On the Metaphysics of Belief

by

Cara Spencer

A.B., Philosophy,
Columbia College, 1992

Submitted to the Department of Linguistics and Philosophy in Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy

at the
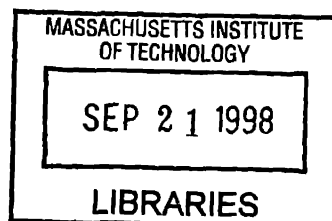
Massachusetts Institute of Technology

September 1998

Signature of Author:_____
Department of Linguistics and Philosophy
September 4, 1998

Certified by:_____
Robert C. Stalnaker
Professor of Philosophy and Chair of Linguistics and Philosophy
Thesis Supervisor

Accepted by: _____
Vann McGee
Professor of Philosophy
Chairman, Committee on Graduate Students

On the Metaphysics of Belief

by
Cara Spencer

ABSTRACT

There is a traditional picture of belief, according to which someone's having a belief is that person's standing in a certain relation to an abstract object, a proposition. My dissertation examines the metaphysical demands that two problems for this picture of belief make on these abstract objects. The first problem comes to us from Frege's "On Sense and Reference," and the second concerns a certain sort of one's beliefs about oneself, which I call "indexical belief."

Frege notes that someone can believe that Hesperus is Hesperus without believing that Hesperus is Phosphorus. It is a short step from Frege's observation to the claim that the sentence "A believes that Hesperus is Hesperus" could be true while "A believes that Hesperus is Phosphorus" is false. Quine has insisted that we cannot address this problem by taking either of these sentences to express a predication of the referent of both names, the planet Venus, as then both sentences would express the same predication of the same object, and this stands in evident tension with the fact that they can differ in truth value. In a recent response to Frege's problem, Mark Richard has taken exactly the route Quine has warned against. In my first chapter, I show that Richard's response is unsuccessful. I then consider the implications of the failure of his response on his proposed semantics for belief and the philosophical motivation he provides for it. I argue that his semantics for belief ascriptions stands in serious tension with its purported philosophical motivation.

In the second and third chapters, I turn to the problem about indexical belief. The task of the second chapter is to identify this problem. To this end, I consider three of John Perry's arguments that the traditional picture of belief cannot accommodate indexical belief. I show that even if these arguments are sound, they give us no reason to think that the problem about indexical belief is in any way unique. I then suggest that there is a special problem about indexical belief, despite the failure of these three arguments to isolate it. Special difficulties attend an account of what it is to retain a tensed belief over time, and I suggest that the special problem about indexical belief is, in brief, a generalization of this problem about retention of tensed belief.

In the third chapter, I raise the problem about indexical belief in a new way, as arising from a tension between several intuitively plausible claims about the relationship between beliefs and desires, on the one hand, and actions, on the other. This presentation of the problem brings out how the problem is special. I survey several solutions to this problem, including one due to Perry. I then argue that Perry's solution cannot characterize the specific kind of similarity between the beliefs of two different people who share an indexical belief. It allows too much to count as shared indexical belief. Thus Perry's solution inadequately addresses what I have suggested is the unique problem about indexical belief. Finally, I suggest that Frege's response to the problem I raise points the way towards an emendation of Perry's account that will enable it to suitably characterize the special way in which different people can share an indexical belief.

Thesis Supervisor: Robert C. Stalnaker
Title: Professor of Philosophy and Chair, Linguistics and Philosophy

Acknowledgments

Table of Contents

Introduction

We find in the later works of Gottlob Frege the view that someone's having a belief is that person's standing in some relation to a certain kind of abstract object, which Frege called a "thought." Like any other affirmation of the existence of abstract objects, Frege's view encounters metaphysical difficulties. How, one wonders, is one thought, or proposition, distinguished from the next? What is the philosophical work they do which justifies the extravagance of admitting them into our ontology? Since they do not enter into causal relations with physical objects such as ourselves, how do we come to stand in any interesting relations to them at all?

But the Fregean view also gives rise to special metaphysical questions. Many have embraced the suggestion, implicit in this view, that we distinguish between beliefs by distinguishing between the propositions that are their objects, i.e. the propositions to which one would stand in the belief relation if one had these beliefs. So if, for instance, believing that Hesperus is Hesperus is not the same as believing that Hesperus is Phosphorus, then the objects of these beliefs must be distinct. And this is so even though it is metaphysically impossible for one to be true and the other false. It seems that the difference between these two propositions must be a epistemic or psychological difference, a difference we might think should be situated in the mind rather than in the world. But now there is a problem, for how can propositions even have the kind of epistemic or psychological properties required to make these sorts of distinctions if they are, as Frege claims, mind-independent entities? Other metaphysical questions arise from the proposition's role as a bearer of truth value. Propositions can be asserted or denied, and they stand in various in logical relations to one another, such as that of implication or incompatibility.

How does this logical role of the proposition constrain our account of what they are, and how they are related to each other? This dissertation examines some of the ways in which these two features of propositions, their epistemic or psychological role on the one hand and their logical or truth-bearing role on the other, constrain our metaphysics.

That propositions should have a logical role at all is somewhat puzzling. Logic, after all, is the study of certain syntactic properties of formal language sentences. But propositions are emphatically not sentences, and it is not clear how syntax as a formal and computational features of sentences could or should inform the more philosophically loaded notion of "logical form," in Russell's sense of that term. Yet at the same time it is clear that if one proposition is to follow from, or to be incompatible with, another, there must be something about these propositions, whether it is structural or more semantical in nature, in virtue of which these relations hold. My first chapter concerns the relationship between the syntax of a sentence and the proposition it can express. There I argue, as against Mark Richard, that logical laws, such as what I call the Principle of Identity, set fundamental constraints on what kinds of propositions certain sentences can express.

I consider a second metaphysical problem about propositions, which concerns what I called earlier their "epistemic" features. What requires propositions to have these "epistemic" features is most sharply evident when we consider their roles as the objects of certain kinds of beliefs about ourselves and our immediate environments. These beliefs are typically expressed with sentences containing indexical expressions such as "I", "here," or "now" or demonstrative expressions such as "this" or "that." Whether we think of a place as "here" or in some other way matters to what we do with respect to this place. For similar reasons, whether one thinks of a certain person as *oneself*, or in some other way, say as the woman in the mirror, matters to what one does. The prevailing attitude is that these kinds of beliefs, which I call "indexical beliefs," present a unique problem for the Fregean picture of belief. Yet the prevailing attitude is in this case puzzling, since

we find that the intuitively attractive statements of the purportedly unique problem about indexical

belief fails to completely explain or to justify the prevailing attitude. I make a case for this claim in

chapter two. Yet it still seems right to say that we do face a special difficulty here. Thus in chapter

three, I present the problem about indexical belief in a way that isolates its particularly problematic

features, which relate to their roles in discourse and the explanation of behavior.

1

Belief and the Principle of Identity

1.   The New Theorists of Reference have argued that proper names, demonstratives, and indexical pronouns are devices of direct reference. That is, the only semantic contribution these expressions make to the propositions expressed by sentences containing them are the objects to which they refer. Some of these same New Theorists also offer what they call a "Russellian" semantics for attitude ascriptions. On the Russellian view, verbs of propositional attitude are two-place predicates which take as arguments pairs of the person to whom the attitude is being ascribed, and the proposition expressed by the sentence following the attitude verb. I shall follow convention in calling a theory that accepts the conjunction of these two claims a "Russellian theory."

The intuitive appeal of the Russellian theory is considerable. Saul Kripke and David Kaplan have both argued in support of parts of the first claim; considered together, their arguments provide powerful support for the view that proper names, demonstratives, and indexicals are directly referential.[1] Further, it is quite generally a desideratum on a semantic theory that the semantic value it predicts for a sentence should respect the surface form of that sentence. On a Russellian semantics for verbs of propositional attitude, these verbs take pairs of persons and propositions as arguments, as the surface form of attitude ascriptions suggests. We need a strong reason to reject the Russellian theory in favor of one that does violence to the surface form. Finally, the prevailing

---

[1]See Kripke [1980] and Kaplan [1989]. The second thesis, about the semantic contribution of propositional attitude verbs, is less widely accepted, and many of those who do accept it offer slightly different versions of it. The *locus classicus* of the view is Salmon [1986]; it is also stated in Soames [1987] and Richard [1983].

alternatives to the Russellian theory posit semantically relevant mentalistic entities such as the sense

of an expression, or an expression's mode of presentation of its referent, entities whose identity

conditions are often less than clear. If we can give a semantics for attitude ascriptions without

having to explain what these things might be, then so much the better for semantics.

It is widely appreciated that the central difficulty facing the Russellian theorist is to explain

how her theory can accommodate the puzzle cases that Frege called to our attention long ago. If

"Hesperus is Hesperus" expresses the same proposition as "Hesperus is Phosphorus," and if

propositions are the objects of belief, then it is not even possible for Ed accept the first and reject the

second. But evidently he can. It is but a short step from here to the observation that (1) can be true

while (2) is false,

(1) Ed believes that Hesperus is Hesperus.

(2) Ed believes that Hesperus is Phosphorus.

and this is a bit of an embarrassment to the Russellian theory I present above, according to which

these two belief ascriptions express the very same proposition.

Quine would insist that we attend to the following facts about (1) and (2):

(a) They do not have the same truth value,

(b) They differ only in that one contains an occurrence of the name "Hesperus" where the other contains an occurrence of the name "Phosphorus," and,

(c) "Hesperus" and "Phosphorus" co-refer.

According to Quine, a consideration of these facts should lead us to think that in the context

of these two sentences, the terms "Phosphorus" and "Hesperus" provide something of semantic

significance besides their referent, to which the embedding context is sensitive, which accounts for

the difference in truth value. If so, we cannot represent what these two sentences have in common

as open sentences with one free variable as the Russellian theorist would have it, since the variable

in such a sentence simply supplies the assigned object, and nothing else that the singular term might

supply. One might put Quine's point this way: The very *meaning* of the variables and quantifiers

requires the truth of the following Principle of Identity ([PI]):

> [PI] Where α is some sentence which may contain some free occurrences of x, and
> $\alpha[y/x]$ is α with y in some or all of the places where x occurs free in α, all
> instances of
> $$(\forall x)(\forall y)(x = y \rightarrow \alpha \rightarrow \alpha[y/x])$$
> are true.

Richard Cartwright calls instances of the schema in [PI] "indiscernibility principles," and

says that all such principles are true, and that furthermore this requirement is a "fundamental

constraint" on the interpretation of α.[2] Quine concurs, identifying any interpretation of α that

violates [PI] as a "distortion."[3] Many of us would be inclined to agree with Cartwright and Quine,

and to think that no semantical, logical, or philosophical consideration could so much as count as a

reason to deny [PI]. Even those of us who reject Quine's and Cartwright's stance that

quantification into opacity is always forbidden would probably draw the line at [PI].

The foundational character of [PI] makes its conflict with the deliverances of the Russellian

theory deeply troubling. And there is a conflict. If the Russellian theorist accepts the ordinary

intuition that (1) can be true while (2) is false, and she thinks that Existential Generalization holds

unrestrictedly, she must also affirm (3) and (4).

(3) $(\exists x)$ Ed believes that x is x.

(4) $(\exists x)(\exists y)$& Ed believes that x is y.

Since the sequences that assign the planet Venus to x and y will satisfy the open sentence $x = y$,

these results appear to require us to affirm (5), which implies that an instance of the schema in [PI]

is false.

(5) $(\exists x)(\exists y)(x = y$ & Ed believes that x is x & & Ed believes that x is y)

2Cartwright [1979] p. 213
3Quine [1960] p. 167

Typically the Russellian theorist responds to this conflict by denying that speaker intuition about (1) and (2) requires us to say that they really can have different truth values. But another option, taken by Mark Richard, is to deny [PI]. In support of this claim, he constructs an artificial language whose predicates are sensitive to ancillary features of the variable, so that the form of the variable in part determines the satisfaction conditions of some open sentences containing it. He then claims that this language has a Tarskian semantics, which ensures that its quantifiers are objectual.[4] The dual sensitivity of the predicate permits the formation of what Richard claims to be false instances of the schema in [PI]. In the first part of this chapter, I argue that Richard has given us no reason to think that [PI] is false, and that he is furthermore wrong in thinking it is false.

To assess Richard's claim that these sentences are instances of that schema, we need to know what counts as an instance of that schema. Schemata are language-specific, so the sentence must be a sentence of a certain language, in this case the language of first-order predicate logic. My strategy will be to show that the language in which Richard constructs putative counterexamples to [PI] is not such a language, hence it contains no instances of the schema in [PI], true or false. I then propose a specific constraint on the open sentences of a language that any language must meet if it is to contain instances of the schema in [PI], and argue that Richard's artificial language does not meet this constraint.

In the second part of this chapter, I turn to Richard's stated motivation for denying [PI]. There, I identify two sources of tension between Richard's semantics for attitude ascriptions and a philosophical motivation he offers for it, both of which are easily seen to follow on the truth of [PI]. First, Richard rightly notes that we have defeasible reason to prefer a semantic theory that accurately reflects the surface form of natural language sentences over one that does not. He

---

[4]Quine introduced the notion of an objectual quantifier to distinguish ordinary quantification from substitutional quantification. An objectual quantifier is one that ranges over objects, and a sentence of the form $\exists x Fx$ is true just in case there is some object within the range of the quantifier which is in the interpretation of F. Substitutional quantifiers have a substitution class of linguistic expressions, so a sentence of the form $\exists \alpha F\alpha$ is true just in case there is some expression e in the substitution class of the quantifier such that F followed by e is a true sentence.

incorrectly infers that this gives us defeasible reason to prefer his theory over competing theories that do not pretend to violate [PI]. Second, according to Richard the only way to respect the claim that our beliefs are about objects is to ensure that the objects our thoughts are about really are constituents of the propositions we believe. This requires him to say that when we quantify in to an attitude context, the initial quantifier must range over ordinary objects, the everyday tables and chairs that our thoughts are about. I shall argue that if Richard is right that we only accommodate the claim that our thoughts are about objects by putting objects directly into propositions, then his account of attitude ascription is seriously flawed, since his technical apparatus cannot deliver a theory that would be acceptable on its own terms.

2.   According to Richard, an objectual quantifier is one "whose semantics is properly given in terms of a Tarski-style definition of satisfaction"[5] . We do not simply give a Tarskian semantics for quantifiers in the absence of a Tarskian definition of satisfaction for the language of which these quantifiers are a part. So when Richard says that objectual quantifiers, so defined, can be part of a language in which [PI] does not hold, he is saying that we can give a Tarskian definition of satisfaction for a language that violates [PI].

