NOÛS 33:4 (1999) 519–557

The Semantics of Belief Ascriptions

MICHAEL MCKINSEY
Wayne State University

Since it was first proposed by Frege (1892), the view that cognitive attitude verbs express mental relations that hold between persons and propositions has dominated discussion of the semantics of such verbs. I will call this view "the relation theory". In the particular case of the verb 'believes', for instance, the relation theory holds that a sentence of the form 'S believes that p' says of the person referred to by S and the proposition expressed by the sentence p that the former bears the relation of believing to the latter. In this paper, I will present an array of evidence against the relation theory, some of it classical and some of it new, and I will argue that this evidence is quite overwhelming and cannot be explained away. I will also propose a new theory of the meaning and logical form of cognitive ascriptions that explains the available evidence. This new theory is based on the concept of linguistic meaning instead of the concept of a proposition, and it provides a compositional account of the meaning of cognitive ascriptions, even though it implies that the cognitive verbs, in their basic senses, do not express relations of any sort. I will conclude by showing that much of the evidence I give against the relation theory can also be applied to refute various recently proposed "contextual" views of cognitive ascriptions.

1. The Russellian theory and the substitution problem.

The relation theory has been proposed in two distinct forms, depending on the view of singular terms with which it has been conjoined. In its original Fregean form, the relation theory is coupled with Frege's thesis that all singular terms express abstract descriptive senses, or modes of presentation. But by now there is almost universal agreement that this thesis of Frege's is false. Kripke's (1972) arguments regarding the modal properties of sentences containing proper names, and Kaplan's (1977) similar arguments concerning indexical and demonstrative pronouns show that these types of singular term are semantically quite unlike

definite descriptions (terms of the form 'the so-and-so').[1] Rather, as many philosophers have urged, the best explanation of the semantic facts about sentences containing names and indexicals would seem to be that terms of these kinds are directly referential, in the sense that the propositions expressed by (simple, non-cognitive) sentences containing such terms are functions solely of the terms' *referents*, and not of any descriptive meanings had by the terms.[2] Terms like proper names and indexicals that are directly referential in this sense I call "genuine terms". Propositions expressed by sentences containing such terms will be called "singular propositions". If we conjoin the direct reference view of names and indexicals with the relation theory of cognitive verbs, we obtain an instance of the relation theory that I will call the "Russellian theory".[3] Since I endorse the direct reference view of names and indexicals, I will assume that the only form of the relation theory that has a chance of being true is the Russellian theory.

One of the main advantages of the Fregean theory is that it provides a straightforward explanation of the apparent fact that substitution of coreferring proper names in cognitive contexts can fail to preserve truth value. By the same token, one of the main difficulties of the Russellian theory is that it implies that such substitution is necessarily truth-preserving.[4] On the direct reference view of names, substitution of one coreferring proper name for another in a simple sentence cannot affect the proposition expressed, since the proposition expressed is solely a function of the name's referent. Thus, since the names 'Hesperus' and 'Phosphorus' both refer to the planet Venus, the following sentences must express the same proposition:

(1) Hesperus is visible in the evening.
(2) Phosphorus is visible in the evening.

But then by the relation theory, the following two belief ascriptions must *also* say the same thing:

(3) The ancients believed that Hesperus is visible in the evening.
(4) The ancients believed that Phosphorus is visible in the evening.

For given that (1) and (2) both express the proposition that Venus is visible in the evening, (3) and (4) both say that the ancients believed this proposition. In general, the Russellian theory implies that substitution of coreferring names in cognitive contexts is always valid, and could not be otherwise. Yet this consequence certainly *seems* wrong. For while (3) is no doubt true, (4) certainly seems to be false, since being ignorant of the fact that Hesperus is Phosphorus, the ancients might well have denied, or at least expressed doubt, that Phosphorus is visible in the evening.

2. Strengths of the Russellian theory.

While apparent failures of substitution raise a serious problem for the Russellian theory, it is important to keep in mind that, as its defenders have correctly em-

phasized, the Russellian theory makes the right semantic predictions in the vast majority of cases in which ordinary names and indexicals occur in the scope of cognitive verbs. (See Soames 1987a, p. 218.) Consider, for instance, the following pedestrian example. The day before the APA convention begins, my friend Larry is standing in line trying to check in at the convention hotel. Like many philosophers, Larry's hair and beard are untrimmed, his clothes rumpled. Seeing Larry, the desk clerk yells to a security guard, "Get that bum outta here!" Observing this, I turn to another friend and say

(5) The desk clerk believes that Larry is a street person.

On the Russellian theory, this use of the sentence 'Larry is a street person' expresses a proposition that is singular with respect to Larry, and the sentence (5) says that the desk clerk bears the relation of belief to this proposition. This is surely a plausible view of what (5) says. Moreover, the view has the desirable implication that (5) says something *de re*, for it implies that (5) ascribes to Larry the property that any object x has just in case the desk clerk believes the singular proposition that x is a street person. It is in fact clear, I think, that anyone using the sentence (5) in this context would be making a *de re* assertion. The speaker would be using the name 'Larry' to refer to a certain man, and using (5) to assert of this man that the desk clerk believes that he is a street person.

A similar point applies to belief contexts containing indexicals, such as

(6) The desk clerk believes that *he* (or: *that man*) is a street person.
(7) The desk clerk believes that *you* are a street person.
(8) The desk clerk believes that *I* am a street person.

Assuming that in each of (6)–(8) the speaker's utterance of the relevant indexical refers to Larry, then uses of (5)–(8) would all make the same *de re* assertion. They would all ascribe to Larry the same property, namely, the property that an object x has just in case the desk clerk believes that x is a street person.

The fact that standard cognitive ascriptions containing names and indexicals are *de re* in meaning should not be confused with a fact about the *scope* of the names and indexicals. I am assuming that in the sentences (5)–(8), the relevant terms are both grammatically and logically in the scope of 'believes'. Thus I am assuming that these sentences are taken to be *structurally* (or logically) *de dicto*. Even so, because of the semantic character of the small-scope names and indexicals, these structurally *de dicto* sentences all turn out to be semantically *de re*. We could put this point by saying that each of (5)-(8), in which the relevant terms have small scope, is semantically equivalent to its structurally *de re* counterpart, in which the relevant term has large scope. Thus (5) is semantically equivalent to:

(9) Larry is such that the desk clerk believes that he is a street person,

(6) is semantically equivalent to

(10) *He* (or:*that man*) is such that the desk clerk believes that he is a street person,

and so on. The important fact is that cognitive ascriptions containing names and indexicals typically do not have *de dicto* readings *as opposed* to *de re* readings. Rather, typically, such sentences' *only* readings are *de re*. The main strength of the Russellian theory is that it has this important fact as a consequence.[5]

3. Russellians' explanation of apparent substitution failure.

Of course, the Russellian theory also has the stronger consequence that, not just most, but *all* cognitive ascriptions that contain names or indexicals have only *de re* readings. This in turn implies that substitution of coreferring names and indexicals in cognitive contexts invariably preserves truth value. What then of the virtually universal intuition that substitution of coreferring proper names can lead from truth to falsehood, as in cases like that of 'Hesperus'/'Phosphorus'?

Most of the defenses of the Russellian theory that I've seen give what is in effect the same explanation of apparent substitution failure. This explanation is based on the hypothesis that normally, the use of a sentence of the form 'S believes that p' will pragmatically convey or conversationally implicate in something like Grice's (1975) sense, that the believer would accept or assent to the imbedded sentence p, although the truth conditions of 'S believes that p' (as specified by the Russellian theory) do not strictly require this. For example, Lois no doubt assents to the sentence 'Clark Kent is a reporter', and so it seems true that

(11) Lois believes that Clark Kent is a reporter.

But Lois would explicitly deny the sentence 'Superman is a reporter', and so it seems false that

(12) Lois believes that Superman is a reporter.

Yet given that (11) is true and that the names 'Clark Kent' and 'Superman' are coreferential, the Russellian theory, as we've seen, implies that (12) is true, contrary to appearances. The Russellian theorist proposes to explain why we have the (on his view mistaken) intuition that (12) is false by appeal to the alleged fact that a use of (12) would pragmatically convey or implicate, though it would not strictly say or logically imply, the falsehood that Lois would assent to the sentence 'Superman is a reporter'.

As far as I know, Tom McKay (1981) was the first to suggest this sort of explanation, and it has since been proposed by several others. But no defender of

the Russellian theory has actually provided any clear account of why or how belief ascriptions like (11) and (12) are supposed to pragmatically convey the information that the believer would accept the relevant imbedded sentence. Some (including Nathan Salmon (1986), Jonathan Berg (1988), and Kirk Ludwig (1996)) invoke Grice's (1975) notion of conversational implicature, but none actually use a Gricean framework to explain how the relevant implicature arises.[6] McKay and Berg both seem to rely on the claim that *typically* the believer will accept the imbedded sentence, so that the assumption or expectation of this acceptance is generated in standard contexts. But McKay also claims that uses of belief ascriptions have "the common secondary purpose" of displaying a sentence accepted by the believer (1981, p. 293), Ludwig seconds the idea that we have this goal in mind (1996, p. 448), and Scott Soames seems to suggest that there is a sort of pragmatic rule or principle in force among speakers to the effect that we should "remain faithful to the words of the speaker unless there is reason to deviate" (1987b, p.119).

Yet even if we grant that we can normally expect with justification that the subject of a true belief ascription would accept the relevant imbedded sentence, there are some clear classes of exceptions to this rule. One important such class, as both McKay and Soames point out, is that of belief ascriptions that contain indexicals. Thus, if Smith believes that he is being attacked by a bear, then he accepts 'I am being attacked by a bear', but not 'he is being attacked by a bear'. (This is McKay's example, 1981, p. 294.) Similarly if Smith believes that I am being attacked by a bear, then Smith accepts not 'I am being attacked by a bear', but rather some other sentence such as 'he is being attacked by a bear'. (Soames gives this sort of example, 1987b, p. 118.) Another sort of case is that of Jones, who believes of Smith—a transvestite—that he is a woman. Of course, Jones does not accept 'he is a woman', but rather, say, 'she is a woman'.

There is a simple reason why a cognitive ascription containing an indexical will not in general pragmatically convey that the agent accepts the imbedded sentence. The reason is just that, as we've seen, such ascriptions are *de re* in significance. Thus the function of an indexical in a belief ascription is simply to refer to an object, so that the ascription can say that a certain agent has a belief of a certain sort that is about that object. It is no part of the indexical's function to imply or suggest that the agent would use any *particular* mode of reference, whether linguistic or mental, to pick out the object of his or her belief. No doubt such a *de re* ascription implies that the agent has *some* way of referring to the relevant object, but it is the nature of a *de re* ascription to remain neutral as to what that mode of reference might be. This is why cognitive ascriptions containing indexicals are exceptions to the Russellians' rule.[7]

But notice that, for the same reason, most standard, ordinary uses of *proper names* in cognitive contexts will also be exceptions to the same rule. Consider our earlier example

(5) The desk clerk believes that Larry is a street person.

Clearly, a use of (5) makes no suggestion that the desk clerk would accept the sentence 'Larry is a street person'. In fact, he would not assent to this sentence, since he has no idea that the man he has in mind is named 'Larry'. In standard, *de re* ascriptions containing small-scope names, the name used will of course be one of the *speaker's* names for the relevant object. But as the example (5) shows, the name need not at all be one of the *believer's* names for that object, and belief ascriptions containing names normally do not suggest otherwise. It seems likely that the vast majority of uses of names in cognitive contexts are like this, and so such uses form a large class of exceptions to the rule that belief ascriptions pragmatically implicate that the believer would assent to the imbedded sentence.

A particularly poignant example of this kind of exception to the Russellians' rule has been described by Jonathan Berg:

> A viewer marvelling at Superman's ability to conceal his identity might remark to another viewer, "Look, there's Superman in his Clark Kent outfit; he's incredibly convincing! *Everyone* thinks he's a reporter—Jimmy Olson, Mr. White—why even that clever *Lois Lane believes that Superman is a reporter*." (Berg, 1988, p. 355; his emphasis.)

In this nice example, our intuition is that the sentence (12) is *true*:

(12) Lois believes that Superman is a reporter.

Moreover, of course, this use of (12) would not at all suggest or implicate the falsehood that Lois would accept the sentence 'Superman is a reporter'. But why not? I think the reason is clear: we can tell from the context that the speaker is using (12) to make a *de re* assertion about Superman, and hence the speaker's use does not implicate that Lois would accept any particular sentences involving any particular ways of referring to Superman.

It should now be clear that the Russellian has failed to explain the intuition that (12) and other similar sentences are false in the standard cases of apparent substitution failure. For if (12) is understood to mean what according to the Russellian theory it *must* mean, then it is understood to be *de re*, and its use cannot have the false pragmatic implicature that the Russellian's explanation requires. To acquire the relevant implicature, apparently, a use of (12) would have to be given a (semantically) *de dicto* as *opposed* to a *de re* reading. By this I mean that the use of (12) would have to be understood semantically as somehow characterizing Lois's belief at least in part in terms of one of Lois's modes of referring or ways of thinking of the (putative) object of her belief (in this case, Superman). This is what I shall mean below when I speak simply of a "*de dicto* reading" of a cognitive ascription like (12).

