strategy clearly fails. The strategy demands that Tashek find noncontradictory global structural properties for the content sentences of the following reports:

(19) Peter believes that Paderewski has musical talent,

[39]Bach is another example: he claims that direct reference is a "plausible principle" (p. 2) but he gives no reason for thinking so. And the principle seems particularly unmotivated in his hands: he combines the coarse-grained direct-reference view of the content of a sentence with a very fine-grained view of the content of the belief that the sentence expresses (p. 18).

(21) Peter believes that Paderewski has no musical talent.

Tashek does indeed claim that the content sentences have such noncontradictory properties but, so far as I can see, says nothing to explain how they could have (pp. 17-19). In the earlier cases, we have found the explanation in differences in names, differences that I have claimed are differences in meaning. But `Paderewski' in (19) is exactly the same name as `Paderewski' in (21), as Tashek himself argues. The logical structures that Tasheck assigns are not there to be found.

What then should we say about this case? I think that the answer is a longish story positing very fine grained meanings (pp. 236-40).

16. Revisionism (Taylor, Bertolet, Aydede)

I define revisionism as follows: "A doctrine is revisionist in a certain area if it revises the current theory, the status quo, in that area; it claims, normatively, that we ought not to do what, descriptively, we do do" (p. 248). Part of the current theory of mind and meaning is that ordinary thought ascriptions explain behavior. So one way to be a revisionist is to deny this. Taylor does deny it, as I have already indicated (sec. 11).[40] He thinks that these ascriptions ascribe the coarse-grained meanings of direct reference that do not come close to explaining behavior. Even my finer-grained meanings involving modes of reference, particularly the ones involving causal modes, are too cognitively austere to do the explanatory job. Taylor shares the widespread conviction that only something very rich and cognitively fine-grained could do the job. Call this "the rich conviction." It plays a role in driving many people, although not Taylor, to holism.[41] Taylor adopts the conviction because he is impressed by the cognitive diversity of people who can share a thought.

Bertolet is not revisionist in this way but he also shares the rich conviction. He thinks that thought ascriptions ascribe to a name not only a coarse-grained meaning but also the contents of the name's "file" - the many associations that the name has with descriptions - and that it is those file contents that explain behavior (sec. 11).

Related to the rich conviction is "the narrow conviction," which also leads to revisionism. Bertolet does not have this conviction (p. 9n) but Taylor may well have. It is the idea that behavior cannot be explained by anything external to the brain, by anything "wide." So, for example, behavior cannot be explained by my meanings involving causal modes of reference that are mostly

---

[40] Taylor (1997) resists the idea that he is revisionist but not, I have argued (1997c, sec. 3), for good reason.

[41] I think that the passages from David Papineau and Ned Block referred to in Coming (pp. 38-9) are examples.

external. The narrow conviction is also widespread and, more than anything else, drives people to methodological solipsism;[42]

Coming's arguments against these convictions and hence revisionism are briefly as follows. First, the properties that are, as a matter of fact, ascribed to mental words by our ordinary ascriptions to explain behavior may often be cognitively austere and largely constituted by factors outside the head: for they are properties of referring to objects by certain modes, often causal modes. These ordinary ascriptions are generally successful. So it is likely that those words have the austere and wide properties ascribed and that those properties really do explain behavior, contrary to the convictions. (So, according to my methodology described in sec. 5, those properties are meanings.). This argument is not conclusive, of course. I also argue against various revisionist attempts to undermine the argument (ch. 5).

A popular reason for revisionism and the narrow conviction appeals to Twin-Earth examples. These examples are thought to show that psychological interest is only in what is common to a person and her twin. So that is all that a psychological theory should be concerned with: it should be concerned with narrow explanations. Consider, for example, the following wide explanation:

Oscar gave water to Mary because he believed that Mary was thirsty and that water relieves thirst.

It is claimed that the psychologically relevant explanation would apply just as well to what Twin Oscar did to Twin Mary on Twin Earth.

Twin-Earth examples are misleading. Explanations of wide behavior - behavior that is partly constituted externally to the behaver - whether of Oscar giving water to Mary, or of a sugar dissolving in lemonade,[43] can be broken down into an internal factor and an external factor. Where the internal factor does most of the work, as perhaps it does in explaining the sugar, we may often prefer to focus on it and hence prefer a narrow explanation. But where the external factor does most of the work, we are likely to prefer the wide explanation. Twin-Earth examples are misleading because Oscar on Earth and Twin Oscar on Twin Earth are identical, and their behaviors are so similar, that we are not encouraged to think that these behaviors may be largely explained by external factors. Rather, we are encouraged to think that the different external factors for Oscar and Twin Oscar are responsible for the relatively minor differences between Oscar's giving water to Mary and Twin Oscar's giving twin water to Twin Mary, but each behavior is almost entirely explained by a fine-grained internal factor shared by Oscar and Twin Oscar (who, of course, share many fine-grained internal properties). So Twin-Earth examples suggest that some fine-grained internal

[42]In Coming I quote examples from Stephen Stich, Colin McGinn, and Brian Loar (pp. 272-3, 300). I am in no position to deny the appeal of the narrow conviction having once embraced it somewhat tentatively myself (1989a). I now think that it is quite mistaken.