Richard sketches a fragment of a formal language which is claimed to both have a Tarskian semantics and to contain a counterexample to [PI]. This fragment differs from more familiar first order languages in only two respects. First, it contains an idiosyncratic predicate F, defined in the first semantic rule. Second, it contains only those expressions necessary to generate the counterexample sentence. In the interests of clarity I present the entire fragment here, but aside from the rule (3a) the fragment is completely familiar. The semantics for the existential quantifier, the conditional, and the identity sign, as well as all of the syntactic rules, are unchanged from these more familiar languages. The semantic rule for the idiosyncratic predicate F allows a true sentence

---

[5]Richard [1990] p. 202

of the form $\exists x \exists y(x = y \,\&\, Fx \,\&\, \&Fy)$, which implies that an instance of the schema in [PI],

$(\forall x_1)(\forall y_2)(x_1 = y_2 \rightarrow Fx_1 \rightarrow Fy_2)$ is false. The fragment is defined as follows:

Fragment of L :

1. Primitive expressions:

        a. First-order variables with numerical indices - $\{x_1, x_2, x_3,...\}$

        b. Constant one-place predicate expression - F

        c. Constant two-place predicate expression - =

        d. Quantifier - $\exists$

        e. Logical Constants - $\rightarrow$, &

2. Syntax:

        a. For all x_i, Fx_i is a formula.

        b. For all x_i and x_j, $x_i = x_j$ is a formula.

        c. If α is a formula and x_i is a variable, then $\exists x_i \alpha$ is a formula.

        d. If α and β are formulae then $\alpha \rightarrow \beta$ is a formula.

        e. If α is a formula then $\&\alpha$ is a formula

3. Semantics:

        In any model **M** <**A**,**I**> for L, **A** is the set of objects in the domain and **I** is a set of ordered

pairs of constant terms in L and subsets of **A**.

        a. Fx_i is satisfied by a sequence s iff

                1)The object in the ith place of s is in **I**(F),

                2) i is an odd natural number.

        b. $x_i = x_j$ is satisfied by a sequence s iff the sequence assigns the same object to x_i and x_j

        c. $\exists x_i \alpha$ is satisfied by a sequence s iff there is another sequence s' like s except maybe in

the assignment to x_i which satisfies α.

d. $\alpha \rightarrow \beta$ is satisfied by a sequence s iff it is not the case that the sequence satisfies α but does not satisfy β.

e. $\&\alpha$ is satisfied by a sequence s iff it is not the case that the sequence satisfies α.

I shall consider in the next section whether this is truly a Tarskian definition of satisfaction for L. According to the clause in the semantics for L that gives the rule for formulae of the form Fx_i, a sequence satisfies the open sentence Fx_i just in case the object assigned to the variable is in the extension of F and the subscript i is an odd natural number. So, for instance, a sequence that assigns to both x_1 and x_2 an object in the extension of F will satisfy Fx_1 but not Fx_2. Since there is no sequence of objects that satisfies Fx_2, the closed sentence $\exists x_2 Fx_2$ is false. However the closed $\exists x_1 Fx_1$ is true, as long as the model contains some object in the extension of F. Since the identity predicate has the standard extensional semantics, the sentence (6) is false in the model M.

(6) $\forall x_1 \forall x_2 (x_1 = x_2 \rightarrow Fx_1 \rightarrow Fx_2)$

M = <**A**,**I**>

A = {Hillary Clinton }

I = { <F, {Hillary Clinton} >, < =, {<Hillary Clinton, Hillary Clinton>}>}

While Richard would grant that natural languages do not violate [PI] in exactly this way, he thinks the possibility of languages like L shows that [PI] is not in any way fundamental to objectual quantification.

3.   Richard claims that in denying [PI], he does not intend to challenge the trivial metaphysical truth that identical objects share all of their properties.[6] At the same time, Richard is not trying to demonstrate what is obviously true, that there are false sentences orthographically identical to instances of the schema in [PI]. But on the standard account of what instances of this schema say,

---

[6]Of course this way of formulating the principle is not the most felicitous, since it appears to concern several objects which have all of their properties in common. But what else can we say? An object has all of its properties and no others? If an object has a property, then it has that property?

there is no middle ground for Richard to occupy. To take issue with Richard's claim without

begging the question against him, we cannot simply help ourselves to this standard account.

Rather, we must attempt a neutral specification of what makes a sentence an instance of this

schema. Since schemata are language-specific, the sentence must be a sentence of a certain

language, in this case the standard first-order predicate calculus. So if Richard thinks (6) is an

instance of the schema, then he must also think that this calculus is flexible enough to

accommodate L, in spite of its nonstandard semantic rule for F and the other semantic

idiosyncrasies to which F gives rise. For instance, F's sensitivity to the form of the variable

interferes with the usual free interchange of variables in sentences, which is exactly what gives rise

to the purported counterexample sentence. When L is enriched to include constant singular terms

as well as variables, the rules that permit substitution of constant singular terms for variables,

Existential Generalization (EG) and Universal Instantiation (UI), do not hold in their usual

generality.[7] So it is at least not obvious that L is a standard first order predicate logic, as it needs

to be to contain an instance of the schema.

Richard suggests that if the quantifiers in L have a Tarskian semantics, then these quantifiers

are objectual, and in this case L is a counterexample to [PI]. So when, according to Richard, is a

semantics for quantifiers Tarskian? He proposes that it must meet the following constraint, [C][8]:

---

[7]The semantic rule for F only tells us what to do if F is followed by a variable, and not by a singular term. An extension of the rule is required, and this does not flow naturally from what is given. F's peculiar feature is that it is sensitive to the form of the variable. Intuitively an extension of that rule should respect this dual sensitivity, but at any event either it will or it will not, and either way we lose a rule of inference. Typically constant singular terms denote only one object in the domain. If we attempt to respect the dual sensitivity of F by introducing multiple names for each object (or some objects) in A, then we must reject universal instantiation, since there are models for an enriched L in which $\forall x_1 Fx_1$ is true but Fa is false. We could also introduce singular terms without trying to accommodate F's dual sensitivity. But on this route, existential generalization fails, since Fa would be true as long as the value of a is in the set F, but $\exists x_2 Fx_2$ is false. I thank Gabriel Uzquiano for bringing this point to my attention. I should also note that I do not point to this idiosyncrasy as an objection to Richard. Clearly to object to L is this way is to beg the question against Richard. My point is simply to indicate that the language is idiosyncratic in the inferences it validates, and to suggest that this puts the onus on Richard to explain why L is a natural extension of more familiar first order languages.
[8]Richard [1990] p. 203

[C] A quantifier Q is objectual just in case all that matters to the truth value of QvS is what sequences satisfy S.

Richard uses condition [C] to argue as follows: The semantics for the quantifiers in L adheres to [C], and this guarantees that the semantics for those quantifiers is Tarskian, hence that the quantifiers are objectual. But so long as its quantifiers are objectual, L can contain instances of the schema in [PI].

I do not think we should accept Richard's argument. As I mentioned above, we do not simply give a Tarskian semantics for the quantifiers without giving a Tarskian semantics for the language of which these quantifiers are a part. Clearly the quantifiers can only have a Tarskian definition of satisfaction if the entire language has one. Thus, if the quantifiers are to have a Tarskian semantics, then the definition of satisfaction for the entire language must meet certain global constraints. And this is where Richard's criterion [C] fails. Even though [C] is exactly what the clause for quantifiers in Tarski's definition of satisfaction requires, [C] alone places no constraints on the satisfaction conditions for open sentences. It permits every function from open sentences of a language to sets of sequences of objects in the domain as a base clause for the definition of satisfaction for the language. For instance, [C] permits the stipulation that a sequence can satisfy Fx_1 and $x_1=x_2$ but not Fx_2, which trivially allows a false sentence orthographically identical to (6). I assume that Richard would not identify any quantifiers prefixed to these open sentences, with these satisfaction conditions, as objectual. If he is willing to identify them thus, he sacrifices what was supposed to be so important about objectual quantifiers, their guarantee that the sentence is *about* objects in the range of the quantifier. Clearly, there is no interesting sense of "about" in which these three sentences are about the object or objects that the sequence assigns to x_1 and x_2. In L, even though the object assigned to the variable is not always the only thing that matters in the determination of the truth-condition of the sentence, it is at least part of what matters. It is a necessary but not sufficient condition on the truth of a sentence of the form Fx_i that the object

assigned to x_i is in $I(F)$. But [C] does not require L to have this feature. Evidently a language

conforming to it can contain false sentences that are orthographically identical to instances of the

schema in [PI]. If conformity to [C] is all that is required of a language to ensure the objectuality of

its quantifiers, then the specifics of Richard's counterexample language do no work in his argument

whatsoever. So even if Richard has shown that a language in which the object assigned to the

variable is part but not all of what matters to the satisfaction of some open sentences containing that

variable, this is no thanks to condition [C].[9]

[C] constrains the semantics for a language, in that it requires that the definition of

satisfaction for the language must be recursive, at least for the sentences containing quantifiers. If

we simply consider this recursive aspect of the truth definition in "The Concept of Truth in

Formalized Languages" (henceforth "CTFL"), we see that Richard's proposed counterexample to

[PI] adheres to it. But this simply shows that it would be a perversion of the clear intent of Tarski's

paper, in particular a failure to attend to the pretheoretical notion of satisfaction that underwrites the

formal notion of satisfaction, to simply consider the recursive aspect of the definition. If (6) is to be

an instance of the schema in [PI], it is not enough that the quantifiers in L have the traditional

semantic rule. We must also understand the sentences α and $\alpha[y/x]$ as open sentences with one

free variable. If Richard can use F to generate a counterexample to [PI], he must show that

sentences composed of F followed by a variable are in fact open sentences with one free variable.

And this the recursive definition does not guarantee.

What constraints does Tarski set on the base clause for the definition of satisfaction, which

correlates the atomic open sentences with sets of sequences of objects in the domain? After he

presents the definition of satisfaction in the calculus of classes, the formal language for which he

defines truth, Tarski makes the following remark about satisfaction:

---

[9]Richard offers [C] in Richard [1987], and again when he presents the argument in Richard [1990]. However in conversation he has suggested a condition on predicates that could supplant [C]. I consider this proposal (labeled [A'] in this essay) in section four, below.

On the basis of Def. 22 [the definition of satisfaction in the class calculus] and the intuitive considerations that preceded it, it is easy to realize that *whether or not a given sequence satisfies a given sentential function depends only on those terms of the sequence which correspond (in their indices) with the free variables of the function*" (my emphasis)[10]

How should we understand Tarski's claim that whether a sequence satisfies an open

sentence depends only on the terms (i.e. the objects) in that sequence that are assigned to the free

variables in the open sentence? At minimum, we can conclude that Tarski adopts constraint [T].

[T] If sequences $s_1...s_n$ differ only in their terms which are not assigned to the free variables in an open sentence r, then either all of $s_1...s_n$ satisfy r or none satisfy r.

This minimal interpretation is perhaps buttressed by the conclusion Tarski immediately draws from

the passage quoted above, that the evaluation of closed sentences does not require us to look at any

of the terms in the sequence. If [T] is all Tarski intends by this remark, then Richard's language L

does appear to meet it. [T] is not equivalent to [C], since it constrains the base clause of the

definition of satisfaction for a language where [C] does not. So Richard may well suggest that by

offering [C] as a criterion of objectuality for quantifiers, he did not intend that [C] alone would

suffice to ensure the objectuality of the quantifiers, rather he intended that [C] *in addition to* other

constraints on the semantics for the connectives and atomic open sentences would ensure

objectuality of the quantifiers.

But in addition to it's being unclear what these other constraints would be, it is also unclear

that the minimal [T] is all Tarski intends by this remark. Tarski might also intend something a bit

stronger, i.e. not merely to deny that whether a sequence satisfies a sentential function can depend

on those terms which do not correspond to the free variables in the function, but also to assert that

whether a sequence satisfies a sentential function depends only on the objects themselves, and what

these objects are like. Given the unclarity of the italicized condition in isolation, we would do well

to examine the considerations that, according to Tarski, lend support to it. He suggests that we can

---

[10]Tarski [1933] p. 194

infer the condition from the definition of satisfaction as well as the "intuitive considerations that

preceded it". What then are these "intuitive considerations" Tarski refers to?

If there are any intuitive considerations about the notion of satisfaction, this must be because

the notion of satisfaction is a familiar notion, not one stipulatively introduced to serve a particular

technical purpose in the truth definition. Tarski indicates as much at the outset of his discussion of

satisfaction, referring specifically to the "usual meaning of this notion in its customary linguistic

usage."[11] This usual meaning is illustrated in talk about objects satisfying certain conditions. He

considers two examples of such talk. The first concerns a sentential function with one free variable.

He says that "we can significantly say of every single *object* that it does or does not satisfy the

given function." (my emphasis)[12] He offers the following schema as an explanation of this claim:

[Schema] For all a, a satisfies the sentential function x if and only if p.

He indicates that we substitute for "p" "the given sentential function (after first replacing the free

variable occurring in it by "a") and for "x" some individual name of this function," and offers

[Instance] as an instance of the schema in [Schema]:

[Instance] For all a, a satisfies the sentential function "x is white" if and only if a is white.

He also considers a special kind of sentential function, an algebraic equation, whose satisfiers are

particular numbers, the roots of the equation.[13] For instance, the number 1 is the only root of the

equation $x + 2 = 3$. Although "$x + 2 = 3$" and "$y + 2 = 3$" are not the same string of symbols, it is

clear that they express the same equation, so any number that satisfies the one will also satisfy the

other. To doubt this is to make a mistake of the most elementary kind. It is not the role of the "x"

and the "y", at least not in algebraic equations, to make substantive restrictions on what numbers

can satisfy the equations they are used to form. These two examples indicate that when Tarski uses

the expression "satisfaction" he has in mind a relation that holds between objects and the conditions

---

[11]ibid. p. 189
[12]ibid. p. 189-90
[13]ibid.

expressed by open sentences. Whether an object meets the condition depends solely on the object

itself. The parallel he draws between open sentences and algebraic equations suggests something a

bit stronger, and more to the point against Richard, that the open sentence, much like the algebraic

equation, is indifferent to the orthographic identity of the variables used to form it. So even though

Tarski speaks of open sentences with n free variables being satisfied by infinite sequences of

objects, rather than by an ordered n-tuple of objects when he gives the definition of satisfaction for

the class calculus, it is only "for the sake of a uniform mode of expression" that he does so. This

convention in no way undermines Tarski's point that, fundamentally, satisfaction is a relation

between objects and conditions that these objects might satisfy. It is clear that Tarski thinks that

satisfaction conditions for atomic open sentences do not simply arise from an arbitrary correlation of

open sentences with sequences of objects, as Richard's account suggests. Such an account could

not be an account of satisfaction, if satisfaction is to be a generalization of the familiar notion Tarski

appeals to.

Now if Tarski simply intended the minimal condition [T] by the italicized portion of the

remark quoted above, then surely these intuitive considerations would be irrelevant. But Tarski

clearly states that he draws the conclusion in the quotation italicized above in part from these

intuitive considerations. Therefore I think we ought not to take Tarski to intend simply [T], which

is all that would be required to generate a recursive definition of satisfaction. Instead we should

take him to intend something stronger, that satisfaction is fundamentally a relation between objects

and conditions that open sentences of the language express.[14]

---

[14]There is good reason to think that Tarski, at least the Tarski of CTFL, has been dragged into the dispute between Quine et. al. and Richard as an unwilling participant. John Etchemendy argues that Tarski had no intention of characterizing the truth definition to characterize the semantics of natural languages. According to Etchemendy, Tarski uses recursion and satisfaction only because the infinity of sentences in the object language and the finite length of sentences in his metalanguage require him to. For Tarski's purposes, a definition of satisfaction for a language which simply consists of a list of instances of the T-schema would work just as well as the one he gives us. If Etchemendy is correct, then at issue in the dispute is an adaptation of a Tarskian definition of satisfaction to serve a purpose quite different from the one he intended, and his remarks about the pretheoretical notion of satisfaction are no longer relevant. See Etchemendy [1988] for these arguments.