To explain the intuition that (12) is false in a case of apparent substitution failure, then, it must be explained why people are ever inclined in the first place

to give *de dicto* as opposed to *de re* readings to cognitive ascriptions containing proper names. But the pragmatic rule suggested by the Russellians cannot be used to explain this, since the rule cannot be correctly applied to a cognitive ascription unless it is first assumed that the ascription is being given a *de dicto* rather than a *de re* reading.[8]

4. Descriptive names and the real possibility of substitution failure.

The unavoidable fact is that when speakers have the intuition that substitution of coreferring names in a given cognitive context fails to preserve truth, the speakers are giving at least one of the cognitive ascriptions a *de dicto* as opposed to a *de re* reading.[9] Now for reasons that I explain below, I agree with the Russellian theory that in many, perhaps most, cases of this sort the semantic facts require that the cognitive ascriptions should be given *de re* readings only. If so, then in many such cases, the inclination to give *de dicto* readings to the relevant sentences is simply mistaken, as the Russellian theory implies.

The trouble with the Russellian theory is that it makes this common and indeed natural mistake seem quite inexplicable, or at least wildly unlikely. For the relation theory of cognitive verbs conjoined with the direct reference view of names makes it *impossible* for a cognitive ascription containing a small-scope proper name to ever have anything other than a *de re* reading. One would expect that speakers who constantly attempt to give impossible readings to cognitive ascriptions would be guilty of some kind of radical semantic incompetence regarding the meaning of either cognitive verbs or proper names or both. Yet the tendency to give *de dicto* readings to cognitive ascriptions containing names is thoroughly pervasive, and is manifested by the most competent speakers. Indeed, this tendency seems to be one that is manifested by *all* competent speakers.

Thus, I would urge that only a view on which it is at least *possible* for cognitive ascriptions containing small-scope names to have *de dicto* readings has a chance of accounting for our common intuitions about such ascriptions. In contrast to the Russellian theory, such a view would also make it at least *possible* for substitution of coreferring names to fail in cognitive contexts. On the other hand, as we've seen, most cognitive ascriptions containing ordinary small-scope names and indexicals have *de re* readings only. We need a view that can reconcile these apparently conflicting facts.

I have proposed elsewhere the outlines of such a view (McKinsey, 1994), and I will try to make the view's semantic details somewhat clearer here. My proposal is based on a sharp semantic distinction between two types of directly referring genuine terms. On the one hand, there are what I call "descriptive names", or names that have descriptive meanings in a language. As we shall see, such names are rare, but they do exist and certainly are at least possible. Cognitive ascriptions containing small-scope descriptive names, I maintain, have *de dicto* readings. On the other hand, there are most ordinary names and indexical pronouns, terms that

have no descriptive meanings in the relevant sense. Cognitive ascriptions containing small-scope occurrences of the latter kinds of terms have *de re* readings only.

Let's consider the possibility that 'Hesperus' and 'Phosphorus' have descriptive meanings. Since they are proper names, these terms are genuine terms and thus they do not express Fregean descriptive senses. However, it is consistent with the assumption that these names are genuine terms to suppose that they nevertheless have descriptive meanings in the sense that their referents are *fixed*, or determined, by semantic rules that are expressible by use of descriptions.[10] Thus perhaps the meaning of 'Hesperus' is given by the rule:

(13) For any token α of 'Hesperus', α is to refer to an object x if and only if x = the heavenly body that appears brightest on the western horizon in the evening,

while the meaning of 'Phosphorus' is given by the different reference-fixing rule:

(14) For any token α of 'Phosphorus', α is to refer to an object x if and only if x = the heavenly body that appears brightest on the eastern horizon in the morning.

Descriptive names are those whose linguistic meanings in a language are given by descriptive reference-fixing rules like (13) or (14).[11] Regardless of whether there are any descriptive names in natural languages, it is obviously *possible* for there to be such names. Moreover, if the names 'Hesperus' and 'Phosphorus' did have descriptive meanings of this sort, it would be intuitively correct to say that (3) is true and (4) is false:

(3) The ancients believed that Hesperus is visible in the evening.
(4) The ancients believed that Phosphorus is visible in the evening.

If, as I think, (3) and (4) do have different truth values (given our assumptions), then this would have to be due to the fact that the different descriptive meanings of 'Hesperus' and 'Phosphorus' semantically convey different ways of thinking of Venus, so that, contrary to the Russellian theory, (3) and (4) end up characterizing the ancients' beliefs in terms of these different ways of thinking, even though the imbedded sentences (1) and (2) express the same proposition.

5. The property theory.

One hypothesis, then, on which it is possible for there to be *de dicto* readings of cognitive ascriptions containing small-scope names, is the hypothesis that features of a sentence's linguistic meaning (as opposed to the proposition it expresses) can play a semantic role in determining what cognitive property is

expressed by a cognitive predicate that contains the sentence. Thus we need to address the question of *how* a sentence's linguistic meaning could play this kind of semantic role.

I'll begin by giving an answer for the special case in which the cognitive ascription has only a *de dicto* reading. I will assume that in all such cases, the imbedded sentence contains no indexicals and no proper names except ones that have descriptive meanings. I will say that such a sentence has a "context-independent" meaning.

Following Wilfrid Sellars, I suggest that when we use a cognitive verb to ascribe a mental state or act, we are classifying that mental state or act in semantic terms (see for instance Sellars, 1969). Since on my view this classification is based on the linguistic meaning of the imbedded sentence, it seems that it must be possible for cognitive acts and states to have semantic properties analogous to a sentence's having a given linguistic meaning. Let us suppose, then, that for any indicative sentence S that has a given context-independent meaning, there is a unique property S_M corresponding to that meaning such that, in any possible world w, a sentence S′ has the same meaning in w that S has in the actual world if and only if S′ has S_M in w.

On the view I'm proposing, cognitive operators like 'thinks that' and 'believes that' do not express relations of any sort. Rather, they form one-place predicates out of sentences (open or closed). A predicate of the form 'thinks that S', where S has a context-independent meaning, expresses the property of having a thought that has the semantic property S_M, where S_M is the property corresponding to the linguistic meaning of S, as just explained.[12] If we assume that beliefs are mental states that dispose persons to have affirmative thoughts (i.e., to make judgments), then a predicate of the form 'believes that S' expresses the property of having a belief that disposes a person to have an affirmative thought that has the property S_M.

Even though the sentences imbedded in the belief ascriptions (3) and (4) express the same proposition, they have distinct linguistic meanings. Thus it follows from the above account that (3) and (4) say different things about the ancients' beliefs, and hence (3) may be true while (4) is false. Notice also that, since the relevant imbedded sentences possess their linguistic meanings independently of whether the terms 'Hesperus' and 'Phosphorus' have referents, it follows from the above account that both (3) and (4) can be true even if the beliefs ascribed fail to be about Venus or any other existing object, and hence even if the imbedded sentences express no propositions at all. Thus, on my account, existential generalization on occurrences of small-scope descriptive names in cognitive contexts is invalid.[13]

I want to now briefly explain how my view will deal with cases in which the imbedded sentence contains indexicals or proper names with no descriptive meanings. Again, I take cognitive predicates formed from such sentences to have *de re* readings only. The basic reason for this is that, since they have no descriptive meanings, the imbedded terms are incapable of semantically conveying anything about the agent's ways of thinking about the terms' referents (if any). Hence, the

only semantic features possessed by the terms that are relevant to characterizing an agent's cognitive states are the terms' referents, and so cognitive predicates containing such terms must be understood *de re*.[14]

It is technically rather difficult to give a clear general account based on linguistic meaning of *de re* cognitive predicates, so I will have to relegate a full statement of this part of my theory to an Appendix. But the main idea is intuitively straight-forward. I propose that a *de re* ascription such as, say, 'Ralph thinks that Ortcutt is a spy' says (roughly) that (1) Ralph has a thought that is about Ortcutt, and (2) that thought has a semantic property that corresponds to the linguistic meaning of *some* (unspecified) sentence of the form '*n* is a spy', where *n* is a name (typically a descriptive name) that refers to Ortcutt. (See the Appendix for a clearer and more detailed formulation.)

This proposal for understanding structurally *de dicto* but semantically *de re* ascriptions is based on the traditional way of understanding structurally *de re* ascriptions, according to which such ascriptions do not give complete semantic characterizations of beliefs. Rather, a *de re* ascription only partially characterizes a belief (1) as being about a certain object or objects (the referent(s) of the non-descriptive names or indexicals in the imbedded sentence), and (2) as sharing the form and predicative content of the imbedded sentence. What a *de re* ascription leaves out that a full or *de dicto* characterization of a belief would include is a semantic characterization of the agent's (typically descriptive) ways of thinking of the objects of his belief that correspond to the occurrences of non-descriptive terms in the imbedded sentence.

Note that on this proposal, in contrast to my earlier proposal for understanding semantically *de dicto* predicates, a *de re* predicate does *not* characterize a belief as having a semantic property corresponding to the specific linguistic meaning of the whole imbedded sentence. Rather, again, a *de re* predicate characterizes a belief partly in terms of the objects it is about, and partly in terms of the semantic property that corresponds to the linguistic meaning of the *predicative part* of the imbedded sentence. If, say, the imbedded sentence is 'Ortcutt is a spy', then the belief is partly characterized as having a semantic property that corresponds to the meaning that all sentences of the form 'n is a spy' have in common, namely, (in this case) the meaning of the predicate 'is a spy'. This allows that the belief in question would be fully characterized, or individuated, by a semantic property that corresponds to a linguistic meaning not possessed by the imbedded sentence.

To distinguish it from the relation theory of cognitive ascriptions, I will call the view just described "the property theory".

Given that I'm using linguistic meanings to play the role usually played by propositions, one might wonder why I don't just say that cognitive operators like 'thinks that' and 'believes that' express relations that hold between persons and linguistic meanings. One reason is that I don't like to hypostatize linguistic meanings. But the main reason is that I don't think that cognitive attitudes like thought and belief are directed toward linguistic meanings. Rather, I think that the only *objects* of the cognitive attitudes are propositions. To put it perhaps paradoxically

(given my rejection of the relation theory), we don't believe linguistic meanings, we believe propositions.

I don't deny that some cognitive verbs express relations that hold between persons and propositions. Thus 'believes' no doubt expresses such a relation when it occurs in contexts of the form 'x believes the proposition that p'. But I would draw a sharp distinction between the sense of 'believes' in such relational contexts and its sense in non-relational contexts of the form 'x believes that p'. In a context of the latter sort, such as (3) ('The ancients believed that Hesperus is visible in the evening'), we characterize an agent's belief non-relationally, in terms of a semantic property that is more fundamental than the proposition expressed by the imbedded sentence. This semantic property is one that *determines* which proposition (if any) is believed, in the way that the linguistic meaning of the sentence 'Hesperus is visible in the evening' determines which proposition (if any) is expressed by that sentence.

From my perspective, cognitive relations such as that expressed by 'believes' in contexts of the form 'x believes the proposition that p' are conceptually derivative from more fundamental concepts like that expressed by 'believes that' in contexts of the form 'x believes that p'. Thus, I would define 'x believes the proposition that p' as 'x has a belief that is individuated by a semantic property S_M such that S_M determines the proposition that p'. (A person's belief is "individuated" by a property F just in case any person y would have the same belief in a possible world w if and only if y has F in w.[15] A semantic property S_M "determines" a proposition p just in case any sentence that had S_M would express p.)

6. A relationist objection.

To support their view of the logical form of cognitive ascriptions, defenders of the relation theory often appeal to the apparent validity of various quantificational inferences involving cognitive verbs, such as the following:[16]

(15) John believes everything that Mary asserts.
Mary asserts that smoking is unhealthy.
∴ John believes that smoking is unhealthy.

On the relation theory, 'that'-clauses are singular terms that refer to propositions, and verbs like 'asserts' and 'believes' express relations between persons and propositions. If we let 'S' be a name of the proposition that smoking is unhealthy, then on the relation theory, (15) has the following valid form:[17]

(16) (∀x)(if Mary asserts x, then John believes x).
Mary asserts S.
∴ John believes S.

Since it implies that (15) has a valid form, the relation theory provides a cogent explanation of (15)'s apparent validity. So it might appear that the validity of such inferences is evidence in favor of the relation theory and provides an argument against the property theory.

However, I think that this apparent advantage of the relation theory is really just an illusion engendered by the fact that the verbs 'believe' and 'assert' happen to have relational as well as non-relational senses. In this respect, 'believe' and 'assert' are actually quite unusual. For most other verbs which take 'that'-clauses as complements do not in fact have relational senses. Consider the verb 'thinks', for instance, as it occurs in a sentence like

(17) Monica thinks that Bill is cute.