[43]I take the example from Sosa 1997. My discussion draws on my 1997c, a reply to Sosa and Taylor 1997.

property carries most of the burden of psychological explanation. I argue that this suggestion is very likely false (pp. 287-92, 309).

Theories of reference, particularly historical-causal theories, may well show that the one shared internal factor could, with an appropriate change in the external context, play just the same role in explaining a vast range of behavior involving not only giving but taking, kicking, and many other acts; involving not just water and twin water but milk, wine, gold and all the other stuffs; and involving not just Mary and Twin Mary but Reagan, France, and any other object that can be named. Narrow meanings may be "promiscuous" in that they can yield any of a vast range of referential values by changing the relevant external context as argument. Contemporary theories of reference make plausible the idea that <u>a great deal of what determines reference, and hence explains intentional behavior, may be determined by what is outside the head</u>. There may often be very little to the narrow meaning of a word; it may be coarse-grained.[44] Hence its contribution to the explanation of behavior may be relatively small. Putting this another way, the "proto-intentional" narrow behaviors that narrow meanings can explain may be so coarse-grained and promiscuous as not to distinguish behaviors involving any named object, any stuff, and so on. The extent to which this is actually the case remains to be seen as we develop theories of reference. It cannot be settled by appeals to Twin Earth.

How is it <u>possible</u> that the explanation of intentional behavior might be such an external matter? According to the Representationalism I urge, the role of a thought in explaining intentional behavior depends on what it represents. We cannot settle a priori the extent to which intrinsic properties of the mind determine this. What matters to the explanation is simply that the thought has its representational property under a mode appropriate to the behavior it is supposed to explain. Theories of reference will tell us about those modes.

Turn next to the rich conviction. This conviction seems to be leading Taylor to major revisionism, one that abandons Representationalism. He is struck by the fact that mental words that share the property ascribed to them by belief ascriptions may differ greatly from each other in their other properties. He emphasizes that people who have words that share the ascribed property may be cognitively diverse. (This is apparently even more true on his view of what is ascribed than on mine.) In the light of this diversity, this property could not do the job of explaining behavior. We need something much finer-grained. Why? Taylor does not say, but presumably the answer is:

[44] Informational theories place significant constraints on reference change as a result of context change because a term can only refer to what it covaries with in the context. Still, there is room for promiscuity, as Murat Aydede in effect brings out in "Pure Informational Semantics and the Narrow/Broad Dichotomy." A <u>purely</u> informational theories will not get reference right: a person's dispositions will not enable her to discriminate between a range of different possible stuffs only one of which is the referent. What makes that one the referent has to be something about the actual context, an historical element that could vary without any change in the instrinsic state of the person's brain.

because the cognitive diversity among people who have words that share this finer-grained property will be much less.

It will indeed be less, but unless we go all the way to holism, which Taylor does not, there will still be some cognitive diversity among those who share the explanatory property. That is, there will be diversity unless we require that the explanatory property of the word is constituted by all of the word's functional relations, with the result that the property can explain the behavior only of functionally identical people. One wonders what basis Taylor could have for his view that the property our belief ascriptions ascribe allows too much diversity for explanatory effectiveness but that holism allows too little.

Very likely, anything differs from anything else in some respect. Yet, anything is similar to anything else in some respect. Among the similarities between things we often find a property that explains the behavior or characteristics of those things. The things that can be explained by this property can otherwise differ. Perhaps, sometimes they will differ a lot, perhaps sometimes a little. But the mere fact that they differ a lot is no reason for concern about the explanation. A penguin is very different from an eagle but the fact that they are both birds explains a lot about them; Clinton is very different from Joyner-Kersee but the fact that they are both Americans explains a lot about them. A car wheel is very different from a Ferris wheel but the fact that they are both wheels explains a lot about them. There is no scientific principle that says: "If a property is to be explanatory, things that share it must only differ to degree n." No whistle is blown when explained things differ a lot. There is no way of telling a priori that so much difference is too much.

It is beside the scientific point that the people who have a mental word with the wide referential property we ascribe are cognitively diverse. What matters is whether ascribing such properties best explain behaviors. The fact that these ascriptions are so successful in ordinary life and the social sciences is evidence that they are good. Do they face rivals that are as good? Taylor does not say precisely what rival he has in mind but he seems headed toward one that ascribes a nonholistic non-reference-determining functional-role meaning. Coming takes a very dim view of this rival (5.11-5.12). First, we have been given no idea how to explain such a meaning. Which of the indefinitely many functional roles of a word constitute its meaning? Second, we have been given no idea how such a meaning could explain intentional behavior. All in all, the evidence is strong that our ordinary ascriptions do indeed provide the best explanation of behavior.