4.   One may protest here that it is not clear what exactly this stronger claim, that satisfaction is

fundamentally a relation between objects and conditions, comes to. What constraint does it place on

the semantics for a language? We can say immediately that whatever constraint this is, it will be a

constraint on the base clause of the definition. And indeed we find that it is the base clause of the

purported definition of satisfaction for L that contains an anomaly in the semantic rule (3a), which

tells how to determine the satisfaction conditions for sentences of the form Fx_i. The syntax and

semantics for all other expressions in L are unchanged from more familiar languages. Clearly

instances of the schema in [PI] can be formed in standard first order languages, so if they cannot be

formed in L, the anomalous rule (3a) must be responsible.

If Richard is to show that instances of the schema in [PI] can be formed in L, then he must

show how this rule is an acceptable variant of the standard rule for sentences of this form. Since

(3a) concerns both F and variables, Richard can argue either that (3a) is a natural extension of

variables or of predicates in first order languages. In an explicit response to this particular problem,

Richard takes the first tack, arguing that the second role of the variables in L is essentially an

extension of their cross-referencing role. I show that his parallel is insufficiently convincing.

However in related work we find an argument for broadening the Quine-Cartwright construal of

predication, which can be extended to argue that F is a one-place predicate which can enter into

instances of the schema in [PI].[15]   I show that this argument fails as well.

Tarski's remarks about satisfaction, quoted above, suggest a conception of variables as

place holders for objects in the domain of quantification. On this conception, the variable serves

simply to present the object assigned to it; it does not present this object in any particular way. On

Tarski's view, it would be a misunderstanding to suppose that an object could satisfy Fx but not

Fy, since the role of the x and the y in these open sentences is simply to present the assigned object

---

[15]Richard [1990] p. 113-6

itself, not in any particular way, not via any mediating mode of presentation. I think it is this feature of the variable that underlies the familiar remark that variables can be rigid designators.[16] If we attend simply to this feature of the variable, it is hard to see how the subscripted x's in L could qualify as genuine variables, as they sometimes supply something besides the variable, i.e. their numerical subscripts.

Although variables in formal languages serve to supply the assigned object, this is not all they do. They can also indicate cross-reference. Even though the same objects satisfy Fx and Fy, these open sentences will be satisfied by different sequences. Sequences assigning an object that is F to x will satisfy Fx, and sequences assigning such an object to y will satisfy Fy. One reason we need an account of satisfaction by an infinite sequence instead of by an object is that atomic open sentences can embed into more complex sentences where the variables may recur free in different places.[17] Variables are used to indicate cross-reference, so whenever a variable appears in an open sentence it must supply the same object. Fx and Fy set the same conditions on an object, but when conjoined with Rxy, they set different conditions on the object assigned to x and the object assigned to y.

Richard hopes that the use of variables as cross referencing devices will help to motivate his view that the variables in L, appearances aside, are consonant with the picture of the variable implicit in Tarski's truth definition. He appears to argue as follows: Variables have many functions in a formal language with a Tarskian truth definition. One of these functions is cross reference. When a variable is used in several places to supply the same object, the variable helps to characterize the condition that the open sentence determines. For instance, the open sentence "x loves y" determines a different condition from the open sentence "x loves x." Since the only difference between these two open sentences is that y occurs in the first but not the second, the

---

[16]See Kripke [1980] p. 49 fn. 16 for this view.

[17]The other reason is to allow all sentences of the object language to be satisfied by the same kind of thing, infinite sequences of objects.

variables themselves do something to characterize the condition that the open sentence determines. So one cannot object to L on the ground that its variables play a dual role, partly supplying an object and also entering into the specification of truth conditions independently of the assignment. Because in playing this second role, they simply help in the characterization of the condition that the open sentence determines, and variables in standard first order languages already do this. The variables in L perform this function in a very different way from the more familiar variables, but essentially the functions are the same.

What Richard calls the "variables" in L, much like the variables in standard first order languages, help to characterize the conditions determined by open sentences containing them. What is not clear is how this point supports the claim that the "variables" in L are not importantly different from regular variables. Richard does not mean to say that anything that helps to characterize the condition determined by an open sentence could be a variable, since every other kind of primitive expression in a formal language also helps to characterize these conditions. Is Richard saying that, given the *other* evident similarities between "variables" in L and variables in standard first-order languages, we should take the fact that both sorts of variables help to characterize the condition that an open sentence determines to suggest that the rule (3a) really does offer a natural extension of the semantics for variables? This construal of the argument just makes it seem question-begging. Whether the expressions of the form x_i are relevantly similar to variables in other first order languages is precisely what is at issue. Richard points to a feature that these expressions have in common with the more familiar variables, a feature they also share with *every other expression* in a familiar first order language. But the mere fact that these expressions have this feature does not suggest that they are variables.

Richard also notes that variables can be used to expand the stock of predicates in a language. For instance, the one-place predicate "loves himself" can be constructed out of the two-place predicate "loves" in quantified sentences by placing the same variable in both of the latter's

argument places.   Do the "variables" in L expand the stock of predicates in L? In a standard first

order language, $\exists x_1 F x_1$ and $\exists x_2 F x_2$ say the same thing, but in L, they do not. The first sentence

says that something is an F and 1 is an odd natural number, and the second says something quite

different, that something is an F and 2 is an odd natural number. Have the "variables" in L allowed

us to use the same predicate letter F to express two different predicates?[18]

The variables can only expand the stock of predicates if what is in the original stock really

are predicates, and this is part of what is at issue. To use the argument above, Richard must already

have established that F truly is a one-place predicate. All we know at this point in the inquiry is that

if (6) is not an instance of the schema in [PI], then the semantic rule (3a), that gives the satisfaction

conditions for open sentences of the form $F x_i$, must be responsible. To convince us that he is really

disagreeing with Quine in offering (6) as a counterexample to [PI], he must show that what look

like predicates and variables in L really are predicates and variables. What Richard's own proposal

suggests is that these may not be two independent tasks.

5.   I noted above that to show that (6) is a counterexample to [PI], Richard must show that the

semantic rule (3a) is a natural extension either of the typical semantics for one-place predicates or

for variables. I argued that Richard's assimilation of the dual role of the variables in L to their

cross-referencing role fails. But Richard argues elsewhere that Quine's conception of predication is

too narrow. Richard thinks that Quine's arguments against quantification into opacity leave us with

a unnecessarily restrictive conception of predication, and it is this restrictive conception that

motivates [PI]. If these arguments are successful, they will give us independent reason for thinking

that F can be used to generate a counterexample to [PI].

Before I turn to Richard's argument, I present Cartwright's reconstruction of Quine's

arguments against quantification into opacity, and bring out the conception of predication they leave

---

[18]Richard does not offer this argument, but perhaps this more specific characterization of the cross-referencing role
of the variable can underwrite the analogy between the variables in standard first order languages and "variables" in L.

us with. I then contrast this conception of predicates with Richard's somewhat looser

understanding of them. Richard argues that certain expressions in English that create opaque

contexts can nonetheless be considered to express predicates. His argument centers on what I call

the "coherence of quantification test." A open sentence passes this test if a natural language

construction that suggests quantification in to the variable position in this sentence is intuitively

coherent. Richard contends that if an open sentence passes this test, then we have good reason to

think that it is a predicate. I argue that, although the intuitions of coherence to which Richard

appeals cannot be ignored, we need not accept his explanation of them.

The familiar arguments from substitution failure do not really attempt to show that all

instances of [PI] are true. We know from consulting the semantic and syntactic rules of a standard

first-order language that all instances of [PI] in that language will be true. The difficult problem,

which substitution failure arguments address, concerns the logical form of natural language

sentences in which substitution fails. These arguments suggest that we should not analyze a

sentence where substitution into a position fails to preserve truth value as an open sentence with a

variable in that position. We do so on pain of contradiction or unintelligibility.

These kinds of considerations lead Quine to adopt the following principle: If we cannot

substitute co-referring singular terms into a sentential position *salva veritate*, then we cannot

objectually quantify into that position either. More specifically, Quine adopts Quine's Thesis

([QT]).

> [QT] If (1)-(3) hold, then we cannot intelligibly substitute a variable for α in S and
> bind this variable with an objectual quantifier placed in front of S. In case (1)-(3)
> hold, let us call α's position in S an opaque position.
> 1) α and β are codesignative singular terms.
>
> 2) S is a sentence containing occurrence(s) of α and S' contains occurrence(s) of β in
> some of the places where S has α and is otherwise exactly like S.
>
> 3) S and S' do not have the same truth-value.

Quine's arguments for [QT] are somewhat obscure. Sometimes the problem appears to be purely logical or semantic, but other times Quine seems to point to a metaphysical objection to quantifying in.[19] The acceptability of quantification in often depends on what kinds of expressions α and β are, and what kinds of contexts we embed them in . Exegetical questions about Quine's arguments aside, we have reason to accept [QT] on solely logical and semantical grounds. In "Indiscernibility Principles", Richard Cartwright treats a paradox like that found in Quine's "Three Grades of Modal Involvement":

(7) It is a truth of astronomy that Hesperus = Phosphorus

(8) Hesperus = Phosphorus

(9) It is a truth of astronomy that Phosphorus = Phosphorus

(7) and (8) are true, but (9) is false. Why does this bar us from quantifying in? Cartwright suggests that we attend to the following problem: (9) is false, so its (wide-scope) negation (10) is true. The identity sign in (8) licenses the universal generalization in (11), and (12) is unobjectionable.

(10) & It is a truth of astronomy that Phosphorus = Phosphorus

(11) $\forall y$ (y = Hesperus $\leftrightarrow$ y = Phosphorus)

(12) $\forall y$(y = Phosphorus $\leftrightarrow$ y = Phosphorus)

If quantification into the context created by the sentence operator "it is a truth of astronomy that" is legitimate, we must affirm (13) and (14), which are existential generalizations on the conjunctions of (11) with (7) and (12) with (10).[20]

---

[19]For the view that the problem is logical or semantic, see Quine [1953] p. 148, Quine [1960] p. 167-8 For the metaphysical objection, see Quine [1953] p. 176, Quine [1961] p. 184.

[20]As I noted above, Richard would rightly object that assuming that EG is universally valid begs the question against him. I present this argument not as an objection to Richard, but as a way to bring out the conception of predication that underwrites it. Many variants of standard first-order languages restrict EG in some way. For instance in free logics, which contain singular terms which do not name anything in the domain, EG is not valid. In modal predicate logics, EG is restricted for various kinds of reasons in different systems. For instance, an object might have a property necessarily in every world in which it exists, so that Nec Fa is true in w, but the object might not exist in every world, so we cannot infer that $\exists x$ Nec Fx is also true in w. Richard has very different reasons for

(13) $\exists x(\forall y)[(y = x \leftrightarrow y = \text{Phosphorus}) \ \& \ $ It is a truth of astronomy that $y = \text{Phosphorus}]$

(14) $\exists x(\forall y)[(y = x \leftrightarrow y = \text{Phosphorus}) \ \& \ \& \ $ It is a truth of astronomy that $y = \text{Phosphorus}]$

But if we affirm both (13) and (14), then we say that there is some object, identical to

Phosphorus, such that it is a truth of astronomy that it is identical to Phosphorus and that it is not a

truth of astronomy that it is identical to Phosphorus. Cartwright concludes that there is "no settling

the question" whether the object identical to Phosphorus is such that it is a truth of astronomy that it

is identical to Phosphorus.[21] This is the conflict which quantification into opacity is meant to give

rise.

Cartwright's argument in (10)-(14) shows why quantification into opacity is unintelligible.

If we attempt to interpret the variables in (11) and (12) as purely referential, and if we hold that

existential generalization is a valid inference rule for these contexts, we must accept both (13) and

(14), which make conflicting demands on the object identical to Phosphorus. (13) states that this

object is such that it is a truth of astronomy that it is identical to Phosphorus, and (14) states that

this same object is such that it is not a truth of astronomy that it is identical to Phosphorus. For this

reason, we cannot accept both (13) and (14).

Cartwright shows that quantification into a context is barred when it could require us to

make conflicting predicate attributions to an object. In so doing, he explicitly connects the notion of

predication with objectual quantification. Quine similarly points out that the paradigm case of

referential position is that of a singular term under predication.[22] Tarski as well appears to assume

such a connection in constructing his truth definition. We saw that in "The Concept of Truth in

---

restricting EG. He can accept the substitutional inference that Fa implies that there is some variable letter v such that $\ulcorner \exists v F v \urcorner$ is true, but he sets constraints on which variable letter it will be.

[21]Cartwright [1987] p. 212. We might say that there is no question to be settled here.

[22]Quine [1960] p. 142

Formalized Languages", the formal notion of satisfaction is modeled on the everyday notion, which is conceptually connected to predication.

Predicates, for Quine, are expressions of a formal or a natural language with a certain kind of semantic rule. In formal languages, the issues are quite simple: an expression F is a predicate just in case its interpretation is a function from objects in the domain of quantification to truth values, and an sentence Fa is true just in case the referent of the singular term "a" is some object that the function associated with F takes to the value true. If the only thing that matters to the truth of an atomic sentence containing F is which object the singular term in the sentence refers to, and whether this object is one of those I(F) takes to the value true, then it can't matter how the object is referred to. If we follow Quine in taking this feature of predicates to be criterial of them, then it directly follows that EG is valid on any atomic sentence composed of an n-place predicate and n singular terms. So on Quine's view of what predicates are, it is difficult to even make sense of Richard's position.

Rather holding the debate on Quine's terms, Richard adopts another criterion for predicates which leaves the question about the truth of [PI] open. Quine thinks that we determine the intelligibility of quantification in to a sentential position somewhat indirectly, by asking whether substitution of co-referring terms into that position preserves the truth value of the sentence. Only if such substitution preserves truth value can we intelligibly quantify in. Richard proposes a more direct route to determine the intelligibility of such quantification in: simply look at a natural language construction which suggests quantification in to this context, and ask yourself whether the sentence is intelligible. I call this the "coherence of quantification test." Where Quine gives a similar treatment to all natural language constructions that do not permit substitution of co-referring terms *salva veritate*, Richard finds a distinction between those for which an quantificational construction is intelligible, and those for which it is not. Consider the following examples:

(15) Clark was sad.

(16) Lois thought Clark looked sad.
(17) Superman was so-called because of the red "S" on his shirt.

In the first two sentences, replacing the singular term "Clark" with a pronoun yields the intelligible

sentences (15') and (16'); replacement with a bound variable yields the slightly clunky but still

meaningful (15") and (16"). A similar replacement for "Superman" in (17) yields (17') and

(17"). Although (17') could be appropriately used if previous discourse supplied a name as an

antecedent for "he," it would not be intelligible outside of such a context. And (17") is clearly

unintelligible.

(15') He was sad.
(15") There is someone who was sad.
(16') Lois thought he looked sad.
(16") There is someone that Lois thought looked sad.
(17') He was so-called because of the red "S" on his shirt.
(17") There is someone who was so-called because of the red "S" on his shirt.