If 'thinks' expresses a relation in (17), then the result of replacing the 'that'-clause in (17) with a term that refers to the proposition expressed by the imbedded sentence should make sense:

(18) *Monica thinks the proposition that Bill is cute.

But to my ear, (18) in fact does *not* make sense. Certainly, as Kent Bach (1997) has pointed out, a sentence like (17) cannot mean the same as a sentence like (18), since the former implies that Monica *believes* that Bill is cute, while the latter does not (p. 225). Bach also points out that while it always makes sense to suspect, fear, or regret that such-and-such, it does not make sense to suspect, fear, or regret the proposition that such-and-such. I would add that it also makes no sense to desire, intend, hope, or wish "the proposition that p". (I might hope that no millenium celebration is violent, but surely, this is not to "hope the proposition that no millenium celebration is violent," whatever this might mean.) It is even dubious, I think, to suppose that it makes sense to "know the proposition that p" or even to "say the proposition that p." (When Monica said that Bill is cute, did she "say the proposition that Bill is cute"? This sounds odd, to say the least.)

Thus the relation theory does not really help to explain the apparent validity of inferences like (15). Consider:

(19) Monica says whatever she thinks.
Monica thinks that Bill is cute.
∴ Monica says that Bill is cute.

Certainly, this inference appears valid. But this apparent validity cannot be explained by the relation theory, since as we've seen, the verbs 'think' and 'say', as they appear in (19), simply do not express relations of the relevant sort. Since the logical forms of the constructions 'x thinks that p', 'x says that p', 'x believes that p', and 'x asserts that p' are surely all the same, these constructions should all be given a uniform treatment. Hence the relation theory's treatment of inferences

like (15), which depends on assigning relational senses to the verbs 'believes' and 'asserts', must be incorrect, since this treatment cannot be correctly generalized to other exactly similar inferences like (19).

The problem of how to understand the quantificational constructions in inferences like (15) and (19) is difficult, and one that I cannot adequately deal with here. The problem is in fact quite general, and goes far beyond constructions involving apparent quantificational binding of positions occupied by 'that'-clauses. Consider, for instance:

(20) John will be whatever Mary wants him to be.
Mary wants John to be rich.
∴ John will be rich.

In this case, the apparently bound position follows the verb 'to be'. But the existence of this quantificational construction gives us no reason whatever to believe that the verb 'to be' expresses some sort of relation, or that the adjective 'rich' is really a noun phrase that refers to an abstract entity.

One possibility is that all of these constructions involve *second-order* quantification. On this idea, the quantifiers range over abstract entities, as the relation theory says, but contrary to the relation theory, the proper substituends of the bound variables are sentences and predicates that express such entities, rather than terms that refer to them. Thus the quantified premises of (15) and (20) would be understood as meaning, respectively:

(21) (∀p)(if Mary asserts that p, then John believes that p); and
(22) (∀F)(if Mary wants John to be F, then John will be F),

(where 'p' ranges over propositions, and 'F' ranges over properties.) Another possibility is that the relevant constructions involve some form of substitutional quantification. On this idea, for instance, the quantified premises of the inferences (15) and (20) might be understood as meaning, respectively:

(23) Every instance of the form 'If Mary asserts that p, then John believes that p' is true; and
(24) Every instance of the form 'If Mary wants John to be F, then John will be F' is true.

Again, I cannot stop to examine these possibilities here. I merely want to point out that, in light of the total evidence available, the existence of the types of quantificational constructions in question neither supports the relation theory of cognitive verbs nor provides a convincing argument against the property theory.

7. Conventional implicature.

Grice (1961 and 1975) discovered an aspect of linguistic meaning that can conventionally determine what is implicated by a speaker's use of a sentence, without thereby influencing what is *said* by such a use. For instance, consider a sentence like

(25) Though Kripke is a philosopher, he's smart.[18]

As Grice points out, a sentence like (25) says the same thing, expresses the same proposition, as the more neutral

(26) Kripke is a philosopher and he's smart.

Yet (25) and (26) have different linguistic meanings, due to a difference in the meanings of the connectives 'though' and 'and'. While these connectives both express the same truth function (namely, conjunction), the former's use, unlike the latter's, conventionally implicates that there is some kind of contrast between the two conjuncts in question. Thus unlike (26), (25) implicates without saying that

(27) Kripke's being a philosopher makes it unlikely that he's smart.

Several other connectives, including 'but', 'yet', and 'nevertheless' resemble 'though' in this respect.

Since the sentences (25) and (26) have different linguistic meanings, the property theory predicts that the *de re* cognitive predicates formed by prefixing (25) and (26) with a cognitive operator such as 'believes that' will express different cognitive properties. This prediction is in fact borne out. Consider the following (blatantly fictional) example. The late physicist Richard Feynman seems to have had a low opinion of the intellectual abilities of philosophers. But imagine that after talking to Kripke for a while one evening, he is pleasantly surprised to learn how much physics Kripke knows (perhaps even the basics of Feynman's own theories!). Later, Feynman says (25) to a colleague. In these circumstances, the following belief ascriptions would both be true:

(28) Feynman believes that though Kripke is a philosopher, he's smart.
(29) Feynman believes that Kripke is a philosopher and he's smart.

Yet it seems clear that (28) and (29) say different things about Feynman's beliefs. For (28), unlike (29), doesn't just implicate, but *logically implies* that

(30) Feynman is committed to believing that Kripke's being a philosopher makes it unlikely that he's smart.

Since (28) logically implies (30) and (29) does not, it is possible for (29) to be true and (28) to be false. This fact both confirms the property theory and provides evidence against the relation theory. For on the relation theory, (28) and (29) must have the same truth value, since again, the imbedded sentences (25) and (26) express the same proposition.[19]

Other cases of conventional implicature provide similar evidence in favor of the property theory and against the relation theory. Consider, for instance, the following pairs of nonequivalent belief ascriptions:

(31a) Smith believes that Jones is an Englishman and is, therefore, brave.
(31b) Smith believes that Jones is an Englishman and is brave.[20]
(32a) Sam believes that even Larry was on time for Kripke's paper.
(32b) Sam believes that Larry was on time for Kripke's paper.

In both pairs, the (a)-ascriptions have logical implications about the agent's beliefs not had by the corresponding (b)-ascriptions. Yet, on Grice's view, which I think is correct, the imbedded sentences in each pair express the same propositions. Again the phenomenon of conventional implicature shows that a sentence's linguistic meaning can play a semantic role in cognitive predicates that is inconsistent with the relation theory.[21]

8. The problem of irrational beliefs.

One of the most important pieces of evidence against the relation theory and in favor of the property theory is provided by an objection to the Russellian form of the relation theory that was recently discovered by some of that theory's most prominent defenders. The objection is based on an important type of example found independently by Mark Richard (1983) and Scott Soames (1985), but the objection was first clearly formulated as a problem for the Russellian form of the relation theory by Nathan Salmon (1986a).[22] Salmon considers a set of sentences of the following sort:

(33a) Lois believes that Superman outweighs Clark Kent.
(33b) Lois believes that Superman outweighs Superman.
(33c) Lois believes that Superman outweighs himself.
(33d) Lois believes that Superman is self-outweighing.

For the sake of discussion, I will follow the Russellians in assuming (what may well be false) that the names 'Superman' and 'Clark Kent' do not have descriptive meanings. Given this assumption, the property theory and the relation theory both imply that (33a) and (33b) are semantically equivalent. On the relation theory, of course, (33a) and (33b) are equivalent because they both assert that Lois believes the same singular proposition. And on the property theory, given that these sentences are *de re*, each sentence says of Superman and Clark Kent (i.e., Super-

man) that Lois believes that the former outweighs the latter. On this reading, (33b) does not, despite appearances, ascribe an irrational belief to Lois. For read *de re*, (33b) could be true by virtue of Lois's having a belief that is semantically analogous to a sentence of the form 'n_1 outweighs n_2' where n_1 and n_2 are perhaps distinct names with perhaps different meanings that both refer to Superman/Clark Kent. The idea is that each occurrence of the name 'Superman' in (33b) serves merely to introduce an object that the relevant belief is said to be about. It is no part of the function of either occurrence of the name to indicate a manner of thinking of Superman, and so (33b), understood *de re*, is neutral on the question of whether the ways of thinking in the agent's belief that correspond to the uses of 'Superman' in the ascription are the same or different ways of thinking. (See the Appendix.)

But with (33c) and (33d), irrationality begins to raise its ugly head. (33c) is in fact problematic, since depending on how reflexives like 'himself' work, it is difficult to say whether (33c) should be assimilated to (33b) or to (33d).[23] So let us just ignore (33c) and concentrate on (33d). (33d) clearly ascribes an irrational belief to Lois, so we may assume that it is false, even though both (33a) and (33b) are true. The property theory easily explains the semantic difference between (33a) and (33b) on the one hand, and (33d) on the other. For on the property theory, each of (33a) and (33b) ascribes to Lois a possibly rational belief of the form 'n_1 outweighs n_2', while (33d) ascribes to Lois a belief that is semantically analogous to a sentence of the form 'n_1 is self-outweighing'. (Again, see the Appendix.) Clearly, any belief of the latter sort is irrational, and since Lois has no such belief, (33d) is false.

But the relation theory has a much more difficult time explaining the difference between (33a) and (33b) on the one hand, and (33d) on the other. For it surely seems that, given the direct reference view of names, the imbedded sentences in all three belief ascriptions must express the *same* proposition:

(34a) Superman outweighs Clark Kent.
(34b) Superman outweighs Superman.
(34c) Superman is self-outweighing.

In particular, it seems obvious (to me at least) that (34b) and (34c) say exactly the same thing.[24] But then, the relation theory has the anomalous consequence that (33d) is *true*. Hence again, the relation theory is false.

Naturally, defenders of the relation theory do not agree. Both Salmon (1986a) and Soames (1987a, 1987b) vigorously defend the claim that sentences having the forms of (34b) and (34c) would *not* express the same proposition. Their basis for this claim seems to be the fact that (34c) contains a meaningful part that expresses a propositional constituent—the property of being self-outweighing—that is not expressed by any part of (34b). However, it is far from clear how this difference between (34b) and (34c) is supposed to show that these sentences express different propositions. One possibility is that Salmon and Soames are prepared to endorse some general principle like

(35) If S and S′ both express singular propositions, then S and S′ express the same proposition only if S and S′ ascribe the same properties and relations to the same objects.

However, the strong intuition that (34b) and (34c) say precisely the same thing can reasonably be regarded as providing a counterexample to the principle (35), especially since (35) seems to have no independent motivation.[25] Moreover, there is a very good reason in favor of saying that sentences related as (34b) and (34c) must in fact express the same proposition. The reason is that one-place predicates formed by use of the operator 'self' seem best understood as *defined* by the two-place predicate to which the operator is applied. Consider (34c) ('Superman is self-outweighing'), for instance. Surely, we cannot understand what it is for an object to be self-outweighing unless we understand this to mean precisely that the object bears the relation of outweighing to itself. And in general, the one-place predicate that results from applying the operator 'self' to a two-place predicate seems to be defined as true of an object x just in case x bears to x the relation expressed by the two-place predicate. Thus it surely seems that for any two-place predicate Φ we have the definition

(36) x is self-Φing $=_{df}$ xΦx.

But given this definition, and assuming that the name 'Superman' has the same meaning in both of its occurrences in (34b), it follows that (34b) and (34c) are just *synonymous*. Hence these two sentences must express the same proposition after all, contrary to the claims of Salmon and Soames.[26,27]

In order to cogently object to this argument, the defender of the Russellian theory must either (1) deny that one-place predicates formed by use of the operator 'self' are defined by the two-place predicates to which the operator is applied, thus rejecting the definition (36); or (2) deny that definitions like (36) result in the synonymy of *definiendum* and *definiens*; or (3) deny that synonymous sentences uttered in identical contexts must express the same proposition. But each of these denials would be a desperate and implausible move. Outright rejection of definitions like (36) makes it impossible to explain how applications of the operator 'self' are understood in terms of the predicates to which the operator is applied. Assertion that definitions do not result in synonymous expressions is contrary to the common and plausible understanding of definitions as abbreviatory devices that add nothing new to our conceptual repertoire.[28] And denying that synonymous sentences uttered in the same context express the same proposition violates the fundamental semantic principle that the proposition expressed by a sentence in a context is determined by its linguistic meaning.