It is of course possible that ordinary ascriptions do not. It is possible that a future psychology will show that "causally deep" explanations ascribe much richer meanings than do our ordinary ascriptions. Taylor is convinced that this is so. But he has not produced any evidence that it is so and I have attempted to show that it is not so.

In conclusion, Taylor urges what is likely to be a very revisionist view solely on the basis of an unsupported conviction that only something cognitively richer than what we ordinarily ascribe to explain behavior could do the job. I have argued that this revisionism and conviction is wrong. So also is the narrow conviction and the revisionism it leads to.

University of Maryland, College Park

REFERENCES

Adams, Frederick, Robert Stecker, and Gary Fuller. 1993. "Schiffer on Modes of Presentation." Analysis 53: 30-4.

Braun, David. 1991. "Proper Names, Cognitive Contents, and Beliefs." Philosophical Studies 62: 289-305.

Davidson, Donald. 1990. "The Structure and Content of Truth." Journal of Philosophy 87: 279-328.

Devitt, Michael. 1974. "Singular Terms." Journal of Philosophy 71: 183-205.

---. 1981a. Designation. New York: Columbia University Press.

---. 1981b. "Donnellan's Distinction." In French, Uehling, and Wettstein 1981: 511-24.

---. 1989a. "A Narrow Representational Theory of the Mind." In Rerepresentation: Readings in the Philosophy of Psychological Representation, ed. Stuart Silvers. Dordrecht: Kluwer Academic Publishers: 369-402. Reprinted in Mind and Cognition: A Reader, ed. William G. Lycan. Oxford: Basil Blackwell, 1990: 371-98.

---. 1989b. "Against Direct Reference." In French, Uehling, and Wettstein 1989: 206-40.

---. 1993. "A Critique of the Case for Semantic Holism." In Philosophical Perspectives, 7: Language and Logic, ed. James E., Tomberlin. Atascadero, CA: Ridgeview Publishing Company: 281-306. Reprinted, with a new "Postscript," in Fodor and Lepore 1993: 17-60. Page references are to reprint.

---. 1996. Coming to Our Senses: A Naturalistic Program for Semantic Localism. New York: Cambridge University Press.

---. 1997a. Realism and Truth. 2d ed. with a new afterword (1st ed. 1984, 2nd ed. 1991) . Princeton: Princeton University Press.

---. 1997b. "Meanings and Psychology: A Response to Mark Richard." Nous 31: 115-31.

---. 1997c. "A Priori Convictions About Psychology: A Response to Sosa and Taylor." In Philosophical Issues, Vol. 8 ed. Enrique Villaneuva. Atascadero: Ridgeview Publishing Company: ....

Donnellan, Keith S. 1966. "Reference and Definite Descriptions." Philosophical Review 75: 281-304.

Dummett, Michael. 1978. Truth and Other Enigmas. Cambridge, MA: Harvard University Press.

Field, Hartry. 1996. "The A Prioricity of Logic." Proceedings of the Aristotelian Society 96: 1-21.

Fodor, Jerry A. 1990. A Theory of Content and Other Essays. Cambridge: MIT Press.

Fodor, Jerry A., and Ernest Lepore, eds. 1993. Holism: A Consumer Update. Grazer Philosophische Studien 46. Amsterdam: Rodopi, B.V.

French, Peter A., Theodore E. Uehling, Jr., and Howard K. Wettstein, eds. 1979. Contemporary Perspectives in the Philosophy of Language. Minneapolis: University of Minnesota Press.

---, eds. 1981. Midwest Studies in Philosophy Vol. 6, The Foundations of Analytic Philosophy. Minneapolis: University of Minnesota Press.

---, eds. 1989. Midwest Studies in Philosophy Vol. 13, Contemporary Perspectives in the Philosophy of Language II. Notre Dame: University of Notre Dame Press.

Kripke, Saul A. 1979. "Speaker's Reference and Semantic Reference." In French, Uehling, and Wettstein 1979: 6-27.

Levine, Joseph. 1993. "Intentional Chemistry." In Fodor and Lepore 1993: 103-34.

Neale, Stephen. 1990. Descriptions. Cambridge, MA: MIT Press.

Quine, W. V. 1953. From a Logical Point of View. Cambridge, MA: Harvard University Press.

---. 1966. Ways of Paradox and Other Essays. New York: Random House.

---. 1981. Theories and Things. Cambridge, MA: Harvard University Press.

Rey, Georges. 1993. "The Unavailability of What We Mean: A Reply to Quine, Fodor and Lepore." In Fodor and Lepore 1993: 61-101.

Sosa, David. 1997. "Meaningful Explanations." In Philosophical Issues, Vol. 8 ed. Enrique Villaneuva. Atascadero: Ridgeview Publishing Company: ....

Taylor, Kenneth A. 1997. "Same Believers." In Philosophical Issues, Vol. 8 ed. Enrique Villaneuva. Atascadero: Ridgeview Publishing Company: ....