Richard thinks that these differences track a semantically and syntactically significant distinction

between the various kinds of opaque contexts, one that Quine's analysis ignores. The first four

sentences are intelligible in a way that the last two are not. We can coherently say that there is

someone Lois took to be sad, but not that there is someone so-called because of the red "S" on his

shirt. The verb "so-called" demands a name as well as an object as antecedent. But since people,

alone, can be sad, or can be taken by others to be sad, Richard thinks we should take expressions

like "Lois thought ___ looked sad" to be predicates just as "___was sad" is. Richard's point is not

that the move from (16) to (16') or to (16") is truth-preserving, since it is not, but rather that (16")

is not the incoherent mess Quine's arguments suggest, but appears perfectly intelligible. What it

appears to say is that there is some object with a certain property, and as such seems to meet an

important part of our intuitive idea of what predicates are. But if this pretheoretical notion allows

for predicates which fail Quine's substitution test, maybe we should conclude not that our

pretheoretical intuitions need to be explained away, but instead that Quine's substitution test

captures no important part of our ordinary notion of predication.[23] But then the formal language in which we represent the logical form of natural language sentences should give way, and allow us to represent all natural language predicates with the same kind of formal language expression. If we accept all of this, then we should also embrace Richard's claim that [PI] is false.

It is clear that Richard points to a significant distinction between the various kinds of opaque contexts. We cannot simply set these intuitions of intelligibility aside. Something must explain them. But the need for some explanation does not commit us to Richard's explanation. We might agree with Richard that the coherence of quantification test captures part of our intuitive notion of predication, without agreeing anything that passes this test is a predicate. Our intuitive notions are typically confused mixtures of many different interrelated concepts, and perhaps Richard is right to isolate the coherence of quantification test as a part of our intuitive notion of a predicate. But this by no means suggests that this test completely captures it. Our intuitive conception of predication also seems to require that for any predicate whatever, every object either does or does not satisfy it, and that if a sentence expressing a predication of an object is true, then the object must exist. But evidently an open sentence can pass the coherence of quantification test without meeting either one of these requirements.

An alternative explanation of the coherence of quantification into attitude contexts will be familiar to readers of Quine's "Quantifiers and Propositional Attitudes." There, Quine grants the coherence of the quantificational construction into attitude contexts, but denies that the logical form of these sentences can be read off the surface form in the straightforward way Richard suggests. According to Quine, propositional attitude verbs can express a triadic relation between a believer, a sequence of objects, and an intension. The singular terms denoting the objects in the sequence (and, of course, the believer) are in referential position, hence quantification into these positions is

---

[23]As any reader of "Quantifiers and Propositional Attitudes" is aware, Quine does not deny the intelligibility of (16′) and (16″). He simply denies that the mere fact that these sentences are intelligible suggests that they have the logical form of a singular term under predication. Richard rejects what he sees to be the syntactic manipulation inherent in Quine's approach, see below.

coherent. The intensions are sealed off from such quantification in, as well as certain other logical operations.

Richard rejects Quine's syntactic manipulations; he does not approve of playing fast and loose with logical form. He takes it to be a virtue of semantic theories generally that they respect the surface form of the natural language sentences. A theory that assigns a sentence that looks like a predication of an object a logical form with two main parts, one giving the meaning of the predicate and the other giving the meaning of the singular term, is *prima facie* preferable to a theory that assigns it a completely different logical form. But presumably Richard would agree that a theory that only reflects the surface form of natural language sentences at the cost of coherence is in no sense preferable to a coherent theory. Richard can only reject Quine's approach on these grounds if there is another *coherent* account that more accurately reflects surface form. In particular, he can only reject Quine's account in favor of his own semantics for attitude ascriptions, which is claimed to violate [PI] in much the same way L does, if he has shown already that [PI] is false. And this, as I have been arguing, is precisely what has not been shown.

Of course, there is a sense of "coherence" in which L is perfectly coherent; we know perfectly well what sentences of L say, and what they say is expressible in sentences of familiar first order languages. The surface forms of the sentences of L and their translations will, of course, differ considerably. But if we accept that L is coherent in this sense, the considerations of faithfulness to surface form still lend no weight to Richard's account. I have suggested that if Richard is to show that L violates [PI], he must show that what appear to be predicates and variables in L really are predicates and variables. But it is only on the assumption that these expressions are predicates and variables that we can faithfully represent natural language predicates and variables by means of them. Since we have no reason to accept this assumption, Richard is left without grounds to prefer his own account of what the coherence of quantification test shows to Quine's account.

6.   Richard offers a condition, [C], that he thinks all languages must meet if they are to contain

instances of the schema in [PI]. I have argued that this condition is too weak to do this work. I

have also argued that the specific arguments Richard employs to show that rule (3a) is a natural

extension of the semantics for variables or for predicates in more familiar languages are not

successful. But perhaps Richard could revise the rule (3a) in some way to evade objections

specifically directed against it. Here, I articulate a general condition that languages must meet if

they are to contain instances of the schema in [PI].

Central to any of the possible semantic rules we can give to predicates in different languages

is the notion of satisfaction of a condition. How can we ensure that the technical notion of

satisfaction of an open sentence is truly a generalization of the intuitive notion? If we consider only

open sentences with one variable free, [S] should constrain the technical notion so that it captures

the intuitive one.

> [S] If an object o satisfies (in the intuitive sense) a predicate F, then the infinite sequences
> with o in the ith place all satisfy (in the technical sense) Fx_i.

[S] makes the role of the variable in open sentences apparent. For one-place predicates and their

associated open sentences of a formal language, the variable simply supplies the assigned object.

In making this role of the variable explicit, [S] highlights the connection between all open sentences

composed of F followed by a variable. These sentences all say the same thing relative to an

assignment of the same object to the variable. Even though these sentences will be satisfied by

different sequences, they will ultimately be satisfied by the same objects.

> [S] generates the following parallel principle [A] which characterizes one-place predicates.

> [A] A predicate F is one place just in case it divides the domain of objects into two classes,
> those that satisfy it and those that do not, an atomic sentence containing F is true if and only

if the object assigned to the variable or the referent of the constant singular term is in the former class.[24]

[A] only characterizes one-place predicates, and I state it here to explain why L does not adhere to this principle, as it contains the expression F. [A] sets two constraints on predicates, and F satisfies only one of them. It does divide the domain of objects into two parts, those that satisfy F and those that do not, but whether an object satisfies an open sentence composed of F followed by a variable does not just depend on the assigned object. So according to [A] F is not a one-place predicate. I offer [A] as a general constraint on one-place predicates. Since it is the clear intention of the schema in [PI] that the substitution instances of α and $\alpha[y/x]$ in the schema in [PI] are to be understood as open sentences with one free variable, anything that can be substituted for these schematic letters must also conform to [A].

Richard may respond to this proposal in two ways. He might charge that requiring all languages in which instances of the schema in [PI] can be formed to conform to [A] begs the question against him. Second, he might reject the proposal in favor of a slightly weaker constraint that L meets. I address these responses in turn.

The first protest rings hollow. [A] motivates [PI], but the two principles say quite different things. It is hard to see how they could be thought to be equivalent, hence hard to see how it is question-begging to offer one in support of the other. First, [A] mentions variables, but its central role is to constrain the notion of a predicate. What motivates [A] is not simply the thought that a variable simple serves to supply its assigned object, rather a plausible conception of predication provides independent motivation. Second, [PI] is language-specific, and [A] is a general principle of predication that applies across languages. Certainly we can formulate different versions of [PI] for different languages, so in this sense what [PI] says is not specific to any

---

[24]A one-place vague predicate might not be thought to divide the domain in this way, as the domain may contain some borderline cases that do not clearly satisfy the predicate, but also do not clearly fail to satisfy it. I shall not discuss this case here, save to say that an account of vagueness that simply left the matter here would say nothing philosophically interesting about vagueness. Vague predicates are puzzling precisely because they appear to violate [A], yet conformity to [A] seems to be criterial of predicates.

particular language. But if, as I suggest, a language must satisfy [A] to even contain instances of

the relevant schema, then Richard must present another argument, rather different from what he has

presented, to show that a language need not conform to [A] to contain instances of the schema in

[PI].

Richard might also reject [A] in favor of a weaker constraint on a language, namely [A'].

[A'] differs from [A] only in that it contains "only if" (here in boldface) where [A] contains "if and

only if".

> [A'] A predicate F is one place just in case it divides the domain of objects into two groups,
> those that satisfy it and those that do not, an atomic sentence containing F is true **only if** the
> object assigned to the variable or the referent of the constant singular term is in that class.[25]

If Richard is right that [A'], rather than [A], appropriately characterizes one-place predicates, then F

is a one-place predicate, so it can be used to form an instance of the schema in [PI]. But there are a

few problems with Richard's view. First, [A'] seems arbitrary. Why should we think it preferable

to [A'']?

> [A''] A predicate F is one place just in case it divides the domain of objects into two groups,
> those that satisfy it and those that do not, an atomic sentence containing F is true **if** the
> object assigned to the variable or the referent of the constant singular term is in that class.

[A'] gives a necessary condition on the truth of atomic sentences formed with F, and [A''] gives a

sufficient condition. *Prima facie*, there is no reason to think that one of them presents a more

defensible criterion of one-place predicatehood than the other. The arbitrariness of selecting [A']

over [A''] counts against choosing one over the other. Perhaps Richard could adopt the disjunction

of [A'] and [A''] (rather than the conjunction, which is just [A]) as criterial of one-place predicates.

This disjunctive condition is not as weak as [C], but still weaker than [A']. But at this point the

choice of criterion appears to be completely unmotivated by any semantical considerations, indeed it

---

[25]Richard has suggested [A'] in conversation. One might see this principle as supplanting condition [C].

fails to be motivated by anything at all, save the desire to work out some construal of predication that can admit false instances of the schema.

7.   Richard claims that our interest in languages like L should be more than theoretical, since natural languages also violate [PI]. In support of this claim, he offers several natural language examples of purportedly false instances of the schema in [PI], such as the following: Suppose that John utters (18).

> (18) Last night, I observed the planet Hesperus and then the planet Phosphorus. I was disappointed; I wanted to observe Phosphorus and then Hesperus.

Richard says we would speak truly in uttering (19), but not (20).

> (19) There are planets x and y. John said that he saw x and then y, but he wanted to see y and then x.

> (20) There are planets x and y. John said that he saw x and then y, but he wanted to see x and then y.

Suppose also that John says nothing else about planets except for what he reports in (18). Richard claims to have ordinary semantic intuition on his side in supposing that given the truth of (18), (19) is true but (20) is false. If Richard is right, then we have an instance of the schema in [PI] which is false, namely:

> (21) $\forall x \forall y$ (x=y $\rightarrow$ (John said that he saw x and then y, but he wanted to see y and then x
> $\rightarrow$ John said that he saw x and then y but he wanted to see x and then y.))

Richard claims that (21) is false under the assignment of Hesperus to x and to y. Even though (21) is false, the existential quantifiers in (19) and (20) are objectual. Hence, according to Richard, we have a counterexample to [PI] in a language whose quantifiers are all objectual.

But is it so clear that ordinary semantic intuition is on Richard's side here? First, we should notice that the only kinds of intuitions about meaning that our semantic theories for natural

languages should explain are those concerning natural language sentences. Certainly there is objectual quantification in English. We also have variables (i.e. pronouns) in English. However, "x" and "y" are not among the pronouns of English. As they are used in (19) and (20), they may not even be expressions of English at all. Further, I am not sure that we can replace the expressions "x" and "y" in (19) and (20) with pronouns of English, particularly if we require that these pronouns get a bound reading. In formal languages, a formula containing different variables may be satisfied by a sequence that assigns the same object to both variables. But in natural languages, if two different pronouns are used in the same sentence, there is often a presumption that the pronouns refer to different objects. So a sentence that used different pronouns to refer to the same object would be at least misleading. If we can find natural language replacements for (19) and (20), then the important semantic intuitions would concern those sentences, not (19) and (20). If there is no replacement possible, then it is not clear what weight we should give to our intuitions about the meanings of (19) and (20). Prior understanding of the semantics for formal languages will no doubt color our intuitions here.

Even if we can legitimately take our semantic intuitions about (19) and (20) into account, it is not clear that Richard is right about our intuitions. Can (19) or (20) be correctly used to report what John said? The first observation is that John is confused. He thinks that there are two objects, Hesperus and Phosphorus, but unbeknownst to him these objects are one. Had John known that Hesperus and Phosphorus are the same planet, he probably would not have spoken as he did in (18). If John had said that he saw Mars before Venus, but wanted to see Venus and then Mars, then Richard could rightly maintain that (19), but not (20), could represent his belief. However, my intuitions about whether (19) or (20) accurately represents John's belief are only firm in the case where John's belief really does concern two different planets. It seems wrong to represent John's belief in such a way as to leave it open whether he thinks it concerns one planet or two. It is of course difficult to respond to an argument from semantic intuition by professing to

have different intuitions, but I hope these considerations cast some doubt on Richard's claim that semantic intuition is clearly on his side.

Finally, even if we can have semantic intuition about (19) and (20), and Richard is right about these intuitions, Richard needs something more to show that (21) is a counterexample to [PI]. To show that (21) is truly an instance of the schema in [PI], he needs to show that the correct semantic theory must treat the quantifiers in (21) objectually. No semantic intuition about (19) and (20) can establish this. Semantic intuition, at least the kind that semanticists need to take seriously, just concerns the truth conditions of natural language sentences. The objectuality of the quantifiers is another matter entirely. Whether the quantifiers are objectual, as Richard rightly points out, depends on whether we give a Tarskian definition of satisfaction for the language. That is, it depends on the technical apparatus we use to interpret the natural language quantifiers within our semantic theory. So views about objectuality of the quantifiers are no part of the intuition a semantic theory is meant to explain, rather they are part of the theory itself. Our semantic theory can accommodate the alleged intuition that (19) is true and (20) is false simply by failing to interpret the quantifiers objectually in this instance, or by tacitly introducing quantification over senses or expressions, as Quine does in "Quantifiers and Propositional Attitudes." For reasons I mentioned above, Richard rejects the manipulations of the surface structure of the sentence that this approach requires.[26] It is not clear that these reasons alone support the kind of semantical manipulations implicit in Richard's approach.

But let us suppose that Richard is correct about semantic intuition, and that the correct semantic theory must interpret the quantifiers objectually. How can his semantics for attitude ascriptions accommodate this?

Richard finds the motivation for a Russellian semantics of attitude ascription compelling, largely because these theories provide a simple way of accommodating the philosophical view that

---

[26]Richard [1987] p. 558

names, indexicals and demonstratives are directly referential. Nonetheless, he is troubled by the

inability of the Russellian view to explain our intuitions about Hesperus-Phosphorus cases. We do

not feel inclined to say of the ancient Babylonian astronomer that he believes that Hesperus is in the

east if he calls the planet he sees there "Phosphorus" and thinks that Hesperus is a different planet

entirely. But on a Russellian theory, "Hesperus is in the east" and "Phosphorus is in the east" both

express the same proposition, and belief is simply a relation between a believer and a proposition.

So according to the Russellian it would not even be possible for someone to believe that Hesperus

is in the east while disbelieving that Phosphorus is in the east.

Richard calls an account of propositional structure and attitude ascription "Russellian" if it

adheres to several basic tenets of Russell's proposal in the *Principles of Mathematics*: Richard

states that these tenets are the following[27]:

> 1. A proposition is structured, and it takes its structure from the canonical sentence used to express it.