In contrast to the relation theory, the property theory gives a simple, straightforward explanation, based on the definition (36), of why a sentence like (33d) ('Lois believes that Superman is self-outweighing') ascribes an irrational belief. On the property theory, (33d) says that Lois has a belief that is semantically

analogous to some sentence of the form 'n_1 is self-outweighing'. This in turn implies (by definition of 'self-outweighing') that Lois's belief is semantically analogous to a necessarily false sentence of the form 'n_1 outweighs n_1', where n_1 has the same meaning in both occurrences. Thus (33d) ascribes an irrational belief to Lois. By contrast, as we've seen, the property theory implies that a *de re* sentence like (33b) ('Lois believes that Superman outweighs Superman') says that Lois has a possibly rational belief of the form 'n_1 outweighs n_2', where n_1 and n_2 are perhaps distinct names with distinct linguistic meanings. Thus the property theory has the desirable consequence that the sentences (33b) and (33d) can have distinct truth values, even though the imbedded sentences (34b) and (34c) express the same proposition.[29]

So it seems to me that the Salmon/Soames response to the problem of irrational beliefs is implausible. Their claims that the relevant sentences in the examples express different propositions are not only counterintuitive, but they require some radically new and as yet unsuggested elaboration of such semantic concepts as those of definition, synonymy, meaning, and proposition. Given the contrasting ease with which the property theory deals with these cases, I suggest that a better response would be to abandon the relation theory in favor of the property theory.

Now that I've presented the property theory in some detail and presented some important new evidence in its favor and against the relation theory, I now need to return to the topic that provided the initial motivation for the property theory, namely, the problem of substitution.

9. Descriptive names and real language.

We have seen that by appeal to the possibility of descriptive names, the property theory can explain how it is at least possible for substitution of coreferring names to fail in cognitive contexts. In this respect, the property theory is superior to the relation theory. But we have yet to see how well the property theory fares in explaining actual cases of apparent substitution failure involving actual proper names. The answer to this question depends in part on how many descriptive names, if any, are found in natural languages, and the fact seems to be that such names, though they exist, are exceedingly rare. (Although in the examples of philosophers, their occurrence is not rare, but frequent.)

One consideration suggesting that most ordinary names have no descriptive meanings is provided by Kripke's (1972) famous Gödel-Schmidt case.[30] Practically the only thing many people have heard about the logician Kurt Gödel is that he discovered the incompleteness of arithmetic. But as Kripke points out, people's uses of the name 'Gödel' would still refer to Gödel even if it had not been Gödel but an unknown Viennese high school teacher named 'Schmidt' who actually discovered incompleteness. Since a similar point can be made regarding all the other achievements for which Gödel is famous, it is clear that the referent of the name 'Gödel' is not determined by any description like 'the discoverer of incompleteness' that might be commonly associated with the name. But then surely, the name 'Gödel' has no descriptive meaning in any public language,

since if it did, there *would* be a commonly associated description that determines its referent.

I have argued elsewhere that the best explanation of our intuitions about the Gödel-Schmidt case and others like it, is that people who know next to nothing about a name's referent would use the name on the basis of a semantic rule that is expressible by use of some "buck-passing" description that involves reference to the speaker, such as 'the one to whom *I've* heard others refer with "Gödel"', or 'the one of whom *I've* heard that he discovered incompleteness'. (See McKinsey 1978a, 1978b, and 1984.) But then again, since such descriptions are private to the speaker, a name's referent is not determined in such cases by any publically associated description, and the name in question would therefore have no descriptive meaning in the public language.

It seems, then, that most of the names used in actual language are not descriptive names. Still, descriptive names do exist, and though they are statistically rare, they are, I think, theoretically important. The common feature of such names is that their referents (if any) are epistemically remote from all speakers in the same way, so that all speakers have to base their reference with the name upon the same narrow set of descriptive assumptions. One excellent example suggested by Kripke (1972, p. 291) is the name 'Jack the Ripper', used to refer to whoever was the murderer of several prostitutes in 1890's London, but about whom nothing else is known. 'Hesperus' and 'Phosphorus' also seem to be good examples, at least as used by the ancient Greeks, since everyone's access to these names' referent was then based solely on visual impressions of the planet in different locations at different times of day. (The English terms 'the Evening Star' and 'the Morning Star' may well be descriptive names that accurately translate the Greek names 'Hesperus' and 'Phosphorus', respectively.[31]) Other good candidates for descriptive names include names of deities like 'Zeus' and 'God'; names of historically remote figures about whom little if anything is known, such as 'Homer' and 'King Arthur'; and pseudonyms, including pen names, at least prior to revelation of the referents' "real" identities.

It is of course always an empirical question as to whether or not a given name has a descriptive meaning in a language. But one good test of whether or not a name is descriptive is provided by the method Kripke used in his Gödel-Schmidt case. In such a test, we consider the description that allegedly determines a name's referent as a matter of its meaning, and ask whether it's possible to use the name in the same way we actually do, to refer to the actual referent, even though some different object satisfies the description in question. If this is *not* possible, then the name is descriptive. Thus consider the absurdity of our asserting such an hypothesis as:

(37) It wasn't Jack the Ripper who murdered all those prostitutes in 1890's London; rather, it was some other man named 'Schmidt'.

Assuming that Jack the Ripper actually existed, there is nothing absurd about the *proposition* expressed by (37). For there is no doubt some possible world in

which Jack didn't commit the relevant murders though someone else named 'Schmidt' did. Rather, what is absurd is our attempting to seriously *assert* this proposition using the sentence (37). For in doing so, we commit ourselves to denying that the referent of 'Jack the Ripper' satisfies the very description that provides the sole basis for determining the actual referent of the name as we use it in (37). Thus if someone named 'Schmidt' was indeed guilty of the murders in question, then he would just *be* Jack the Ripper, and hence (37) would be false. (The falsity of (37) is guaranteed by its own semantics, though it expresses a proposition that is possibly true.)

In sum, the test is this: if you run a Gödel-Schmidt test on a name and a description, and assertion of the relevant hypothesis is absurd, then the name is descriptive. On the other hand, if—as in the Gödel-Schmidt case—no such absurdity is provided by any commonly associated description, then the name is not descriptive. I think it is clear that on application of this test, many ordinary proper names, including those mentioned above, turn out to have descriptive meanings. To give another example, it seems clear that the name 'God' is descriptive, once we apply the Gödel-Schmidt test. For surely, it would be absurd to say, for instance

(38) It wasn't God who created the heavens and the earth by intelligent design; rather, it was some other powerful being named 'Schmidt'.

10. Substitution failure in real language.

Since there really are descriptive names in actual languages, there will be real cases of substitution failure involving coreferring names, and such failures of substitution will be given a semantic explanation by the property theory. The most important such case is that of 'Hesperus'/'Phosphorus'. In fact, the cogency of its explanation of the failure of substitution of these names in cognitive contexts is perhaps the property theory's greatest victory. On the other hand, given the admitted rarity of descriptive names in actual languages, it might appear that this is a pretty hollow victory. Such a conclusion, however, would be premature. For the model provided by the 'Hesperus'/'Phosphorus' case, and the semantic facts that underlie it, go a long way toward explaining our intuitions about cases in which substitution of coreferring names appears to fail, even when it in fact succeeds.

Let us consider an example involving a pen name. Pen names seem to be good candidates for descriptive names, since at least prior to revelation of the referent's "real" identity, a pen name N would seem to have its referent universally determined by a description of the form 'the author of works published under the name N'. For instance, St.-John Perse was a famous French poet who eventually won the Nobel Prize in Literature, but 'St.-John Perse' was the pen name of Alexis Léger, who was at least equally well known for his brilliant career in the French diplomatic service between World Wars I and II.[32] Suppose that in the 1930s (long before the identity was known) Claudette is a student who admires both the

career of Alexis Léger and the poetry of St.-John Perse. Then (39) would be true while (40) is false:

(39) Claudette believes that Alexis Léger is a brilliant diplomat.
(40) Claudette believes that St.-John Perse is a brilliant diplomat.

On the property theory, (39) is a true *de re* ascription, since 'Alexis Léger' is an ordinary name with no descriptive meaning. But (40) is a *de dicto* ascription, since 'St.-John Perse' is a descriptive name, and understood *de dicto*, (40) is false: Claudette has no belief with a semantic property corresponding to the linguistic meaning of the sentence 'St.-John Perse is a brilliant diplomat'.

So it is easy for the property theory to explain substitution failure when a descriptive name is substituted for a coreferring non-descriptive name. However, apparent substitution failure in the opposite direction is more difficult to explain. Consider:

(41) Claudette believes that St.-John Perse wrote *Anabase*.
(42) Claudette believes that Alexis Léger wrote *Anabase*.

Though we may assume that (41) is true, most would no doubt be inclined to say that (42) is false, on the grounds that Claudette is not aware of Alexis Léger's secret identity as a poet. However, both the Russellian theory and the property theory imply that (42) is *de re*, since again, 'Alexis Léger' is surely just a typical ordinary name with no descriptive meaning. But read *de re*, (42) seems true, since Claudette no doubt believes *of* St.-John Perse, that is Alexis Léger, that he wrote *Anabase*.

What then can explain the strong inclination to try to give (42) a *de dicto* reading that it does not have? First, perhaps, we are so inclined because there is nothing in the context that suggests a *de re* reading, that suggests in other words that we are trying to characterize Léger himself, as opposed to Claudette's ways of thinking about him. Since (41) has a *de dicto* reading, we go on to try to read (42) analogously. Second, as I have argued elsewhere, we are strongly predisposed to give *de dicto* readings to cognitive ascriptions whenever we can, because such readings ascribe properties that individuate mental states, and are thus more informative regarding the fundamental semantic properties of such states than *de re* readings. In the absence of any suggestion of a *de re* reading, we thus think of the *de dicto* reading as the "default" interpretation (see McKinsey, 1994, p.320). Third, and most important, the semantics of 'believes that' and the semantics of proper names in general are both, as I have argued, consistent with a sentence like (42)'s *having* a *de dicto* reading. It is entirely natural to think that in (42) we are trying to classify Claudette's belief in terms of the meaning of the sentence 'Alexis Léger wrote *Anabase*' and to think that this meaning is a function of the (descriptive) meaning of the name 'Alexis Léger'. Of course, as Kripke has taught us, the name in question has no such meaning. But an ordinary speaker

could not be expected to know Kripke's point about names. It is in fact quite natural and plausible (though wrong) to assume that a name like 'Alexis Léger' has a descriptive meaning, especially in a case like this, in which a person has two names, each associated with a distinct identity. But once one has made this assumption, the semantics of the operator 'believes that' *requires* one to think that (42) has a *de dicto* reading, and given our predisposition in favor of such readings we will then try to read (42) *de dicto* and hence conclude that it is false.

The sole mistake made by one who tries to read (42) *de dicto* is just the simple and natural mistake of assuming that a particular name has a descriptive meaning when it doesn't.[33] Note that even this mistake is based on the *correct* assumption that names *can* have descriptive meanings. Beyond this point, the mistaken impression that (42) has a *de dicto* reading is then generated by a fully *correct* understanding of the semantics of cognitive operators like 'believes that'. Thus my explanation of speakers' mistaken intuitions regarding apparent substitution failure requires only minimal error, in contrast to the Russellian theory which as we saw earlier, seems to require some kind of radical semantic incompetence.[34]

11. The "no proposition" problem.

The final sort of consideration that I wish to offer against the relation theory and in favor of the property theory is based on two general types of cognitive ascription that can be true, even though the sentence in the scope of the cognitive operator fails to express a proposition. The first type of ascription, one that I alluded to above in section 5, is that of cognitive ascriptions containing small-scope descriptive names that may lack referents, such as:

(43) Al believes that God is omniscient.

Atheists and agnostics would be perfectly willing to say something like (43), confident that they had made a true statement about Al's beliefs, independently of the question of whether God exists. Hence it surely seems that (43) could be true, even if God does not exist. But since the name 'God' is directly referential, the sentence 'God is omniscient' expresses no proposition unless God exists. Thus on the relation theory, (43) cannot be true unless God exists. So again, the relation theory is false.

Sentences like (43) pose no problem for the property theory, according to which (43) ascribes to Al a belief that has the semantic property S_M, where S_M corresponds to the linguistic meaning of the imbedded sentence 'God is omniscient'. Since this sentence has its meaning independently of whether the name 'God' has a referent, the property theory implies that (43) can be true, even if God does not exist.

A defender of the relation theory might reply that while (43) literally has only a *de re* reading, we would be loathe to interpret an atheist or agnostic as intending such a reading, since doing so would commit them to believing that God exists.

Hence we conclude that they must have a non-literal, *de dicto* interpretation in mind when they say (43). However, this reply ignores the fact that a theist would be happy to assert (43) even if taken *de re*, and so the reply has the consequence that an atheist and a theist would not both be able to use (43) both literally and with the intention of saying the same thing. But surely, this is wrong. Surely, in fact, both atheist and theist can happily agree that Al believes that God is omniscient, and so both could use (43) in a sense that commits neither to the existence of God.

Since there are, I believe, many descriptive names in actual languages, there is a wide range of cases similar to (43) that can be used to make the same point. (Just think of belief ascriptions containing such names as 'Santa Claus', 'King Arthur', and 'Homer', for instance.) But because the existence of descriptive names may still seem controversial to some, it will be useful to consider a different type of case that I've described elsewhere (McKinsey, 1986) involving not names but anaphoric pronouns. This type of case also raises a "no proposition" problem for the relation theory.

Consider such sentences as

(44) Oscar wishes he had caught the fish that got away.

I wish to restrict my attention to readings of (44) on which it does not imply that any fish actually did get away from Oscar. (What "got away" might have been an old boot or a rock.) What is the content of the wish that such a reading of (44) would ascribe to Oscar? It cannot be a proposition of the form

(45) Oscar caught at t the fish that got away from him at t,

since such a proposition would be contradictory, while the wish that (44) ascribes to Oscar is clearly consistent. My suggestion is that we should adopt a proposal made by Peter Geach (1967) for understanding similar cases, and take (44) as meaning

(46) Oscar assumes that just one fish got away, and Oscar wishes it had been the case that he caught it.

Contexts of this sort, in which a pronoun in the scope of one cognitive operator refers back to a quantifier antecedent in the scope of a different cognitive operator, I call contexts of "mental anaphora". (46) has the advantage of implying neither that any fish actually got away from Oscar nor that Oscar's wish is inconsistent, and so (46) seems to capture the most natural way of reading (44).

However, as Geach pointed out, (46) raises a new problem, namely, that of how to understand the anaphoric pronoun 'it' as it occurs in (46). I think that in many of Geach's cases that are structurally like (46), the anaphoric pronoun is going proxy for a definite description that is recoverable from the pronoun's

quantifier antecedent. (See McKinsey 1986, pp. 162–163.) But I've chosen (46) precisely because this cannot be so: again, if 'it' were proxy for 'the fish that got away (from Oscar at t)', then (46) would ascribe an inconsistent wish to Oscar, which it does not do. Nor can 'it' be a variable bound by its quantifier antecedent 'just one fish', for in (46) 'it' is not in the *scope* of this antecedent. (The quantifier 'just one fish' lies in the scope of 'assumes that', whose scope does not extend beyond (46)'s first conjunct.)

But if the second occurrence of 'it' in (46) is neither a bound variable nor proxy for a definite description, then how is the pronoun functioning? I think that the only possibility is that, in this context, 'it' is functioning as what Gareth Evans (1977) called an "E-type" pronoun, a rigid genuine term whose referent is fixed by the description recoverable from its quantifier antecedent. Consider the fact that Oscar could express the very wish ascribed to him in (46) by uttering the words

(47) Just one fish got away. Would that it had been the case that I caught it.

It seems clear that the final occurrence of 'it' in (47) is being used as an E-type pronoun. For in uttering (47) Oscar is attempting to express his wish that he had caught a certain fish, and so what would make his wish come true is the truth of the singular proposition that he caught *that very fish*. Thus, in uttering (47), Oscar is using 'it' as a directly referring genuine term whose referent is fixed by the description 'the fish that got away'.[35]

(46) ascribes to Oscar an assumption and a wish based on that assumption. The ascription is accomplished by using the words of (47), words that Oscar could use to express the very same wish. So it surely seems that in its final occurrence in (46), the pronoun 'it' must have the same meaning as it has in (47), and so in (46) it must also be an E-type pronoun.

Still, there is a stark contrast between the ways the pronoun 'it' functions in (46) and (47). For in (47), the pronoun has the normal function of a genuine term, namely, that of introducing a referent into what is said by the speaker, while in (46), the same pronoun, with the same meaning, clearly does *not* introduce a referent into what is said by the speaker. For again, (46) can be true, even if Oscar's assumption that just one fish got away is false; and if this assumption is false, then the second occurrence of 'it' in (46) has no referent. In these circumstances, the sentence 'he caught it', as it occurs in (46), would therefore express no proposition. And yet (46) would still be *true*. Once again, we have a case of a cognitive ascription that is true, even though the sentence in the scope of the cognitive operator expresses no proposition. And so again, we have a counterexample to the relation theory.

It is worth emphasizing the similarity between the semantic behavior of E-type pronouns in the context of mental anaphora, and the semantic behavior of descriptive names in the scope of cognitive operators. A descriptive name like 'God', for instance, would normally be used in a non-cognitive sentence like 'God is omniscient' to introduce a referent into the (singular) proposition that the speaker

wishes to assert. If no such referent exists, then the sentence simply expresses no proposition, and the sentence is thus neither true nor false. But when such a term occurs in the scope of a cognitive operator, as in a sentence like (43) ('Al believes that God is omniscient'), the name's descriptive meaning becomes the primary contribution to what is said by the whole sentence, the name's referent drops out as irrelevant, and the sentence can express a true proposition even if the name has no referent. As we've just seen, a similar change in function happens when an E-type pronoun occurs in the context of mental anaphora.[36]

The semantic similarity between descriptive names and E-type pronouns allows cognitive contexts containing the one sort of term to shed light on the meaning of cognitive contexts containing the other. In a cognitive ascription containing a descriptive name, we use the name to represent (but not refer to) a mental act whose reference is fixed in a manner analogous to the way the name's reference is fixed by its descriptive meaning. But then, for any such cognitive ascription, we can more explicitly reveal what is said by using mental anaphora to express the relevant descriptive assumption. Thus, suppose that the reference of the name 'God' is fixed (as a matter of its meaning) by some such description as 'the being who created the heavens and the earth by intelligent design'. Then the belief ascription (43) can be understood to say exactly the same thing as the following instance of mental anaphora:

(48) Al assumes that just one being created the heavens and the earth by intelligent design, and Al believes that it (that very being) is omniscient.

On the other hand, we can use the concept of a proper name's descriptive meaning to provide a clear account of the cognitive properties ascribed in contexts of mental anaphora. Thus suppose that S is a sentence prefixed by a cognitive operator such as 'thinks that', and that S contains one or more occurrences $e_1 \ldots e_n$ of E-type pronouns, each of whose antecedent quantifier phrases does not occur in S but does occur in the scope of an affirmative cognitive operator, so that $e_1 \ldots e_n$ all occur in the context of mental anaphora.[37] On the assumption that S contains no occurrences of genuine terms lacking context-independent meanings other than $e_1 \ldots e_n$, the property ascribed by the predicate 'thinks that S' can be specified as follows.[38] Let S* be a possible sentence obtained from S by replacing each occurrence e_i of $e_1 \ldots e_n$ by a (possible) descriptive proper name whose reference is fixed by a semantic rule expressible by use of the description recoverable from e_i's quantifier antecedent. Then 'thinks that S' ascribes the property of having a thought that has the semantic property S^*_M, where S^*_M corresponds to the linguistic meaning of S*, as explained above in section 5.

12. Contextual theories of belief.

In recent years, certain "contextual" variations on the relation theory have been widely discussed. Like the relation theory, these views also say that 'believes' and other cognitive verbs express relations that hold between persons and prop-

ositions (and perhaps other things as well). But in an effort to accommodate real substitution failure, these views allow cognitive ascriptions to have additional, contextually determined semantic features. Perhaps the most important of these views is the "hidden-indexical" theory, first proposed by Stephen Schiffer (1977 and 1987), and also proposed in a slightly different form by Crimmins and Perry (1989). (Schiffer (1992) has since become a forceful critic of his theory.)

On the hidden-indexical theory, a sentence of the form 'x believes that p' expresses a proposition of the form

(49) (∃m)(Φ∗m & x believes the proposition that p under m),

where 'm' ranges over modes of presentation, and as Schiffer puts it, "Φ∗ is an implicitly referred to and contextually determined type of mode of presentation" (1992, p. 503). The main advantage of this type of view is that it allows for substitution failure. Thus, utterances of 'Lois believes that Superman can fly' and 'Lois believes that Clark Kent can fly' can have different truth-values, because these utterances can say that Lois believes the proposition that Superman can fly under distinct types of modes of presentation.

But while it may overcome the substitution problem, the hidden-indexical theory nevertheless succumbs to all of the other difficulties we've raised for the relation theory. Thus, the hidden-indexical theory is refuted by the "no proposition" problem, since like the relation theory, it also implies that a belief ascription is true only if the ascription's imbedded 'that'-clause expresses a proposition. But the cases I've cited involving descriptive names and mental anaphora show that this consequence is false.

Also, cases involving irrational beliefs raise the same problem for the hidden-indexical theory as for the relation theory. In order to accommodate purely *de re* belief ascriptions, the hidden-indexical theory must allow for limiting cases in which a belief ascription says merely that an agent believes a certain proposition under some mode of presentation or other. (See Schiffer, 1992, pp. 503–504. In such a case, we can let the relevant instance of Φ∗ be a property trivially had by all modes of presentation, such as the property of being self-identical.) Now suppose that in a given context, a speaker uses (33d) to make a purely *de re* ascription:

(33d) Lois believes that Superman is self-outweighing.

Then, on the hidden-indexical theory, this use of (33d) is true, provided that Lois believes the proposition that Superman is self-outweighing under some mode of presentation or other. But the latter can be true, even if all of Lois's beliefs are completely rational. Thus it might be true in the same context that

(33a) Lois believes that Superman outweighs Clark Kent,

where this is true because Lois believes the proposition that Superman outweighs Clark Kent under a mode of presentation involving distinct modes of presentation that she associates with the names 'Superman' and 'Clark Kent'. But then, since the proposition that Superman outweighs Clark Kent just *is* the proposition that Superman is self-outweighing, it follows from the hidden-indexical theory that when (33a) is true for this reason, (33d) is also true *de re*. But this consequence is clearly false: even understood *de re*, (33d) is false, because it ascribes to Lois an irrational belief that she just does not have.

The reason why ascriptions of irrational beliefs like (33d) pose the same problem for the hidden-indexical theory as for the relation theory, is that in the limiting case of purely *de re* ascriptions, the hidden-indexical theory just reduces to the relation theory. And for the same reason, my earlier argument based on conventional implicature refutes the hidden-indexical theory as well as the relation theory. Thus suppose a speaker uses (50) to make a *de re* assertion about me:

(50) McKinsey believes that though Kripke is a philosopher, he's smart.

Now I insist that this use of (50) would just be a plainly false report of my beliefs. Yet on the hidden-indexical theory, (50) would have to be *true*, since the *de re* ascription (51) would be true:

(51) McKinsey believes that Kripke is a philosopher and he's smart.

Since (51) is true, it follows (according to the hidden-indexical theory) that I believe the proposition that Kripke is a philosopher and he's smart under some mode of presentation or other. But then, since this proposition just *is* the proposition that though Kripke is a philosopher, he's smart, it follows that (50) must be true too. But it's not.

In addition to the problems it shares with the relation theory, the hidden-indexical theory has serious problems of its own, some of which have been forcefully pointed out by Schiffer (1992). Given all of these problems, I think we are justified in concluding that this type of contextual theory is false.

A quite different sort of contextual theory has been proposed by Mark Richard (1990). Richard agrees with defenders of the relation theory that the cognitive verbs all express relations. But there are some crucial differences. First, the objects of belief are not propositions as standardly understood. Rather, they are "fusions" of the constituents of Russellian propositions with words that refer to those constituents. Consider a belief ascription like 'Lois believes that Superman can fly'. On Richard's view, the 'that'-clause in this ascription determines an abstract pair of pairs, each of which consists of an expression in the 'that'-clause together with its referent:

(52) ⟨⟨'can fly', being able to fly⟩, ⟨'Superman', Superman⟩⟩

Richard calls such fusions 'RAMs' (for 'Russellian annotated matrices'). Then the theory is that a belief ascription is true (in a given context) if and only if the RAM determined by its 'that'-clause gives the "right" translation (in the context) of a RAM in the believer's representational system. (Richard 1990, pp. 136–142.) Note that on this view, what counts as a "right" translation will vary from context to context. Thus, Richard treats cognitive verbs like 'believes' as indexicals that express different relations relative to distinct contexts.

Like the hidden-indexical theory, the main motivation behind Richard's theory seems to be that of overcoming the substitution problem. But also like the hidden-indexical theory, Richard's view shares one of the major defects of the relation theory: it runs afoul of the "no proposition" problem. This is because unless the sentence in the scope of 'believes' expresses a proposition, it also will not determine a RAM. For instance, unless God exists, there is no such RAM as

(53) ⟨⟨'is omniscient', being omniscient⟩, ⟨'God', God⟩⟩

Thus on Richard's theory, a belief ascription like (43) ('Al believes that God is omniscient') cannot be true unless God exists and its 'that'-clause expresses a proposition. But again, this consequence seems obviously false.

Does Richard's theory have problems with examples involving irrational beliefs and conventional implicature, in the limiting case of purely *de re* ascriptions? In Richard's theory, the limiting case is obtained by allowing that in some contexts there are no restrictions on which RAMs can provide accurate translations of a believer's RAM, other than that the two RAMs "amount to the same Russellian proposition" (Richard 1990, p. 138; see also p. 136). However, Richard's account seems to also require that two RAMs must contain the same number and types of expressions, if one is to "represent" the other. So it may well follow from his view that the RAM determined by 'Superman outweighs Clark Kent' could not be represented by the RAM determined by 'Superman is self-outweighing'. If so, then Richard's theory apparently avoids problems raised by irrational beliefs.