> 2. The contents of ordinary proper names, demonstratives and indexicals are the referents of these expressions.

> 3. The contents of expressions other than singular terms are functions or relations.

> 4. The content of "believes" is a two-place relation between believers and propositions that meet (1)-(3).

Any theory of this kind will have difficulty explaining our intuitions on the case described

above. Since the content of a proper name is just its referent, "Hesperus is in the east" and

"Phosphorus is in the east" express the same proposition. Since belief is just a relation between a

believer and a proposition, there will be no way of explaining how one person could believe one but

not the other. Richard wants to formulate a theory that respects some of the main tenets of and

motivations for a Russellian theory but has something more satisfying to say about these cases.

---

[27]Richard [1990] p. 109

Richard's account is complicated, but most of the complexity is brought in to explain the

Hesperus-Phosphorus cases, in which two attitude ascriptions intuitively differ in truth value but

the sentences only differ in containing different proper names of one object. Part of this apparatus

drops out of the picture in his treatment of ascriptions containing pronouns in place of proper

names, so I will not present all of it here.[28] Essentially, Richard identifies propositions with

ordered n-tuples composed of ordered pairs of the syntactically simple components of a sentence

and their Russellian contents. He calls such an ordered n-tuple a "Russellian annotated matrix"

(henceforth "RAM"), and the ordered pairs that compose it are called "annotations." For instance,

the proposition the English sentence "Twain is dead" expresses is represented as [R], the ordered

pair of the syntactically simple components of the sentence, that is, "Twain" and "is dead", each

paired with their Russellian contents.

[R] <<"Twain", Twain>,<"is dead", being dead>>

We can use sets of RAMs to represent everything that a speaker believes. Richard calls sets of

RAMs of this kind "representational systems."

Richard thinks that "Ed believes that S" is true only if that the RAM expressed by S

determines, under what Richard calls a correlation, a RAM in Ed's representational system. A

correlation is a function which takes annotations to annotations which have the same second term,

that is, which preserve reference. For instance, a function taking the annotation <"Twain", Twain>

to <"Clemens", Twain> is a correlation, since the second term is the same in both annotations. So

for "Ed believes that Twain is dead" to be true, there must be some function that maps <"Twain",

Twain> onto another annotation which also has Twain as its second member (for instance,

<"Clemens", Twain>) and which also maps <"is dead", being dead> onto another annotation with

---

[28]We can construct cases like the Hesperus-Phosphorus case that do not require the use of two different names for
the same object in two attitude ascriptions. Suppose I see someone walking through the halls of Building 20, and I
judge this person to be tall. Later, I see the same person turning the corner, only this time I don't judge him to be
tall. I might think "That guy is taller than this guy" without thinking "This guy is taller than himself". The fact
that these cases do not turn on the presence of proper names in the ascription suggests a *prima facie* problem for a
semantic theory that treats the two cases differently.

the same second member (maybe here it is <"is dead", being dead>), and the resulting RAM

<<"Clemens", Twain>, <"is dead", being dead>> is in Ed's representational system.

      If the singular term in S is a pronoun instead of a name, then the necessary condition is also

sufficient. Richard thinks that variables, like other singular terms, are syntactically simple

components of sentences. So on his view, a sentence containing a variable expresses a RAM

containing an annotation for the variable. He says that an open sentence such as (23) determines the

RAM (24) relative to an assignment of the object, say Twain, to the variable.

      (23) x is dead

      (24) <<"x", Twain>,<"is dead", being dead>>

Of course, in English, we do not ascribe a belief in an open sentence relative to some assignment to

the variable. We do, however, ascribe beliefs using sentences like (25).

      (25) There is something Ed believes to be red.

Richard thinks the ascription in (25) is true if and only if there is some object X such that the RAM

in (26) is in Ed's representational system under some correlation.

      (26) <<"x", X>,<"is red", being red>>

That is, Ed must have in his representation system some RAM like (26) except perhaps in having

another singular term in place of "x". So (25) is true if there is some object Ed believes to be red.

In general, Richard's account will validate all existential generalizations where the variable inside

the belief context is bound from the outside.[29]

      What does Richard's semantics say about (19) and (20)? When we consider sentences

containing more than one occurrence of the same variable, we see that since correlations must be

functions, what look like false instances of the schema in [PI] are possible. Richard exploits the

fact that the variables x and y occur in different places in (19) and (20). He says that (19) and (20)

determine the RAMs in (27) and (28), respectively.

---

[29]An obvious problem with validation of every existential quantification into an attitude context is the replacement of an empty name with a variable. Richard specifically defers discussion of this problem.

(27) <<"O",O>,<"x",V>,<"y",V>,<"y",V>,<"x",V>>

(28) <<"O",O>,<"x",V>,<"y",V>,<"x",V>,<"y",V>>

Here, V is the planet Venus, and "O" abbreviates the sentential function in (19) and (20). Since

correlations are functions, they must map both occurrences of <"x",V> to the same annotation, and

both occurrences of <"y",V> to the same annotation. But since there is no way of mapping one

annotation onto both occurrences of <"x",V> and one annotation onto both occurrences of <"y",V>

in (27) that yields something of the form of (28), there will be no guarantee that a person who

believes (27) will also believe (28). If there is no such guarantee, then there is some circumstance

in which (19) is true and (20) false, yielding the purportedly false instance of [PI] in (21).

8.   At the outset of this chapter I stated two aims: first, to demonstrate that Richard has offered

us no reason to reject [PI]; and second, to identify two sources of tension between Richard's

semantics for attitude ascriptions and part of his philosophical motivation for it. I have made my

case for the first, I turn now to the second.[30] Although Richard's semantic theory, like L, contains

false sentences that look like instances of the schema in [PI], their appearance is deceptive. These

theories contain expressions that look like, and are called, "predicates" and "variables," but these

names are simply misleading. The result is that sentences of both languages mask the structure of

what they express. When the real structure of what is expressed is made apparent, it is far from

clear that the theory is motivated by certain considerations Richard brings to bear in support of it.

    In particular, Richard argues in "Quantification and Leibniz's Law" (henceforth QL) that we

have certain defeasible grounds to prefer his own semantics for attitude ascriptions over Quine's

"Quantifiers and Propositional Attitudes" proposal, sketched above. The semantic theory that

---

[30]According to Richard, his semantics for attitude ascriptions violates [PI], but not in quite the same way as L
purportedly does. So Richard presumably thinks that the truth of [PI] would present a significant technical problem
for this semantics for attitude ascriptions. One might then assume that this is the tension I had in mind, however
this assumption would be mistaken. As is clear, I think that [PI] is true, hence I think it is a mistake to say that
any language violates it. Further, I think it is a mistake to suppose that the truth of [PI] presents any technical
problem whatsoever for Richard's semantics of attitude ascription.

Richard defends in QL is not the same as the one he offers in *Propositional Attitudes* (henceforth PA). However, in the latter, Richard offers similar grounds for preferring his new theory over the Quinean theory, which is there called a "Fregean" theory.[31] In neither case, however, do the particular considerations offered give us reason to prefer Richard's theory over the alternative.

Richard notes that predictive accuracy is not the only ground we have to assess different semantic theories of natural languages. He takes it to be a desideratum on a semantic theory that it predicts logical forms for natural language sentences that respect the surface form of these sentences. Of course, other considerations, such as theoretical unity and overall simplicity, can override. But at all events if semantic theory A respects surface form while semantic theory B does not, we should, *ceteris paribus*, prefer A to B. Richard suggests that both the straightforwardly Russellian proposal endorsed in QL and the less straightforwardly Russellian proposal of PA respect the surface form of attitude ascriptions in a way that a particular alternative does not, hence we should, other things equal, prefer his theories to these particular alternatives.

How can Richard's PA proposal offer a more faithful representation of the surface form of certain attitude ascriptions than what he calls, in PA, the "Fregean" alternative? Richard notes that sentences like (29) are naturally represented by the existential quantification expressed in the Russellian analysis (30), and not by the Fregean treatment of (29) in (31):

(29) There is a man such that Odile believes that he is dead.

(30) $(\exists x)(x$ is a man & Odile believes that x is dead).

(31) $(\exists x)(\exists y)(y$ is a man & x is a sense that presents y & Odile believes ^{s}x is dead$^s)$[32].

---

[31]See Richard [1990] p. 67 for Richard's exposition of this version of the Fregean theory and p. 150-51 for the discussion treated here.

[32]The "^{ss}" are sense-quotes which are to be understood on an analogy with Quine's corner quotes For instance, if "x" is a variable ranging over senses which has been assigned the sense of some proper name t, then ^{s}x is tired^s refers to the sense of the sentence "t is tired".

Richard thinks his own treatment of (29) is more akin to the Russellian (30) than to the Fregean (31). On his view, (29) is true just in case there is a man X such that (32) represents one of the RAM's in Odile's RS.

(32) <<"is dead", being dead>, <"x", X>>

Richard's treatment of (29), like the Russellian treatment, involves only one explicit quantification, quantification over objects. Since (29) itself also appears to involve only one quantification, this appears to be a point in favor of both Richard's treatment and the Russellian treatment in (30) over the Fregean (31), at least from the perspective of structural similarity between the natural language sentences and their representations in the semantic theory. But when Richard's treatment is completely spelled out, it is not obvious that its treatment of quantification into intensional contexts is as straightforward as it appears. For Richard, (29) is true just in case there is a RAM in Odile's RS which is like (32) except that, first, in place of the X there is some man, and second, in place of the "x" there is some proper name, indexical, demonstrative, or variable that has the second term of the annotation, i.e. the man Odile believes to be dead, as its Russellian content. Now Richard's treatment of (29) appears to involve two quantifications, one over objects and another over expressions. It is in this sense more closely akin to the Fregean treatment (31) than the Russellian (30). Rather than explicitly introducing the second quantifier as a syntactic object, Richard builds it into the semantics, specifically into the representation relation, in that one RAM represents another just in case they are identical except perhaps in the first member of some of their constituent annotations.

What is at issue here, as well as in Richard's claim to have produced a counterexample to [PI], is whether and to what extent the syntactic categories of a language can accommodate semantic deviation. It is not easy to say what it is for the logical form that a semantic theory predicts for a natural language sentence to respect the surface form of that sentence. One would think that the formal language representation r of a natural language sentence s should at least have semantic

constituents corresponding to the semantically significant constituents of s, and that r should have no other semantic constituents. The sort of correspondence between the semantically significant expressions of r and those of s is not simply a one-to-one mapping; the correspondence should be between expressions of the same semantic type. For instance, a quantifier in s should correspond to a quantifier in r, and a predicate in s should correspond to a predicate in r. As I have argued above, whether a formal language expression should count as a member of a certain semantic category, such as predicates or quantifier, depends on the sort of semantic rule it has. If its semantics deviates too far from what is typical for expressions in that semantic category, then it cannot be counted as an expression in that category. Hence, although we can use such expressions to give a semantics for natural language sentences, we cannot say that we reflect the surface form of these sentences by using such expressions to represent their semantic constituents.

The second tension between Richard's semantics for attitude ascriptions and his philosophical motivation for it issues in part from his claim that his semantics for attitude ascriptions preserves what is attractive about the Russellian theory while delivering a more acceptable response to the puzzle cases that cause it so much trouble. Russellianism requires that the objects that our thoughts are about are constituents of the propositions we believe. It also requires that the quantifier in a sentence like "There is someone Alice believes to be happy" is an objectual quantifier, and that the objects it ranges over are the sorts of everyday objects that Alice might believe to be happy. Yet, if the atomic open sentences of the language are interpreted so as to yield what look like violations of [PI], then there is no interesting sense in which an open sentence under an assignment of objects to its variables is *about* those objects.

Richard's official view is that the objectuality of the quantifiers of a language has nothing to do with the base clause of the definition of satisfaction for the language. All that is required to guarantee the objectuality of the quantifiers is that the quantifiers themselves should have a recursive semantics. This requirement is stated in Richard's condition [C], considered above. As I noted

above, [C] sets no constraints on the base clause of the definition of satisfaction for the language, hence it permits *any* function from sets of sequences to atomic open sentences as a base clause. For instance, condition [C] allows functions according to which Fx is satisfied by all sequences, and Fy is satisfied by none. I noted above that if such base clauses are permitted, then there is no natural sense of "about" in which the sentences of the language are about the objects in the domain of quantification. The problem for Richard's account of propositions, then, is that there is no natural sense in which they are about the everyday objects that we naturally take our beliefs to be about.

Much of the preceding attempts to show that Richard has given us no good reason to think that there are false instances of the schema in [PI], and further that there is good reason to think that there could not be a false instance of this schema. Richard claims that his denial of [PI] is neither a denial of the metaphysical platitude that identical objects have all of their properties in common nor an affirmation of the truism that we can construct a language in which what looks like an instance of the schema is false. I have argued that no other route is open to him.

Since his semantics is meant to be broadly Russellian, in that the objects the thought is about are actually constituents of the proposition, he wants the quantifiers in his formal theory to be objectual. Hence, he needs to show that a language containing only objectual quantifiers can still contain false instances of the schema in [PI]. If his semantics for attitude ascription requires that a language like [L] have a Tarskian semantics, then his failure to show that a truly Tarskian semantics for such a language is possible will make a defense of his theory as a Russellian theory untenable.

2

Three Arguments about Belief

1.   Some time ago, John Perry raised a new problem for what might be called the "traditional

picture of belief," according to which someone's having a belief is that person's standing in a

certain relation, the belief relation, to a proposition.   Perry's problem concerns indexical belief,

for which I give the following rough semantical characterization: an indexical belief is a belief

expressible with a sentence containing an indexical pronoun such as "I," "here," or "now," or a

demonstrative pronoun such as "this," or "that."[33]

On the contemporary philosophical scene, one finds the view that Perry has presented us

with a *special* problem for the traditional picture, a problem that non-indexical belief does not

present. As with many views that one finds in this way, it is not clear who lays claim to it. Perry

does not explicitly commit himself one way or the other. He asserts that "the essential indexical

poses a problem for various otherwise plausible accounts of belief,"[34] and in illustrating the

problem he restricts himself to examples of indexical belief as I have characterized it above. David

Lewis explicitly claims that a certain kind of belief presents a unique problem for his favored

version of the traditional picture, on which propositions are understood to be sets of possible

worlds.[35] He says that beliefs in this special class are the *only* ones that require him to say that

sets of possible individuals, rather than sets of possible worlds, are uniformly the objects of the

---

[33]I intend this characterization simply as a convenient way of indicating the sorts of beliefs that are at issue. I do not think that expressibility with indexicals is at all essential to indexical belief. For instance, I think that someone whose language contained no indexical expressions might still be able to have what I call an "indexical belief."
[34]Perry [1979] p. 4
[35]Lewis [1979]

attitudes. Thus, he suggests that he takes the problem they raise to be a unique one. Although Lewis uses the same sorts of examples as Perry does to illustrate the problem, he is not explicit about what sorts of beliefs are members of this problematic class.[36] So although he does think that indexical belief presents a unique problem, it is not clear that he would endorse my semantical characterization of the class of beliefs that presents this problem. Hector-Neri Castañeda is most explicit in claiming to identify a special problem about the logical form of *ascriptions* of indexical belief, but his points do not carry over in any obvious way to indicate a special problem about indexical belief itself.