However, the problem raised by conventional implicature does seem to arise. For with no additional "translation" restrictions, the RAM determined by 'Kripke is smart though he's a philosopher' does, on Richard's view, represent the RAM determined by 'Kripke is smart and he's a philosopher' (since 'though' and 'and' refer to the same truth-function). Thus Richard's theory has the false consequence that a sentence of the form 'A believes that Kripke is smart though he's a philosopher' is true in a context with no restrictions, if the referent of 'A' has the RAM determined by 'Kripke is smart and he's a philosopher' in her representational system.

Richard's theory also faces at least one other serious problem that is unique to it. Since every RAM is partly composed of linguistic expressions, it is a consequence of Richard's theory that every belief ascription refers to some of the words that occur in it, and hence is partly quotational. But I think that this consequence

is pretty obviously false. It is no doubt true that many belief ascriptions are quotational, even when they contain no explicit forms of quotation.[39] But if *every* belief ascription is quotational, then it follows that no two belief ascriptions in different words or from different languages could ever express the same proposition, since each ascription would refer to words that the other does not refer to. But surely this is false. Surely, for instance, such pairs of sentences as (54) and (55) not only can, but often do in fact express the same proposition:

(54) Larry believes that snow is white.
(55) Larry glaubt, dass Schnee weiss ist.

Richard is of course aware that pairs of belief ascriptions like (54)-(55) must, on his view, express different propositions. He not only endorses this consequence, but he endorses the much stronger view that *every* pair of distinct sentences, of whatever languages, must express distinct propositions. (Richard 1990, p. 154.) This in turn is because he decides to *identify* propositions with RAMs (though I cannot see that he gives any cogent reason for doing so). As a terminological decision, it seems to me that Richard's proposal has nothing to recommend it. If we adopt a sense of 'proposition' according to which it is impossible for any two sentences to ever express the same proposition, then we take away from the concept of a proposition its primary form of application, both depriving it of its main function in semantics and removing whatever intuitive content it might have had.

Of course the fact that it is a trivial consequence of Richard's idiosyncratic terminology that no pairs of sentences like (54)-(55) ever "express the same proposition", is of no interest. What is important is that in the *standard* sense of 'proposition', it is a consequence of his view that no two belief ascriptions containing different words can ever express the same proposition. And again, I take this consequence to be obviously false.[40]

From my point of view, 'contextual' theories like Schiffer's and Richard's are mere complications, or epicycles, of a type of theory that is false in principle. These epicycles are false for their own unique reasons, but more fundamentally, they are false because they share the basic false assumption of their parent theory, namely, the assumption that 'believes' and other cognitive operators, taken in their basic senses, express relations. In this paper, I have presented what I take to be an overwhelming array of evidence against this assumption. I have also attempted to describe the basic features of an alternative view, the property theory, that I believe provides an adequate account of all this evidence. On the property theory, the predicates formed by cognitive operators like 'believes that' are used to semantically classify cognitive acts and states, where these semantic classifications are primarily based, not on the concept of a proposition, but on the concept of linguistic meaning.[41]

APPENDIX: *The Semantics of* De Re *Cognitive Predicates*

By a "*de re*" cognitive predicate, I mean a predicate of the form 'Cs that S' that is *semantically de re*, even though I assume that the imbedded sentence S is entirely in the scope of the relevant cognitive operator 'Cs that' (so that the predicate is structurally, or logically, *de dicto*). What makes such a predicate *de re* is that the imbedded sentence S contains n occurrences $o_1 \ldots o_n$ of m genuine terms ($m \leq n$), each of which is either a variable free in S, a proper name with no descriptive meaning, a deictic indexical or demonstrative, an anaphoric indexical whose antecedent is a genuine term, or a rigid E-type pronoun or demonstrative whose antecedent quantifier phrase is not in the scope of any cognitive or modal operator. Let each of the objects $x_1 \ldots x_n$ be either the referent of the corresponding occurrence o_i in S, or the referent relative to an assignment *a* of objects to the variables, if o_i is an occurrence of a variable. (If any occurrence o_i *lacks* a referent, then the relevant predicate 'Cs that S' expresses no property.)

A mental act or state *A* will be said to be *about* an n-tuple of objects $\langle x_1,\ldots,x_n\rangle$ just in case *A* is individuated by a semantic property S_M, and there is an n-place relation R^n such that S_M (together perhaps with facts about the world) determines the singular proposition that $<x_1,\ldots,x_n>$ bear R^n to each other. (Again, a semantic property S_M *determines* a proposition p just in case any sentence that had S_M would express p.)

Now consider the (perhaps infinite) class *I* of *instances* of the relevant sentence S, each of which is obtained from S by replacing each occurrence o_i by an occurrence of a proper name n_i that has the same referent as o_i. Here, each n_i is not assumed to be a name of English, but rather may be any *possible* name of any possible extension of English. It is allowed that distinct occurrences of the same term in S may be replaced by distinct names, and while the names n_i may have (and will be assumed to typically have) descriptive linguistic meanings, this is not an absolute requirement (for a reason to be given below). The so-defined class *I* of instances of S determines a class M_I of semantic properties, each of which corresponds to the linguistic meaning (in the relevant extension of English) of an instance of S in *I*.

Then we may say in the present kind of case that a cognitive predicate of the form 'thinks that S' ascribes (relative to an assignment *a*) the *de re* property of having a thought that both (1) is about $\langle x_1,\ldots,x_n\rangle$, and (2) is individuated by some semantic property in the class M_I. Here, the class of semantic properties M_I is determined by the class *I* of instances of S, as defined above. A predicate of the form 'believes that S' in this sort of case ascribes the property of having a belief that disposes a person to have (in certain circumstances) thoughts satisfying the conditions (1) and (2) just mentioned.

While the theory just stated is admittedly a bit complicated, the basic idea behind it is simple and corresponds to the traditional way of understanding *de re* ascriptions. The idea is that a *de re* belief ascription does not specify any *particular* belief of the agent's. Rather, such an ascription says that the agent has a belief of a certain sort that is about a specified object or objects. For example, on the theory just stated, a *de re* ascription such as 'Ralph thinks that Ortcutt is a spy' says that Ralph has a thought which (1) is about Ortcutt, and (2) is individuated by a certain sort of semantic property. In this case, the relevant sort of property is determined by the class of sentences of the form 'n_1 is a spy', where n_1 refers to Ortcutt. Thus, while the ascription does not mention any sentence or form of sentence, it says in effect that Ralph has a thought of the form 'n_1 is a spy', where n_1 is a name that refers to Ortcutt. Similarly, a *de re* belief ascription like (36b) ('Lois believes that Superman outweighs Superman') says in effect that Lois has a belief about the pair ⟨Superman,

Superman⟩ of the form 'n_1 out-weighs n_2' where each of n_1 and n_2 is a name that refers to Superman.

Different specific theories of *de re* ascriptions can be obtained by adding semantic restrictions on the types of names used to define the instances of an imbedded sentence S that in turn determine the relevant class of semantic properties for a given predicate of the form 'Cs that S'. For example, causal theorists might want to restrict the names in the instances to those whose meanings can be given in causal-theoretic, rather than in descriptive, terms. For reasons given elsewhere, I would disagree with such a restriction. (See McKinsey, 1978a, 1978b, 1983, and 1984.) I prefer a view on which the names in the relevant instances are all one of two types: either they have descriptive meanings, or they have no meanings of any sort, only referents.

I have to allow that a belief could be individuated by a semantic property corresponding to the linguistic meaning of a sentence that contains a name with no meaning, for the following reason. Suppose that I believe that I am hungry. Then I have a belief about myself of the form 'n_1 is hungry'. But notice that it would be wrong to require that n_1 be a descriptive name that refers to me, since I of course do not have to refer to myself on the basis of any description, but can refer to myself directly, from a first person perspective. In this case, it seems to me that my belief is of the form 'n_1 is hungry', where n_1 has no descriptive meaning, but only a referent, namely me. In my theory of *de re* ascriptions, then, I must allow that a sentence S containing a proper name n_1 will have a specific linguistic meaning, even when the name n_1 itself has no meaning. In such a case, if S* comes from S by substituting another name n_2 for n_1 in S, then I will say that S and S* have the same meaning if and only if n_2 also has no meaning but n_1 and n_2 have the same referent.

Suppose that a given thought is individuated by a semantic property corresponding to the meaning of a sentence of the form 'n_1 is hungry', where n_1 is a name with a referent, but no meaning. Such a thought, we might say, would be "irreducibly *de re*", since it would be individuated in part by a certain object, the referent of n_1. It is a matter of some controversy as to whether thoughts are often, or ever, irreducibly *de re* in this sense. Using the phenomenon of mental anaphora, I have argued elsewhere (McKinsey, 1994) that our thoughts about ordinary, external objects cannot be individuated in terms of those objects, and so are not irreducibly *de re* with respect to any such objects. Thus, thoughts about external objects must be individuated by semantic properties that correspond to the linguistic meanings of sentences containing descriptive names that refer to those objects. However, I would allow that thoughts about what Russell called "objects of acquaintance", such as oneself, could be irreducibly *de re* in the sense explained. (See McKinsey, 1991.)

Notes

[1]Here and below, I count as "indexicals" all personal, impersonal, and demonstrative pronouns, as well as demonstrative descriptions of the form 'that F'.

[2]See for instance, Donnellan (1974), Kaplan (1977), McDowell (1977), Evans (1979), McKay (1981), McKinsey (1984), Salmon (1986), and Soames (1987).

[3]I call this theory "Russellian", since Russell held the relation theory and also emphasized the importance of genuine terms and singular propositions (see Russell (1904) and (1912)). But note that Russell denied the direct reference view of ordinary names, and so he would have rejected the theory of belief that I'm calling "Russellian".

[4]This was raised as a problem for the Russellian theory by Plantinga (1978), McKinsey (1978), and Ackermann (1979).

[5]A related significant strength of the Russellian theory is that, by allowing cognitive predicates containing small-scope names and indexicals to have *de re* readings, the theory explains how quantification into opaque contexts can be meaningful and thus provides a cogent solution to Quine's (1956) problem about quantifying-in. The alternative to the Russellian theory that I propose below in section 5 also has this advantage. In McKinsey (1998), I argue at length that, contrary to Quine's thesis, a correct semantics of cognitive verbs should imply that quantifying into opaque contexts is meaningful.

[6]Berg (1988) comes closest to providing a Gricean account. Stephen Schiffer (1987) provides a useful and detailed criticism of one possible attempt to use Gricean implicature to explain away one kind of apparent substitution failure. Although the type of Gricean explanation that Schiffer criticizes is different from the one I'm discussing here, my criticism below of the standard Russellian explanation owes much to Schiffer's discussion.

[7]Soames (1987b, pp. 118–119) seems to believe that there must always be some particular semantic or pragmatic "excuse" for deviating from the believer's words. Thus in the cases proposed by McKay and Soames in the previous paragraph, Soames would appeal to the convention that one should refer to oneself only by use of the first person pronoun, to explain why deviation from the believer's words is required. Similarly, in my transvestite example, Soames might try to appeal to the fact that the referent is male to explain why the speaker is required to use the masculine pronoun 'he' and thus must deviate from the agent's words. Perhaps, then, Soames would disagree with my more general explanation, according to which *de re* ascriptions invariably allow us to deviate from the believer's words, and would perhaps insist that some more specific excuse is always necessary. In fact, however, there often seems to be no such specific excuse. In my transvestite case, for instance, the speaker after all *could* have used words that the believer Jones would accept, for the speaker could have said 'Jones believes that that person is a woman'. Yet in spite of this, we find the sentence 'Jones believes that he is a woman' or even 'Jones believes that that transvestite is a woman' to be perfectly acceptable ascriptions. Here there really seems to be no "excuse" for departing so radically from words that the believer would accept, other than the mere fact that the ascription is intended to be *de re*.

[8]A referee for *Nous* objected to my argument on the grounds that a defender of the Russellian explanation can just say that in cases like Berg's, the normal implicature that the speaker would accept the imbedded sentence is simply *cancelled* by knowledge of the context, just as conversational implicatures can always be cancelled, on Grice's view. But this objection falsely assumes, contrary to Grice's view, that a sentence like (12) can have conversational implicatures independently of any literal meaning that it might be taken to have. Again, my point is that when (12) is understood to have the literal *de re* meaning that according to the Russellian theory it must have, its use simply fails to have the relevant implicature. This failure cannot result from implicit cancellation of an implicature that use of the sentence, given its literal meaning, would normally have; for given its literal meaning, use of the relevant sentence always fails to have any such implicature that could be cancelled. Hence, in order to have such a cancellable implicature at all, the sentence must first be taken to have a *different* (*de dicto*) meaning. Of course, the fact that speakers ever take the sentence to have this different meaning in the first place cannot be explained by appeal to an alleged implicature that use of the sentence only has when it is taken to have this different meaning!