At any event, the claim that indexical belief presents a unique problem for the traditional picture of belief, which I shall call "the uniqueness claim," is *prima facie* plausible, so it is worth considering what, if anything, can be said in favor of it. What is problematic about the uniqueness claim is that the arguments offered to demonstrate that there is a problem about indexical belief at all typically show that various kinds of non-indexical belief present exactly the same problem for the traditional picture.

In what follows, I consider three arguments that indexical belief presents a problem for certain features of a traditional picture of belief. I show that even if they are sound, they give us no reason to adopt the uniqueness claim. I then consider what we should conclude from the failure of these arguments to establish the uniqueness claim. I show that the third argument, explicitly focusing on the special semantical features of the sentences that express indexical belief, does point towards a way in which indexical belief presents a unique problem for an account of what communication of indexical belief consists in, and thus a unique problem for any account of belief. Thus I conclude that we do have reason to adopt the uniqueness claim.

---

[36]Since he gives a unified treatment of the objects of the attitudes, taking all of them to be sets of possible individuals, there is no reason for him to precisely delineate the kinds of attitudes that require this revision of his version of the traditional picture.

2.   We can briefly state the central claims of the traditional picture of belief. First, according

to this picture belief is a relation between a believer and a proposition, or more precisely,

> [Belief] A's having the belief that s is A's standing in the belief relation to the
> proposition that s.

Second, there is something about the internal state of a thinker, which we call, without

making any claims as to its nature, a belief state. Someone has a belief, say the belief that grass is

green, in virtue of being in a certain belief state. Our belief states are among our intrinsic

properties, to which we have immediate non-inferential access. Thus, belief states are supposed to

provide for what seems quite natural, that we have *a priori* access to our beliefs, and that by citing

a person's beliefs we can explain his or her behavior.[37]

That beliefs have a psychological component is something no one could deny, but the

traditional picture offers us a particular understanding of the semantical relevance of the

psychological component of belief. A complete explanation of the way in which the psychological

component of belief is supposed to have its particular semantical relevance is beyond the scope of

this exposition of the picture. But the view has a consequence which is of interest to us here. That

is, on this picture, belief states alone determine the truth conditional content of the belief someone

has in virtue of being in that state. So the traditional picture of belief includes the following Belief

State Criterion:

> [Belief State Criterion] If A and B are in the same belief state, then the propositions that
> they believe in virtue of being in that belief state have the same truth condition.

The dominance of the New Theory of Reference in contemporary thought about belief and

language makes it hard to see how anyone could find [Belief State Criterion] at all plausible, but

this criterion quite naturally falls out of a Fregean approach to belief. On the Fregean view,

thinkers have "mental fixes" on the objects their thoughts are about simply in virtue of having

---

[37]There is good reason for finding the notion of a belief state deeply problemmatic. I discuss these reasons in some detail, as they bear on the problems and arguments I consider, in section 4 of this chapter. At this point, I mean only to adopt a somewhat vague notion found in the literature on belief.

those thoughts. A "mental fix" is necessarily a fix *on* a particular object or property, just as a road

is necessarily a road that goes somewhere. Only when the thought is fully determined by means of

these "mental fixes" on objects and properties is it appropriate to use language to identify its

content. Since it is possible to have two distinct "mental fixes" on the same object, it is possible

for a thinker to have truth conditionally equivalent, but cognitively distinct, beliefs. But the nature

of the mental fix makes cognitively equivalent and truth conditionally distinct beliefs impossible.

In the next six sections, I consider three arguments that this picture of belief, which I call

throughout "the traditional picture," cannot accomodate indexical belief. That is, the phenomenon

of indexical belief is a *counterexample* to this traditional picture. I then argue that even if these

arguments are sound, they do not provide any obvious means of support for the uniqueness claim.

3.   Although it carries a commitment to these purely internal belief states, the traditional picture

gives us no guidance so far on how we can find out that two people are in the same belief state.

Perry proposes that we can sidestep some of the difficulties involved in saying what belief states

are if we have some way of classifying agents into groups, all of whom are in the same belief

state.   He proposes, in particular, that we classify an agent's belief states using sentences that

these states dispose him or her to utter.

> When we believe, we do so by being in belief states. These states have typical
> effects, which we use to classify them. In particular, we classify them by the
> sentences a competent speaker of the language in question would be apt to think of
> or utter in certain circumstances when in that state. To accept a sentence S is to be in
> a belief state that would distinguish speakers who would think and utter S from
> those who would not.[38]

I shall adopt Perry's notion of a speaker's acceptance of a sentence in what follows. This notion is

no part of his account of the traditional picture, rather it is a central part of the theory he suggests

we adopt *in place of* the traditional picture. Below, I model a psychological criterion of

---

[38]Perry [1980] p. 53-4

individuation for belief states, [Psychological Criterion], along the lines suggested by Perry's

sentential classification scheme. This criterion should not be understood as a part of the traditional

picture of belief. Rather, it is meant to be an *independently* plausible principle.

> [Psychological Criterion] Fix a language L; if two speakers of L, A and B, accept
> the same sentence of L, then A and B are both in the belief state that this sentence
> classifies.

The sentences of L to which [Psychological Criterion] applies must be free of obvious

kinds of ambiguities and context-dependence that undermine its intuitive force. If A and B both

utter "I'll meet you by the bank," and A means to refer to a place that charges high fees, and B

means to refer to the edge of a river, then clearly A and B are not in the same belief state with

respect to this sentence. Less clear-cut cases of context-dependence also pose a threat to

[Psychological Criterion], so must be excluded from its purview. For instance, suppose A and B

sincerely utter the sentence "John ought to have been here by now." According to [Psychological

Criterion], they are both in the belief states that this sentence specifies. Yet it is possible that A

meant that he expected John to have been there by then, and B meant that John was under some

moral obligation to have been there by then. In this case, it would be unnatural to say that A and B

are in the same belief state. Importantly, the context-sensitivity of indexicals is *not* supposed to

threaten their univocality in the way that these context-dependencies do. A person who uses an

indexical to express a belief is supposed to be thinking of the referent of that indexical in a certain

way, and this way of thinking of the referent associated with an indexical is again supposed to be

constant throughout its uses, even though the object so thought of differs from use to use.

Perry suggests that if we accept [Psychological Criterion], then we should also say that

indexical belief is a counterexample to the traditional picture of belief. How is this so? Perry

invites us to consider the case of Hume and the delusional Heimson, who thinks he is Hume.

"Let us imagine David Hume, alone in his study, on a particular afternoon in 1775, thinking to himself, "I wrote the *Treatise*." Can anyone *else* apprehend the thought he apprehended by thinking this? First note that what he thinks is true. So no one could apprehend the same thought, unless they apprehended a true thought. Now suppose Heimson is a bit crazy, and thinks himself to be David Hume. Alone in his study, he says to himself, "I wrote the *Treatise*." However much his inner life may, at that moment, resemble Hume's on that afternoon in 1775, the fact remains: Hume was right, Heimson is wrong. Heimson cannot think the very thought that Hume thought to himself, by using the very same sentence."[39]

Perry uses this example to state what I call the Belief State Argument. Consider the following pair

of propositions:

p = the proposition Hume asserts in uttering "I wrote the Treatise."

q = the proposition Heimson asserts in uttering "I wrote the Treatise."

Suppose that Hume believes p and Heimson believes q. According to [Psychological Criterion],

Hume and Heimson have these beliefs in virtue of being in the same belief state, as clearly Hume

and Heimson are both speakers of English, they utter the same sentence which is free of any

obvious ambiguities that undermine the intuitive force of [Psychological Criterion]. Thus, according

to [Belief State Criterion], the beliefs they have in virtue of being in this belief state are truth

conditionally equivalent. Yet it is evident that p and q differ in their truth conditions, p is about

Hume, and q is about Heimson. Our supposition, that p and q are the objects of Hume's and

Heimson's beliefs, in that p and q are the propositions to which they stand in the belief relation,

thus implies something false. Hence, on the assumption that the other premises of the Belief State

Argument are true, the supposition itself is also false. We thus conclude that the Hume-Heimson

case is a counterexample to the traditional picture of belief. In the next section, I will consider how

we should best respond to the problem posed by the Belief State Argument, and whether the

problem is uniquely posed by indexical belief.

---

[39]Perry [1977] p. 16-17

4.   The Belief State Argument is, I think, evidently valid. The conjunction of the traditional

picture with [Psychological Criterion] cannot accomodate the Hume-Heimson case. What I shall

consider in this section is whether this problem is uniquely raised by indexical belief. But it will

also be useful in passing to consider how we ought to respond to the problem raised by the Belief

State Argument. So I shall begin with a consideration of our options in response to this problem.

[Belief] appears to be a neutral principle, but its very neutrality between competing accounts

of belief, and of propositions, masks more substantial disagreements among those who accept it.

These disagreements, however, are irrelevant to the issues at hand, as everyone who accepts

[Belief] also thinks that propositions have truth conditions, and it is simply this assumption about

the nature of propositions that gives rise to the problem. And clearly this consequence of [Belief] is

well-motivated on intuitive grounds. Beliefs can be true or false, so it is natural to say that the

objects of belief have truth values. Ordinary speech suggests that belief ascriptions state that a

certain sort of relation holds between a person and the referent of a that-clause, i.e. a proposition.

These intuitions lend *prima facie* support to [Belief], so we ought not reject it if there is a more

intuitively compelling solution to the problem.

Fortunately, there are compelling reasons to doubt both [Psychological Criterion] and

[Belief State Criterion]. I shall consider only a few such reasons here. First, contrary to

[Psychological Criterion], there may be no interesting correlation between an agent's internal belief

states and the sentences he or she accepts. But despite this quite general problem about the

classification of belief states, it is difficult to deny the specific intuition about the Hume-Heimson

case, that Hume and Heimson are in the same belief state, even if this is incidental to the fact that

they accept the same indexical sentence. Granting this leaves us no choice but to reject [Belief State

Criterion]. The Belief State Argument itself forces only a minor modification to [Belief State

Criterion], but certain further considerations provide reason to reject the claim that a purely internal

belief state can determine anything outside of the mind. After discussing these three reasons for

rejecting the two contentious premises of the Belief State Argument, I show that even if we accept the notion of a belief state that this argument requires, the argument still lends no support whatsoever to the uniqueness claim, since beliefs expressed with natural kind terms present the very same problem for the traditional picture.

What reason is there to reject [Psychological Criterion]? Let us suppose for the moment that there are wholly internal belief states that determine, either partially or entirely, the propositional contents of belief. Can a speaker's utterances of indexical sentences be used to identify these belief states, as [Psychological Criterion] claims? Certain sorts of sentences containing indexical expressions do not correspond in any interesting way to the belief states of those who utter those sentences. "He" and "it" present particularly clear cases--there is just no interesting sense of "psychologically similar" in which all those who utter, say, "He's standing next to it" are psychologically similar. We have no reason to think that all such agents think of the referents of "he" and "it" in similar ways, no reason to think that the beliefs they express by uttering that sentence dispose them to act in similar ways under similar circumstances, in short, no reason to think that they are in the same internal belief state.

Beliefs expressed with "this" and "that" present a similar, though less immediate, difficulty for Perry's claim. As I characterized indexical belief at the outset of this chapter, such beliefs are indexical. And the characterization is quite natural; at least as a starting point. Much as there is a characteristic way of thinking of oneself associated with "I," and a characteristic way of thinking of a location associated with "here," there is a characteristic way of thinking of objects demonstratively identified, and these beliefs are paradigmatically expressed with sentences containing "this" or "that." How do these beliefs present a special problem for [Psychological Criterion]? Consider this situation: Lisa and Judy are standing next to a vase full of decaying flowers. Both say "that smells atrocious." Are they in the same belief state? Suppose that Lisa and Judy want to avoid bad-smelling things, and both move away from the vase of flowers. We can then explain why both

move away from the vase by citing the beliefs that they expressed in uttering "That smells atrocious." Suppose further that Lisa is standing to the right of the vase, and Judy is standing to the left of it, so in moving away from it, they move in different directions. But since Lisa and Judy both accept the same sentence ("That smells atrocious") we are at a loss to explain why they move in different directions. But clearly this difference is a difference in what they do, an intentional difference, and as such, there should be a belief-desire explanation for it.

Suppose there was some other indexical sentence that Lisa accepts and Judy does not. Would this save the assumption of [Psychological Criterion] that two people who assertively utter the same indexical sentence are in the same belief state? Not necessarily. First, if we ask which of Lisa's or Judy's beliefs explains why both recoiled from the vase, the most likely answer would be some belief expressible with the word "this" or "that", and not some description of their locations with respect to the vase in absolute space. Even though they are in different positions with respect to the vase, and they are aware of this, this difference in their beliefs does not appear to be manifest in a difference between the indexical sentences they accept. And this is to be expected. Our beliefs about our immediate environment, particularly about the objects around us and the spatial relations between them, affect our actions in subtle ways that outrun our linguistic capabilities. We can maneuver in and manipulate our surroundings in response to information gleaned from perception, long before we are able (if we are ever able) to state that information with the degree of precision that explains the exact actions we perform. Thus, it is not at all clear that indexical sentences appropriately categorize belief states as [Psychological Criterion] states.

This success of this objection to Perry depends on the somewhat contentious claim that the beliefs that explain a person's behavior directed towards an object in that person's environment are the sorts of beliefs that person would express with "this" or "that". One might object to this claim on several fronts. One might say, for instance, that one only explains behavior directed towards objects in the immediate environment of the agent by citing a number of indexical beliefs, not just a

single belief of the sort expressible with "this" or "that". Gareth Evans, for instance, holds that we cannot expect to give an account of each kind of indexical or demonstrative reference without bringing in the other kinds. He thinks that the way we think of ourselves in using "I", for instance, is inextricably bound up with the way we think of our locations in space using "here".[40] Further, the claims I have made here about demonstrative belief do not appear to be extensible to certain sorts of indexical belief, specifically the sorts of indexical belief expressible with sentences containing "I," "here," and "now." Even if there are many ways of thinking of a oneself associated with the pronoun "I", and many ways of thinking of one's spatial location associated with "here", all of these ways play the same role in the explanation of certain kinds of behavior. So it is convenient, and appropriate, to follow Perry in classifying at least this narrow class of belief states by means of the "I" and "here" sentences we use, or would use, to express them. Since the case I borrow from Perry concerns beliefs expressible with "I", the objection to [Psychological Criterion] from examples using pure demonstratives does not in any obvious way threaten the use of [Psychological Criterion] in this argument. There is, at least, a narrower principle which restricts [Psychological Criterion] to those indexical sentences containing indexicals like "I" or "here" or "now." And this narrower principle would not be subject to the objection I have noted here.

Granting the narrower principle leaves us no choice but to reject [Belief State Criterion], and this is indeed Perry's solution. What motivated the [Belief State Criterion] was Frege s idea that our internal states somehow reach outside of our minds and home in on the objects that our thoughts are about. The Hume-Heimson case shows this idea to be somewhat suspect. One might merely conclude that the Fregean idea, though essentially correct, stands in need of minor revision and qualification to accommodate the Hume-Heimson sort of case. On this diagnosis of the problem raised by the Belief State Argument, we can respond to it by claiming that internal belief states *partially* determine the truth conditions of beliefs, in that these states determine some

---

[40]See Evans [1980] chs. 6-8 for this view.

sort of function from contexts to propositional contents. A revision to the [Belief State Criterion] in this spirit would allow us to accommodate the problematic features of the Hume-Heimson case. The revised criterion would allow us to say that Hume and Heimson have in common determines a naturally describable function from contexts into contents. The difference in their contexts accounts for the difference in the propositional content of their beliefs.

Robert Stalnaker and others have proposed that this minimal revision will not do. According to these critics, the same Fregean idea motivates both [Belief State Criterion] and the minimal revision to it, and this idea is fundamentally misguided. Thus, the problem raised by the Belief State Argument requires a more radical rejection of [Belief State Criterion].[41] The faulty assumption in both principles is that there is some aspect of belief which is impervious to environmental variance, and which determines, either entirely or in part, the propositional content of the belief. If there were such things as purely internal belief states, then it would be possible to re-identify them in other possible worlds where the social and environmental facts are different. But since this belief state is supposed to be insensitive to all social and environmental facts, we cannot use these facts to re-identify the belief state. How, then, are they to be identified? For familiar reasons, we cannot identify them with physical states of the brain. Identifying belief states as functional states is no help, as a state is a functional state in virtue of its relational, i.e. non-intrinsic, properties. Using intersubjective identity to identify them simply begs the question.

The problem reappears when we consider how an agent's internal state would determine the truth conditions, in that world, of the belief that this state supposedly determines. In possible worlds which bear a very close resemblance to the actual world, it is easy to see what the truth conditions of the re-identified belief would be. And it is exactly these sorts of worlds Hilary Putnam and Tyler Burge use to motivate externalism.[42] But as the world starts looking less like

---

[41]See Stalnaker [1990], [1991] for these views.
[42]See Putnam [1975] and Burge [1979b] for these arguments.

the actual world, it becomes harder to say how we could determine the truth conditions of the re-identified belief. Consider a possible world which differs from the actual world only in that some bodies of watery liquid are made of H_2O and others are made of XYZ. Would the "water" beliefs of the residents of this world be about H_2O, or XYZ, or some combination of both, or perhaps neither? Or consider a world in which the oceans, lakes, and bathtubs are all filled with a substance that plays the same role for its native population as water plays for us, but has the consistency of melted caramel. We can imagine that its residents cannot tell the difference between this substance and real water, even though their discriminative abilities are otherwise very acute. Suppose we can peer into the brains of the residents, and we find that they have beliefs which are "internally identical" to our water beliefs. Are their "water"-beliefs about this caramel-like substance? As Stalnaker has noted, it is hard to see how these questions could be answered.

If the assumption that belief states can be identified independently of their propositional content it is false, then the Belief State Argument does not bring to light a problem uniquely raised by indexical belief. Rather, it brings to light what follows from a mistaken assumption about belief states, which has nothing particular to do with indexical belief, as this notion is characterized at the outset of this chapter. But let us suppose that this assumption is warranted, and that we can respond to the problem posed by the Belief State Argument simply by adopting the minimal revision of [Belief State Criterion] suggested above. On these assumptions, would the Belief State Argument isolate a problem for the traditional picture uniquely raised by indexical belief?

Above, we saw that there is good reason to think that [Psychological Criterion] is not plausible for certain kinds of indexical sentences, in particular sentences containing the demonstratives "this" or "that". For most indexical sentences, clearly there is no interesting sense of "same belief state" in which everyone who assertively utters the same indexical sentence, say, "He is my mail carrier," is in the same belief state. The various individuals who assertively utter this sentence might think of the various referents of "he" in the similar ways, but nothing that

makes it appropriate for them to use this sentence to express this belief requires them to do so. Beliefs appropriately expressed with the demonstratives "this" and "that" have in common a type of mode of presentation of the referent of the demonstrative. But as we saw, there are also differences between the belief states associated with demonstrative thought which matter to the explanation of the agent's actions, but which do not correspond to any difference in the sentence one would use to express the belief. Indeed, it seemed that the criterion was only remotely plausible when restricted to a specific class of indexical sentences, such as those containing "I," `"here," or "now."

This does not suggest, however, that this specific class of indexicals uniquely raises the problem, for there may be other *non*-indexical terms that exhibit these features as well. Two features of indexicals like "I," "here," and "now" give rise to the problem. First, the referent of a use of such an expression is contextually determined. This is a feature they share with many other sorts of terms, such as demonstratives, natural kind terms, and perhaps proper names as well. This feature guarantees that the beliefs expressed with these terms will have different truth conditions, so long as the contexts of utterance are relevantly different. Thus, according to the [Belief State Criterion], agents who have these beliefs must have them in virtue of being in different belief states. Second, these indexicals present their various referents in similar ways. Two people who refer to different places as "here" will nonetheless think of these places in similar ways, for instance as "the place where I am now." These similarities in the way these indexicals present their various referents make the relevant restriction of [Psychological Criterion] plausible. So the problem is uniquely raised by indexical belief, or properly speaking a subclass of indexical belief, only if indexicals are the only sorts of natural language expressions that have these two features.

It seems likely that indexicals are not the only sorts of expressions that have these features. Natural kind terms plausibly have both features as well. As Putnam and Burge have shown,

natural kind terms have an indexical component, in that the references of uses of such terms are partially determined by the social and natural environment of utterances of them. So clearly natural kind terms have the first feature. It is a bit more difficult to say whether they have the second feature. Putnam and Burge both argue from thought experiments in which two phenomenally identical people are shown to have beliefs that differ in truth conditional content. That they are phenomenally identical guarantees that they will be in the same total belief state, but the [Psychological Criterion] is concerned with partial, not total, belief states. The Belief State Argument does not require that Hume and Heimson are in the same total belief state, Hume might well believe that the colonists should revolt, and Heimson could disagree. Rather, the argument requires only that Hume and Heimson are in the same partial belief state, the belief state determined by the indexical sentence "I wrote the *Treatise*."

There are some problematic features of the extension of the Belief State Argument to the case of beliefs expressible with natural kind terms, but nonetheless I think the extension is possible. If I am right, then the Belief State Argument lends no support to the uniqueness claim. To state the extension, and to bring to light what might be problematic about it, we need an example that parallels the Hume-Heimson case. I borrow Perry's form of words to make clear where the analogy lies.

> Suppose Hume is again in his study, only this time he is contemplating a glass of water. In a prescient burst of genius, he thinks to himself, "water is H_2O." Can anyone else apprehend this thought? Some of us can, but not everyone. For instance, Twin-Hume, sitting in his study, is also contemplating a glass of what he calls "water." In what is this time an unfortunate mistake, he thinks to himself, "water is H_2O." But again, however the psychological role of Twin-Hume's belief state resembles the psychological role of Hume's belief state, the fact still remains: Hume is right, but Twin-Hume is wrong. Twin-Hume cannot think the very thought that Hume thought to himself, by using the same sentence.

Does this case present exactly the same problem as the Hume-Heimson case for the joint accommodation of [Belief], [Belief State Criterion] and [Psychological Criterion]? Take p' to be the belief Hume expresses when he says "water is H_2O" and q' to be the belief Twin-Hume

expresses when he utters this same sentence. Propositions p' and q' are distinct, since their truth

conditions differ. Can a "water"-sentence determine a belief state in the same way that an "I"-

sentence purportedly does? The central motivation for [Psychological Criterion] is intuitive; it just

seems that those who accept the same "I"-sentence are psychological similar in an important way.

I think it is plausible to suppose that the intuitions at work here carry over to the case of natural

kind terms. Although Hume and Twin-Hume are thinking about different substances, as Hume is

thinking of H_2O and Twin-Hume is thinking of XYZ, but they are thinking about them in the same

way, as a colorless, odorless liquid that fills the lakes, streams, etc. The word "water" seems to

determine a way of thinking of its referent much as "I" does.

If these intuitions are correct, it would appear that far from presenting the *distinctive*

problem about indexical belief, this argument against the traditional picture points to its quite

general inadequacies. But Perry also offers another kind of reason, and example, to suggest that

the object of an indexical belief cannot be a proposition. These considerations are sufficiently

remote from the argument presented above to suggest that Perry has in mind a different argument

entirely. I consider this argument in the next section.

4.   The second argument is directed against a slightly different version of the traditional

picture, on which the objects to which we stand in the belief relation are Fregean propositions.

This view accepts both [Belief] and [Belief State Criterion], and conjoins these with [Descriptive

Sense] and [Frege's Criterion], both stated below. Frege's own view, as well as broadly Fregean

accounts of the proposition, are paradigm cases of the traditional picture of belief I outlined earlier.

So any investigation of what can be said in support of the uniqueness claim must take Perry's

argument into consideration.

Perry's argument is the following. Perry provides substantial supporting argument for

both premises, but this is its basic form.

(1) The Fregean sense of any singular term is descriptive.

(2) Indexicals do not have descriptive senses.

(3) Hence, the propositions expressed with sentences containing indexicals are not Fregean propositions.

I will address two questions about this argument. First, is it sound? Second, does it lend any support to the uniqueness claim? As my primary concern does not rest in the details of Frege's view, or in the details of what is called the "Fregean view" in so far as this differs from Frege's own view, I simply grant first premise of this argument, that the Fregean sense of any singular term is descriptive, offering only an explanation of what the premise says. My main interest in Perry's argument is the supporting argument for the second premise, that indexicals do not have descriptive senses. I shall argue that although the premise is eminently plausible, and that there are sound arguments that can be brought in its support, these arguments are not to be found in the letter of Perry's arguments. I first present Perry's arguments as I find them in the text, and then offer reconstructions of these arguments which circumvent some problems in their initial formulation. Finally, I argue that even though we do have reason to think that indexicals do not have descriptive senses, these reasons provide no support for the uniqueness claim.

A proposition, according to the Fregean view that Perry targets, is composed of senses, all of which are the senses of the primitive expressions in the sentence that expresses the proposition.[43] According to the semantical characterization of indexical belief I offered at the outset of this chapter, Φindexical belief differs from non-indexical belief in that the former is expressible with sentences containing indexicals and demonstratives, many of which are singular terms. Other sorts of words are common to expressions of both indexical and non-indexical thought. So any explanation of the traditional picture's inability to accommodate indexical belief in

---

[43]What if different sentences express the same proposition? This is a real problem for the view as I state it here. However I will have nothing to say about this particular problem in what follows.

particular must describe the traditional picture's account of singular terms in natural languages, to

make clear why it cannot accommodate indexicals.[44]

As the first premise of Perry's argument indicates, the sense of a singular term is

descriptive on the Fregean view. There is considerable disagreement about precisely what this

comes to, but I will make no attempt to adjudicate between the parties to the disagreement. Rather,

I will simply use Perry's conception, without considering whether it is licensed by anything Frege

says, or whether it is the most plausible account of descriptive sense, Frege's views aside.   A

difficulty remains, however, even in adopting Perry's conception of descriptive sense, as there are

some tensions between various things he has to say about it. For instance, at the outset of his

discussion of the Fregean theory, Perry offers a fairly standard reading of Frege, on which

Fregean propositions are partially individuated by the three criteria below:

> (1) If someone can take different cognitive attitudes towards what a sentence s
> expresses and what a sentence s' expresses, then s and s' express different
> propositions.

> (2) If s expresses a truth and s' expresses a falsehood, then s and s express different
> propositions.

> (3) If "A believes s" is true and "A believes s'" is false, then s and s' express
> different propositions.

But when Perry moves on the argue that propositions expressed with demonstratives and

indexicals cannot be Fregean propositions, he does not target these three criteria explicitly. Rather,

he targets the first premise of the argument sketched above, that all singular terms have descriptive

senses. Perry takes this to mean that the sense of a singular term is either identical to, or

"intimately related to" the sense of a natural language definite description.[45] That this is his

interpretation of Frege is amply illustrated in the text of "Frege on Demonstratives."

---

[44]Since some indexicals are not singular terms, for instance adverbs and suffixes indicating tense, a complete
appraisal of the Fregean treatment of indexical belief will consider its treatment of these sorts of expressions as well.
However, we can show that the Fregean view cannot accommodate indexical belief simply by showing that it cannot
accommodate a certain subset of indexical beliefs, those expressible with sentences containing indexical singular
terms.
[45]Perry [1977] p. 14

In the case of proper names, Frege supposes that different persons attach different senses to the same proper name. To find the sense a person identifies with a given proper name, we presumably look to his beliefs. If he associates the sense of description D with Gustav Lauben, he should believe: Gustav Lauben is D.[46]

[C]an we not say that for each person the sense of the demonstrative "today" for that person on a given day is just the sense of one of the descriptions D (or some combination of all the descriptions) such that on that day he believes: Today is D. [47]

When we look to Perry's arguments that the Fregean theory cannot accommodate demonstratives, we see that the operative assumption is that each use of a singular term will be synonymous with a natural language definite description.[48] So the Fregean view of descriptive sense that Perry attacks is best understood as the conjunction of (4) and (5), below:

(4) An utterance of a singular term A has a descriptive sense if and only if there is a non-indexical definite description whose constant sense is that expressed by A in c.

(5) All utterances of singular terms have descriptive senses.

The Descriptive Sense thesis [Descriptive Sense] is a consequence of (4) and (5) that allows a clear view of the structure of Perry's central argument.

[Descriptive Sense] For any natural language sentence "a is F," where a is a singular term, and any context of utterance c, there is a non-indexical definite description "The G" such that "The G is F" has the same sense as "a is F" has when uttered in c.

---

[46]ibid., p. 15

[47]ibid. In the text between the two passages I have quoted, Perry raises a problem for the claim that demonstratives such as "today" have senses. As their referents shift with context, so must their senses, if the sense of a demonstrative is to determine its referent. I consider this problem below.

[48]In a footnote to these passages added in 1993, Perry claims that the problem demonstratives present for Frege in no way depends on the view that singular terms have descriptive senses as Perry has characterized them. "The significant point," he says, "is not that the demonstrative could not be regarded as an abbreviation for a description. It is rather that the sense of a demonstrative cannot be one that determines its reference independently of context. It does not matter whether these senses are identified by definite descriptions, or expression of some other type, or cannot be identified linguistically at all." (Perry [1993] p. 15 fn. 4.) Perry may be correct about this, but the arguments he presents against a Fregean treatment of demonstratives and indexicals in Perry [1977] and Perry [1979] do not explicitly target the more general view that Perry implicitly attributes to Frege in this footnote, that the sense of any singular term must determine its reference independently of context. I restrict my consideration to the arguments Perry presents, without speculating on how they might be directed against this more general view.

Why should theses (4) and (5), and their consequence [Descriptive Sense], say what it is for an utterance of a singular term to have a descriptive sense? Why not just say what it is for a singular term to have a descriptive sense *simpliciter*? On Perry's version of the Fregean theory, the sense of any singular term determines its reference. Even on the weakest possible understanding of "determination," on which it just means that senses correspond many-one with references, demonstratives and indexicals present a prima facie problem. The references of these expressions shift with context, and since sense determines reference, their senses must shift with context as well. Hence, [Descriptive Sense] accommodates this by allowing that different utterances of the same singular term can have different senses in different contexts.