[9]I admit that my discussion here does not prove this. Perhaps a Russellian might come up with some other pragmatic explanation that does not have to assume that the relevant cognitive ascriptions are being read *de dicto*. However, every pragmatic explanation of apparent substitution failure that I have heard, read, or thought of myself does make this assumption. In every such explanation, the alleged false implicature that purportedly explains the apparent substitution failure simply does not arise when it is kept firmly in mind that the relevant sentence is being understood *de re*.

Consider for instance the following Russellian explanation, which might seem fairly plausible. In cases of apparent substitution failure of the 'Clark Kent'/'Superman' variety, there are two names of

the same object, and the two names are (psychologically, if not semantically) associated with distinct identities or "dossiers" of information. A speaker's decision to use the name 'Superman' as opposed to the name 'Clark Kent' in a sentence like (12) ('Lois believes that Superman is a reporter') strongly suggests that it is the 'Superman'-dossier that is relevant to what the speaker wishes to convey. Thus a use of (12) would have the false implicature that Lois believes that Superman is a reporter, under her 'Superman'-dossier. Thus (12) seems false.

I think it is quite clear that this explanation works only if the relevant sentence (12) is being tacitly read *de dicto*. When, on the contrary, it is kept firmly in mind that the sentence is being understood *de re*, as in the scenario described by Berg, then it is also understood that the speaker is using the name 'Superman' solely to introduce its referent, and *not* to suggest anything about the particular ways in which the believer might think of the referent. In the absence of any such suggestions, I take it, a use of (12) will simply have no relevant false implicatures.

Again, I cannot prove that no pragmatic explanation of the Russellian sort will work, but the evidence I've given makes it highly probable that this is so. (My thanks to Jane Heal, whose question led me to see the need for this note.)

[10]Kripke (1972) stresses the possibility that while names are not synonymous with descriptions, their referents might be fixed by description. See also McKinsey (1984, 1986, and 1994).

[11]In general, I identify the linguistic meaning of a genuine term, whether the term is a proper name or an indexical pronoun, with a semantic rule that determines the term's referent. The rules that govern proper names determine a name's referent independently of context, while the rules that govern indexicals are token-reflexive. On the other hand, I identify a term's *propositional* meaning, or contribution to the proposition expressed, with the term's referent. See McKinsey (1984), pp. 491–498, and p. 512, note 13; see also McKinsey (1987), pp. 16–19. My distinction between linguistic and propositional meaning corresponds roughly to Kaplan's (1977) distinction between character and content.

[12]Here and below, I often use single quotes to enclose expressions that contain metalinguistic variables, as in " 'thinks that S' ". In such cases, the quotes are understood as quasi-quotes, or Quine's corners, and do not form names of the expressions within the quotes. Context should suffice to resolve this ambiguity.

[13]I will provide a fuller discussion of the properties of cognitive predicates containing descriptive names below in section 11. I also give a detailed account of such predicates and their important implications for the philosophy of mind in McKinsey (1994). As I point out there, on the assumption that 'Hesperus' is a descriptive name, my view implies that it can be true in a given possible world w that a person A believes that Hesperus is F, even though in w, A's belief is not about any object, so that the belief has no proposition as its content (this would happen if in w no object uniquely satisfies the description that actually fixes the reference of 'Hesperus'). Moreover, a person might, in a given world w*, believe that Hesperus is F, even though in w* this belief is about a different object than in the actual world (this would happen if in w* a planet other than Venus—Mars, say—uniquely satisfies the relevant description). Keep in mind that, on the assumption that 'Hesperus' is a descriptive name, it is a consequence of my view that an ascription of the form 'A believes that Hesperus is F' must be understood *de dicto* as opposed to *de re*.

A referee for *Noûs* correctly pointed out that these features of my view imply that there could be a possible world in which the following three propositions are jointly true:

(a) John believes that Hesperus is F.
(b) Hesperus is not F.
(c) Everything John believes is true.

The referee suggested that this consequence is paradoxical, but in my view there is no paradox. In a world w* in which (a), (b), and (c) are jointly true, the belief that we (in the actual world) ascribe by use of (a) would not be about Hesperus (Venus), but about a different planet, Mars, say. Thus the belief ascribed by (a) would in w* not have the proposition that Hesperus (Venus) is F as its content. Rather,

the belief would have as its content the proposition that John would refer to in w* as "the proposition that Hesperus is F", and this would of course be the proposition that *Mars* is F. Thus my view implies that in w* it is *false* that

(d) John believes the proposition that Hesperus is F.

So I would maintain that, from our point of view in the actual world, (a) is perfectly consistent with the conjunction of (b) and (c). What is inconsistent with this conjunction is not (a), but rather (d), which my view implies would be *false* in the circumstances. From my perspective, the referee has just mistakenly conflated the non-equivalent ascriptions (a) and (d). The admitted plausibility of this conflation is one of the main motivations behind the relation theory. But it is a mistake, nevertheless. For further discussion and defense of my view that the constructions 'x believes that p' and 'x believes the proposition that p' are not equivalent, see the final two paragraphs of this section and the whole of section 6 below. In fact, of course, all of the evidence that I present in this paper against the relation theory is also evidence against the equivalence of these two constructions.

[14]For a fuller account see McKinsey (1994), pp. 315–321.

[15]For an argument that beliefs are not always individuated by the proposition believed, see McKinsey (1994).

[16]See, for instance, Schiffer (1992), pp. 504–505.

[17]See Bach (1997), p. 223.

[18]This is a variation on an example suggested by Böer and Lycan (1980), p. 455.

[19]Dorothy Edgington pointed out to me that the kind of contrast implicated by use of 'but' or 'though' can vary depending on context. In her example, our aim is to construct a list of dumb philosophers. Someone says, "Well, Kripke is a philosopher, but he's smart." Here the contrast implicated is not between the conjuncts, but rather between the conjunction and the goal of finding a dumb philosopher. A referee for *Noûs* made the same point with the example of someone who says, "Though Kripke is a philosopher, he's smart," by way of giving a counterexample to the foolish claim that no philosophers are smart.

I am inclined to think that this correct point does not affect my contention that sentence (28) logically implies sentence (30), an implication that still seems intuitively to hold, whatever the context happens to be. I would suggest that application of a cognitive operator to a conjunction formed by 'but' or 'though' has the effect of "sealing off" the conjunction from other possible features of the context, and thus forces the implicated contrast to be one between the conjuncts of the imbedded sentence. But even if I'm wrong about this, my basic point remains unaffected: a belief sentence like (28) will still logically entail commitment by the believer to *some* implicated (contextually salient) contrast. And of course the existence of *any* such entailment is inconsistent with the relation theory and supports the property theory. Moreover, there are many other Gricean examples, like (31) and (32) below, where the relevant conventional implicatures do not vary from context to context, and which can also be used to show that linguistic meaning is relevant to cognitive ascriptions in a way that is inconsistent with the relation theory.

[20]This example is Grice's (1975, p.25).

[21]In reply to this argument, it might be objected that Grice's view of conventional implicature is just mistaken. It might be claimed, for instance, that (25) and (26) do not really express the same proposition, since (25) but not (26) expresses the following triple conjunction:

(i) Kripke is a philosopher and he's smart, and the former fact makes the latter unlikely.[42]

(This possible objection was suggested to me independently by Mark Huston and Lawrence Powers.) But in support of Grice's view, notice that the following is intuitively true in the Feynman example:

(ii) Feynman was surprised to discover that though Kripke is a philosopher, he's smart.

Since one can't be surprised to discover falsehoods, the truth of (ii) implies that (25) is true (in our example). Yet (i) is of course false, since (surely!) its third conjunct is false. So contrary to the

objection, (25) and (i) don't express the same proposition. Moreover, since (25) is true (in spite of having a false implicature), the true proposition it expresses must just be the proposition expressed by the conjunction (26), as Grice's view says.

[22]See also Soames (1987a, 1987b) and McKay (1991).

[23]See the discussions of reflexives by Salmon (1986a, 1992), Soames (1990, 1994), and McKay (1991).

[24]Here and below, I ignore the fact that the relevant sentences may be tensed, and so may express different propositions when uttered at different times. For the sake of discussion, just assume that any tensed sentences are all uttered at the same time, or are being evaluated relative to the same time.

[25]Indeed, when I was a graduate student at Wayne State in the 1960s, I was taught that other pairs of sentences related to each other as (34b) and (34d) do provide counterexamples to the principle (35). As I remember, the sentences used to illustrate this point were 'Socrates is identical with Socrates' and 'Socrates is self-identical'. I am fairly certain that it was Alvin Plantinga who used a pair of sentences like this as a counterexample to a principle like (35).

[26]Of course, two synonymous sentences (even the same sentence) can be used to express different propositions in different contexts, provided that the sentences contain indexical elements in their meaning. I am assuming that any such indexical elements are irrelevant in the present cases. For instance, if we suppose that (34b) and (34c) are tensed, then I'm assuming that they are uttered at the same time, so that being synonymous they must express the same proposition.

I should note that one very striking feature of the property theory is that substitution of one synonymous sentence for another in a cognitive context can fail to preserve truth, as in the case of going from (33b) to (33d). I don't see this as a problem, since this is clearly what happens when (33b) is understood *de re*. Note that this example involves a kind of "limiting case" of synonymy, since the component name 'Superman' is really being understood to have *no* meaning, only a referent (see the Appendix). When two sentences with context-independent meanings are synonymous, the property theory implies that they are intersubstitutable everywhere preserving truth.

[27]It may be of use to mention another kind of case discussed by Salmon that may make the same point even more forcefully. (Salmon, 1989, pp. 262–263; see also McKay, 1991, pp. 730–731.) Consider the following:

(ia) Lois believes that Superman can fly and Clark Kent cannot fly.
(ib) Lois believes that Superman can fly and Superman cannot fly.
(ic) Lois believes that Superman both can fly and cannot fly.

Once again, to avoid the absurdity of attributing to Lois the irrational belief ascribed by (ic), the defender of the relation theory must claim that the sentences imbedded in (ib) and (ic) express different propositions:

(iib) Superman can fly and Superman cannot fly.
(iic) Superman both can fly and cannot fly.

And again, it seems to be about as clear as anything ever gets that (iib) and (iic) say precisely the same thing. Surely, to say that an object has a given conjunctive property is just to say that a certain conjunction of propositions about that object is true.

[28]At one point (in a footnote), Soames indicates that sentences containing defined expressions are exceptions to his principles about propositional identity (1987a, p. 239, note 34). But then it is a mystery why he would claim that (34c), which contains the defined expression 'is self-outweighing', expresses a different proposition than (34b).

[29]A referee for *Nous* suggested that one of my criticisms of the Russellian theory also applies to the property theory. On the property theory, sentences of the form 'Mary believes that n is self-*R*ing' and 'Mary believes that n*R*n' may differ in truth-value, even though 'n is self-*R*ing' is definitionally equivalent to 'n*R*n'. It follows from this, according to the referee, that "the definition in question is not purely abbreviatory, but rather expands our conceptual repertoire." However, I fail to see why the

non-equivalence of the belief-sentences in question shows that introduction of predicates of the form 'self-*R*ing' by the definition in question would expand our conceptual repertoire in any way. As I've tried to make clear in the text, the property theory relies on the assumption that any sentence of the form 'n is self-*R*ing' has precisely the same meaning as a sentence of the form 'n*R*n' to explain how a belief sentence of the form 'Mary believes that n is self-*R*ing' could ascribe an irrational belief that is semantically analogous to a necessarily false sentence of the form 'n*R*n'. So it is clear that, on the property theory, our concept of a belief of the form 'n is self-*R*ing' would *not* be a new concept, since it would be precisely the same as a concept we already had, namely, the concept of a belief of the form 'n*R*n'. On the property theory, as we've seen, a *de re* sentence like 'Mary believes that n*R*n' would by contrast *not* ascribe a belief of the form 'n*R*n', and so it can be true even though 'Mary believes that n is self-*R*ing' is false. But this does not imply that the latter sentence involves any expansion of our conceptual repertoire.

[30]Donnellan (1970) described a similar example.

[31]A surprising number of philosophers (including a referee for *Nous*) have objected to my suggestion that 'Hesperus' and 'Phosphorus' have linguistic meanings corresponding to descriptive rules like (13) and (14), by emphasizing that they themselves constantly fail to remember which of the two names is associated with the evening and which is associated with the morning. Such a failure of memory is no doubt due to the fact that the objectors tend to be native speakers of English, while the names in question are names of Greek (native speakers of Greek of course do not have this problem). I thus invite English speaking sceptics about descriptive names to everywhere substitute the name 'the Evening Star' for 'Hesperus' and 'the Morning Star' for 'Phosphorus'.

[32]Pamela McKinsey gave me this example.

[33]On the theory of names that I've proposed elsewhere (McKinsey, 1984), this mistake is even more natural and easy to explain. On this theory, whenever a person uses a word as a name, the person is following a descriptive reference-fixing rule of the form illustrated by the rules (13) and (14) in the text. Typically, however, the rule is idiosyncratic to the speaker, and so it cannot provide a linguistic meaning that could be semantically conveyed by the name in the public language. For instance, the description involved in the rule might make essential reference to the speaker, such as 'the one to whom I've heard others refer with "Gödel'". Nevertheless, when a speaker follows such a rule in using a word as a name, it will be *as if* the speaker were speaking a language that contains the descriptive rule in question, and so the two kinds of situations will be indistinguishable as far as the speaker's linguistic intuitions are concerned. Thus, on my description theory of names, every speaker's use of a name will have the intuitive "feel" of a descriptive name, and so it is easy for any cognitive context containing a small-scope name to intuitively "feel" as if it had a *de dicto* reading. Of course, when the descriptive rule governing the name-use is idiosyncratic to the speaker, the use cannot semantically convey a way of thinking that could be ascribed to a believer, and so the assumption that the name has a descriptive meaning is mistaken. But on my theory of names, the tacit acceptance of such a mistaken assumption is almost unavoidable.

[34]An additional large class of real substitution failures can be accounted for by the property theory, if we make the plausible hypothesis that many such failures are due to implicit quotation of imbedded items. Suppose that Sally has heard of Cary Grant and knows that he was a famous movie actor, but she doesn't have the slightest idea that his original name was 'Archibald Leach'. Then (i) is true, but (ii) seems false:

(i) Sally believes that Cary Grant was a famous movie actor.
(ii) Sally believes that Archibald Leach was a famous movie actor.

In a recent paper, Steven Rieber (1997) suggests that substitution of coreferring names can fail when one or both of the names occurs in a context of implicit "complex quotation", in which the name is simultaneously both mentioned and used. On this idea, (ii) might be understood to mean

(iii) Sally believes that 'Archibald Leach' was a famous movie actor,

which Rieber (p. 272) suggests would in turn mean

(iv) Sally believes that 'Archibald Leach' refers to someone who was a famous movie actor.

In my view, many of the standard kinds of examples of substitution failure are really cases involving implicit quotation. Such cases are not of any theoretical interest, since in such cases the relevant names are not really coreferential, being used as names of themselves and not of their standard referents.

[35]Evans (1977) maintains that all unbound pronouns whose referents are determined by descriptions recoverable from their quantifier antecedents are rigid designators. In my 1986 article "Mental Anaphora" I showed that the evidence Evans gives in support of this generalization is inconclusive, and that in many of the cases he gives to support his view, the pronouns in question must in fact be going proxy for non-rigid definite descriptions (McKinsey, 1986, p. 161 and p. 174, note 4). Soames (1989) also criticizes Evans' view, on the same kinds of grounds that I had given. But Soames concludes from his discussion that "E-type pronouns are not, in fact, rigid designators" (Soames, 1989, p. 145). I, on the other hand, went on to provide a new argument, based on my example (47), that there are in fact rigid E-type pronouns, as Evans believed, even though many of the pronouns that Evans called 'E-type' are in fact non-rigid. (For the complete argument, see McKinsey, 1986, pp. 166–167.) My example (47) is one of a wide range of cases that I had adduced in unpublished work of 1974 to show that there are rigid anaphoric pronouns of the type that Evans later dubbed 'E-type'. I believe that the evidence provided by my examples shows that there are rigid pronouns of this kind, even though I agree with Soames that the evidence Evans gave does not show this.

[36]My explanation of the fact that a descriptive term's referent drops out as irrelevant when the term occurs in the scope of a cognitive operator, is based on the priority (mentioned above in section 10) that we give to (semantically) *de dicto* readings of cognitive predicates. The main function of such predicates is to say as much as possible about the fundamental semantic properties of our cognitive states, and so we give priority to interpretations on which such predicates express properties that can *individuate* a cognitive state. When a term has a descriptive meaning, this meaning is relevant to individuating a cognitive state, but its referent (if any) is not relevant to individuating the state (as I argue in McKinsey 1994). Hence in a cognitive ascription, a descriptive term's referent drops out as irrelevant to the classification of the cognitive state in question. (For fuller discussion of this point, see my 1994, pp. 219–221.)

[37]By an "affirmative" cognitive operator, I mean any operator C such that 'x Cs that S' logically implies 'x believes that S′.

[38]If S does contain other such terms, then the property ascribed can be specified by combining the method to be described for E-type pronouns with the method for interpreting *de re* predicates described in the Appendix.

[39]See note 33 above, and Rieber (1997).

[40]Other difficulties for Richard's theory have been raised by Schiffer (1990), Crimmins (1992), Sider (1995), and Soames (1995).

[41]Earlier and shorter versions of this paper were presented at the Universities of Oxford, Glasgow, St. Andrews, Bristol, Cambridge, Moscow, Uppsala, Helsinki, and Barcelona, to the Institute of Philosophy of the Russian Academy of Sciences, and to the Central Division of the American Philosophical Association (Chicago, May, 1998). I am grateful to the participants at those occasions for their helpful comments. I am especially grateful to Scott Soames, who gave insightful comments at the APA session, and whose constructive criticisms have in general led to many improvements in the paper. I am also grateful to two referees for *Noûs* for their salutary criticisms. For useful discussions of these matters, I am grateful to Traugott Schiebe, and to my colleagues Lawrence Lombard, Lawrence Powers, Susan Vineberg, and especially to David Shier and Paul Wagoner. I am also grateful to the members of my 1997 seminar in the philosophy of language, especially to David Baggett and Mark Huston. Some of the work on this paper was supported by a 1995 summer research award and by a sabbatical leave during 1998–99; I am grateful to Wayne State University for both of these

sources of support. Finally, I am indebted to the Subfaculty of Philosophy of Oxford University, to Corpus Christi College, to the staff at 10 Merton Street, and especially to Martin Davies, for providing me with a superb environment in which to work during 1998–99.

References

Ackermann, Diana: 1979, "Proper Names, Propositional Attitudes and Non-Descriptive Connotations," *Philosophical Studies* 35, 55–69.

Bach, Kent: 1997, "Do Belief Reports Report Beliefs?" *Pacific Philosophical Quarterly* 78, 215–241.

Berg, Jonathan: 1988, "The Pragmatics of Substitutivity," *Linguistics and Philosophy* 11, 355–370.

Böer, Steven E. and Lycan, William G.: 1980, "Who, Me?" *The Philosophical Review* 89, 427–466.

Crimmins, Mark: 1992, "Context in the Attitudes," *Linguistics and Philosophy* 15, 185–198.

Crimmins, Mark and Perry, John: 1989, "The Prince and the Phone Booth: Reporting Puzzling Beliefs," *Journal of Philosophy* 86, 685–711.

Donnellan, Keith: 1970, "Proper Names and Definite Descriptions," *Synthese* 21, 335–358.

Donnellan, Keith: 1974, "Speaking of Nothing," *Philosophical Review* 83, 3–31.

Evans, Gareth: 1977, "Pronouns, Quantifiers, and Relative Clauses (I)," *Canadian Journal of Philosophy* 7, 467–536.

Evans, Gareth: 1979, "Reference and Contingency," *The Monist* 62, 161–189.

Frege, Gottlob: 1892, "On Sense and Reference," in *Translations from the Writings of Gottlob Frege*, P. Geach and M. Black (eds.), (Oxford: The Oxford University Press, 1966).

Geach, P.T.: 1967, "Intentional Identity," *Journal of Philosophy* 64, 627–632.

Grice, H.P.: 1961, "The Causal Theory of Perception," in Paul Grice, *Studies in the Way of Words* (Cambridge, Mass.: Harvard University Press, 1989), 224–227.

Grice, H.P.: 1975, "Logic and Conversation," in Paul Grice, *Studies in the Way of Words* (Cambridge, Mass.: Harvard University Press, 1989), 22–40.

Kaplan, David: 1977, "Demonstratives," in J. Almog, J. Perry, and H. Wettstein (eds.), *Themes from Kaplan* (Oxford: Oxford University Press, 1989), 481–563.

Kripke, Saul: 1972, "Naming and Necessity," in D. Davidson and G. Harman (eds.), *Semantics of Natural Language* (Dordrecht: D. Reidel), 253–355.

Ludwig, Kirk: 1996, "Singular Thought and the Cartesian Theory of Mind," *Nous* 30, 434–460.

McDowell, John: 1977, "On the Sense and Reference of a Proper Name," *Mind* 86, 159–185.

McKay, Tom: 1981, "On Proper Names in Belief Ascriptions," *Philosophical Studies* 39, 287–303.

McKay, Tom: 1991, "Representing *De Re* Beliefs," *Linguistics and Philosophy* 14, 711–739.

McKinsey, Michael: 1978a, "Names and Intentionality," *Philosophical Review* 87, 171–200.

McKinsey, Michael: 1978b, "Kripke's Objections to Description Theories of Names," *Canadian Journal of Philosophy* 8, 485–497.

McKinsey, Michael: 1983, "Review of *Designation*, by Michael Devitt," *Canadian Philosophical Reviews* 3, 112–116.

McKinsey, Michael: 1984, "Causality and the Paradox of Names," *Midwest Studies in Philosophy* 9, 491–415.

McKinsey, Michael: 1986, "Mental Anaphora," *Synthese* 66, 159–175.

McKinsey, Michael: 1991, "The Internal Basis of Meaning," *Pacific Philosophical Quarterly* 72, 143–169.

McKinsey, Michael: 1994, "Individuating Beliefs," *Philosophical Perspectives* 8, Logic and Language, 303–330.

McKinsey, Michael: 1998, "The Grammar of Belief," in F. Orilia and W. J. Rapaport (eds.), *Thought, Language, and Ontology* (Kluwer), 3–24.

Plantinga, Alvin: 1978, "The Boethian Compromise," *American Philosophical Quarterly* 15, 129–138.

Richard, Mark: 1983, "Direct Reference and Ascriptions of Belief," in N. Salmon and S. Soames (eds.), *Propositions and Attitudes* (Oxford: Oxford University Press, 1988), 169–196.
Richard, Mark: 1990, *Propositional Attitudes* (Cambridge: Cambridge University Press).
Rieber, Steven: 1997, "A Semiquotational Solution to Substitution Puzzles," *Philosophical Studies* 86, 267–301.
Russell, Bertrand: 1904, "Letter to Frege," in Frege, Gottlob, *Philosophical and Mathematical Correspondence*, ed. G. Gabriel, H. Hermes, F. Kambartel, C. Thiel,, and A. Veraart, abr. B. McGuiness, tr. H. Kaal (Chicago: University of Chicago Press, 1980), p. 169.
Russell, Bertrand: 1912, *The Problems of Philosophy* (New York: Oxford University Press, 1959).
Salmon, Nathan: 1986, *Frege's Puzzle* (Cambridge, Mass.: The MIT Press).
Salmon, Nathan: 1986a, "Reflexivity," in N. Salmon and S. Soames (eds.), *Propositions and Attitudes* (Oxford: Oxford University Press, 1988), 240–274.
Salmon, Nathan: 1989, "Illogical Belief," *Philosophical Perspectives* 3, Philosophy of Mind and Action Theory, 243–285.
Salmon, Nathan: 1992, "Reflections on Reflexivity," *Linguistics and Philosophy* 15, 53–63.
Schiffer, Stephen: 1977, "Naming and Knowing," *Midwest Studies in Philosophy* 2, 28–41.
Schiffer, Stephen: 1987, "The Fido-Fido' Theory of Belief," *Philosophical Perspectives 1, Metaphysics*, 455–480.
Schiffer, Stephen: 1990, "The Relational Theory of Belief [A Reply to Mark Richard]," *Pacific Philosophical Quarterly* 71, 240–245.
Schiffer, Stephen: 1992, "Belief Ascription," the *Journal of Philosophy* 89, 499–521.
Sellars, Wilfrid: 1969, "Metaphysics and the Concept of a Person," in Karel Lambert (ed.), *The Logical Way of Doing Things* (New Haven: Yale University Press).
Sider, Theodore: 1995, "Three Problems for Richard's Theory of Belief Ascription," *Canadian Journal of Philosophy* 25, 487–513.
Soames, Scott: 1985, "Lost Innocence," *Linguistics and Philosophy* 8, 59–71.
Soames, Scott: 1987a, "Direct Reference, Propositional Attitudes, and Semantic Content," in N. Salmon and S. Soames (eds.), *Propositions and Attitudes* (Oxford: Oxford University Press, 1988), 197–239.
Soames, Scott: 1987b, "Substitutivity," in J. Thomson (ed.), *On Being and Saying: Essays for Richard Cartwright* (Cambridge, Mass.: The MIT Press), 99–132.
Soames, Scott: 1989, "Review of *Collected Papers* by Gareth Evans," *Journal of Philosophy* 86, 141–156.
Soames, Scott: 1990, "Pronouns and Propositional Attitudes," *Aristotelian Society Proceedings* NS 90, 191–212.
Soames, Scott: 1994, "Attitudes and Anaphora," *Philosophical Perspectives* 8, Logic and Language, 251–272.
Soames, Scott: 1995, "Beyond Singular Propositions?" *Canadian Journal of Philosophy* 25, 515–549.