Allowing that the sense of a singular term can shift with context is not, of course, to lift all constraints on how the various contextually given senses of a singular term-type are related. First, the sense a singular term has in a context is always constrained to pick out the referent of that singular term. Second, anyone who understands a language knows how to determine the referent of an indexical or demonstrative term in any context of its use, even if he or she does not know who or what the referent of the expression is in some context. This suggests that if indexicals and demonstratives have senses, these senses must shift with context according to some rule that competent language users know, either explicitly or tacitly. Perry's argument does not assume the second constraint, in that it aims to establish that no description could ever provide the sense for an indexical, not just that there is no expression-specific rule that determines different descriptive senses in different contexts.

To establish that indexicals do not have descriptive senses conforming to [Descriptive Sense], Perry also uses a psychological criterion of difference for propositions. He supposes that if attributing the belief p to someone would explain behavior that attributing the belief q would not explain, then p and q are distinct. For instance, Perry invites us to consider the case of a messy shopper, whose torn sack of sugar in his grocery cart is leaving a trail on the supermarket floor.

After following the trail of sugar up and down the supermarket aisles, the shopper soon realizes

that *he* is the messy shopper he is trying to catch. After coming to this realization, he reaches into

his cart and adjusts his sugar sack. All along, our shopper believed that the person with the torn

sugar sack was making a mess, but he then came to believe something different, something he

would express by saying "I am making a mess." Why are these different beliefs? Attributing one

to our shopper explains his action of reaching into his cart to adjust his sugar sack, but attributing

the other to him only explains this action on the assumption that he has the further belief he would

express by saying "I am the shopper with the torn sack."

Perry does not mean to suggest that we explain the shopper's action by attributing only one

belief to him, the belief he would express by saying "I am making a mess." This explanation, like

the other, is only successful on the assumption that our shopper has many other beliefs and

desires, for instance the desire that supermarket floors be kept free of sugar trails, the belief that

there is a sack of sugar in his cart, the belief that if he adjusts the sugar sack in a certain way he can

prevent the sugar from spilling out, the belief that sugar tends to fall out of holes in the bottom of

upright bags of sugar, but not out of holes on the top of such bags, and the list goes on and on.

Some of these beliefs and desires, Perry would be quick to note, would find natural expression in

indexical sentences. But just as clearly some of them would not be naturally expressed in these

terms. So we should not take Perry so literally when he says that ascribing an indexical belief to

an agent can explain his or her action, while ascribing a non-indexical belief only explains this

action on the assumption that the agent has some further beliefs and desires. For no action can be

explained by ascribing a single belief or desire to its agent. All such explanations assume that the

agent has many other particular beliefs and desires.[49]

---

[49]Perhaps Perry assumes that belief-desire explanations of behavior require a background of other beliefs and desires
tacitly ascribed to the agent. On this understanding, p and q are different objects of belief if, on the assumption that
the agent has a set of beliefs and desires A, ascribing the belief p explains this agent's action, but ascribing the belief
q does not explain it. If this is what Perry intends, it still does not evade an objection like the one presented in the
text. On this revised view, whether attributions of p and q explain a particular action depends on which beliefs and
desires are contained in the set A.

How then should we understand Perry's psychological individuation criterion for propositions? Perry wants to demonstrate that the belief he expresses in saying "I am making a mess" is not the same as the belief that the shopper with the torn sack is making a mess, or the belief that the only bearded philosopher in the Safeway is making a mess. As Frege would note, someone can believe the indexical belief while disbelieving, or staying neutral about, the latter two non-indexical beliefs.[50] So one way to make the distinctions between propositions of interest to Perry is to use Frege's individuation criterion, [Frege's Criterion].

> [Frege's Criterion] p and q are distinct if and only if it is possible that someone takes different cognitive attitudes towards them.[51]

We are now in a position to state Perry's argument's in support of the second premise of the argument with which I began, that indexicals do not have descriptive senses.[52] Suppose that I develop a desire for a cup of coffee, and that I express this desire at this time in a conventional way, by saying, "I want some coffee." Call the context of my utterance C.[53] If the object of my belief is a proposition, as propositions are characterized on this version of the traditional picture, then "I" as I use it in C must have a descriptive sense. That is, there must be some definite description which refers to me, say "the G," such that I cannot take different cognitive attitudes towards the proposition expressed by "The G wants some coffee" and the proposition I express by

---

[50] Of course, not everyone could be in this belief state. For example, someone who believed that he was the shopper with the torn sack, that is, who believed what he expressed in saying "I am the shopper with the torn sack," could not simultaneously believe what he would express in saying "I am making a mess" and disbelieve that the shopper with the torn sack is making a mess. But what makes it impossible for him to take different cognitive attitudes towards these propositions is a contingent fact about his belief states, that he happens to have the belief he would express by uttering "I am the shopper with the torn sack." Certainly someone who did not have this belief could take different cognitive attitudes towards these propositions.

[51] For instance, if someone believes p and disbelieves, or remains neutral on, q, then that person takes different cognitive attitudes towards p and q.

[52] Perry first considers an argument directed against the claim that an indexical must have the same sense in every context of use, and shows that there is no description or conceptual ingredient that will always pick out the correct referent. He then considers a response to his argument, an indexical need not have the same sense in every context of utterance if it is to have a sense at all. It is his second formulation of the argument, which meets this response, that I consider here.

[53] I understand a context to include a specification of the speaker.

uttering "I want some coffee" in C. Perry argues that there can be no such description "The G," hence "I" in this context does not have a descriptive sense.

What reasons are there for thinking that there could be no such description? Perry notes that for any candidate description "the G," we can imagine my believing the proposition I express in uttering "I want some coffee" without believing that the G wants some coffee. Since I can take different cognitive attitudes towards these propositions, they are, by [Frege's Criterion], distinct. But if they are distinct, then it must be that "the G" is not synonymous with this utterance of "I," as this is the sole difference between the sentences used to express these propositions. Hence "the G" does not provide the descriptive sense for this utterance of "I".

The problem with Perry's argument is that it rests on a scope fallacy. The argument grants that indexicals shift their senses with context, but it does not take seriously what it is to grant this. Granting this requires Perry to allow that any two utterances of the same indexical, even in very similar contexts, could have different senses. Allowing this vastly limits the kind of counterfactual reasoning we can bring to bear in establishing that an indexical sentence s uttered in a context c does not express the same proposition as non-indexical sentence s'. In particular, we cannot establish that these propositions (what s expresses in c and what s' expresses) are distinct by showing that there is some context c' in which s and s' express different propositions. For this inference is only valid if the indexical sentence s expresses the same proposition in c and c', and this is exactly what Perry has called into doubt.

Once he grants that the sense of an indexical shifts with context, e.g. that "I want some coffee" might express one proposition when uttered in c and a different proposition when uttered in c', Perry needs to establish [Descriptions and Context] to support his claim that indexicals do not have descriptive senses.

> [Descriptions and Context] For all definite descriptions "The G" and contexts of
> utterance c, "I want some coffee" uttered in c does not express the same proposition
> as "The G wants some coffee."

What his argument establishes, however, is not [Descriptions and Context] but the weaker

[Descriptions and Context*].

> [Descriptions and Context*] For any definite description "the G," there is some
> context of utterance c such that "I want some coffee" as uttered in c does not express
> the same proposition as "The G wants some coffee."

Although Perry s argument, as it stands, does not yield his desired conclusion that indexicals do

not have descriptive senses, the conclusion is nonetheless eminently plausible. There is nothing,

either in the semantics of indexicals used to express indexical belief, or in the internal states of the

agents whose beliefs are thus expressed, that determines the referents of indexicals used to express

the belief. Both semantics and psychology require facts about the agent's context to determine

which objects his or her indexical beliefs are about. Yet descriptions are supposed to secure their

references without mediation of contextual facts.[54] According to the psychological story of what it

is for agents to think of objects descriptively, such thought requires no contextual mediation either.

The way an agent thinks of an object descriptively is fixed no matter what the agent's context

happens to be. So despite the technical problem in Perry's supporting argument for the second

premise of his central argument, we have reason to accept the conclusion of this central argument,

that Perry's version of the Fregean view cannot accommodate indexical belief.

Since the letter of Perry's supporting argument for the second premise does not establish it,

it does not establish that indexical belief presents any problem at all for Perry's version of the

Fregean view, let alone a unique problem for this view. But even a sound argument for the

second premise would lend no support to the uniqueness claim. Familiar arguments show that

proper names do not have descriptive senses either. Since proper names are singular terms of

natural language, [Descriptive Sense] requires that for any proper name there is some non-

indexical definite description which provides its sense. So, for instance, [Descriptive Sense]

requires that there is some non-indexical description which provides the sense of "Hesperus". But

---

[54]Of course, this too is an idealization. Many uses of a definite description "the G" refer to the contextually salient G.

again, whichever description we pick, say "the second planet from the sun," we can imagine

someone believing that Hesperus is in the east, without believing that the second planet from the

sun is in the east, and this seems to follow for whatever description we might consider. But if so,

then proper names violate [Descriptive Sense] in exactly the same way indexicals do. Thus, it

seems that Perry's argument does not rely on any specific features of indexical identification,

rather it tracks the quite general difference between descriptive and direct reference that Saul

Kripke has noted about names. As such, it is old news. If indexical belief falsifies the Fregean

theory because indexicals do not have descriptive senses, surely belief expressed with proper

names falsifies it for the very same reason. Thus, this argument, like the Belief State Argument,

gives us no reason to think that indexical belief presents a special problem for this second version

of the traditional picture.

5.   A third argument that the traditional picture of belief cannot accomodate indexicals arises

from a committment of the apparently netural principle [Belief], which I restate here.

> [Belief] A's having the belief that s is A's standing in the belief relation to the
> proposition that s.

[Belief] is a schema, of which we form instances by replacing "s" with a sentence of English. If

[Belief] is a fully general principle which applies to any belief whatever that A might have, then for

all of A's beliefs, there ought to be a true sentence of the form "A believes that s," which reports

that A has the belief in question. Thus for any proposition whatever, there ought to be a sentence

that expresses it. I formulate this sententialist view explicitly as [Sent]:

> [Sent] For all propositions p, there is some sentence s such that s expresses p whenever s
> is used.

But this identification strategy fails for indexical belief, since a sentence containing an

indexical will have different propositional contents in different contexts of utterance. The English

sentence "Hesperus is in the east" by itself identifies a proposition, where the sentence "I want

some coffee" does not. The question about which proposition utterances of it express remains

open, since the sentence alone does not identify any particular proposition.[55]

The argument that the object of an indexical belief cannot be a proposition if [Sent] is true

resembles the second argument stated above in broad outline.

(1) [Sent]

(2) There are no sentences that express the objects of indexical belief whenever they are used.

(3) Thus, the objects of indexical belief are not propositions.

The first premise, [Sent], seems to fall directly out of [Belief]. The second premise is further

supported by other supplementary arguments, in which appeals are made to other features of the

traditional picture that have been specified above.

The main argument for the second premise largely parallels the argument I considered in

section 4 that indexicals do not have descriptive senses. Suppose Fred asserts "I want some

coffee" in c. What proposition is the object of his belief? According to [Sent], there must be some

sentence that expresses this proposition whenever it is used. The sentence cannot be the one Fred

uttered, since that expresses different propositions on different occasions of use. For the same

reason, it cannot be any other sentence that contains any other expression containing indexicals in

place of "I", such as "the person here now" or "the only person named 'Fred' in this room." If

there is such a sentence, it must be an eternal sentence, one which picks Fred out non-indexically.

According to [Frege's Criterion], p and q are distinct if and only if it is possible for someone to

take different cognitive attitudes towards them. So if [Frege's Criterion] and [Sent] are both true,

then there must be some eternal sentence such that it is impossible for anyone to take different

cognitive attitudes towards what it expresses and what Fred's utterance of "I want some coffee"

expresses in c. Is there such a sentence? Plausible candidate sentences, if there are any, will be

---

[55]One might wonder whether we really do have a clearer idea of what proposition an eternal sentence expresses, simply because we know that there is some sentence that always expresses it.

identical to the sentence Fred uttered, except in containing a non-indexical expression in place of

"I". For the reasons presented in section 4, the expression cannot be a definite description that

picks out Fred. However, [Sent] does not require that the expression replacing "I" is a definite

description, it requires only that the expression refer to Fred on every occasion of its use, and that

it present Fred in the same way that his utterance of "I" presents him in c. But as we have ruled

out both indexical expressions and definite descriptions, what is left? The only possibility is a

name of Fred. But names are first ambiguous, and second, do not present their referents in the

special way that "I" does. So it appears that there is no sentence that eternally expresses what Fred

said. According to [Sent], what Fred expressed is not a proposition.[56]

One might think that if the special problem about indexical belief only arises on the

assumption that for every proposition there must be some sentence that eternally expresses it, that

the problem could hardly reveal anything interesting about the peculiarities of indexical belief, or

the inadequacies of the traditional picture. Indeed, it reveals little else but the evident fact that

indexicals shift reference with context. However I think it is worthwhile to consider Perry's

proposal here more seriously, as I think that not even the sententialist assumption will deliver a

theory for which indexical belief presents a problem at all, let alone a unique problem.

To see this, we must ask how to understand [Sent]. If it simply requires that there be some

sentence that anyone can use to express the proposition, why can't we just offer some sentence,

and stipulate that it eternally expresses exactly the proposition that expressed with a particular use

of an indexical sentence? Or, why couldn't we simply offer another singular term in place of the

indexical in the sentence, and stipulate that this term has a constant meaning, the same meaning that

---

[56]One might have granted right at the outset that indexical belief does present a problem for [Sent] while noting that the problem is hardly unique to indexical belief. Almost every natural language expression is context-sensitive in some way, so a sentence will not typically express the very same proposition on every occasion of its use. If the proposition is not always expressible with the sentence that happened to express it on some occasion of use, then we have no reason to think that there must be some *other* sentence that will eternally express this proposition, as it holds of all sentences that they typically express different propositions in different contexts. The violation of [Sent] is most apparent in the case of indexical belief, but once we appreciate the context-sensitivity of most natural language expressions we see that most kinds of belief also violate [Sent].

the indexical has in the context? Does either one of these stipulations provide the sentence s that

[Sent] claims to exist? If they do, then the conclusion is blocked. The revised argument does not

establish that there is a problem about indexical belief, *a fortiori* it does not establish that there is a

unique problem. I formulate the claim that blocks the argument from [Sent] as [Stip], below.

> [Stip] For any proposition p, if p is expressible in a language L, then a sentence s
> can be introduced into L, whose meaning is exhausted by the following rule: s
> expresses p on every occasion of its use.

I suggested above an extension of [Stip] for indexical singular terms. Since the indexical is the

only expression in the sentence causing the difficulty with [Sent], one might also hope to block the

argument from either of these principles by stipulatively introducing an expression that has the

same sense as the indexical, as permitted by [Stip*], below.

> [Stip*] For any indexical singular term t of a language L, if t can be used in some
> context c, then another term t' can be introduced into L, whose meaning is
> exhausted by the following rule: t' is synonymous with t-in-c.

When is an expression synonymous with t-in-c? Synonymy is a relation between words, two

words are synonyms if they have the same meaning. Indexical expressions present a difficult case

for the very notion of meaning, hence for synonymy as well, as there is a sense of "meaning" in

which indexicals have a constant meaning throughout their uses, and another sense of "meaning"

on which they do not. To extend the notion of synonymy to indexicals, we need not decide which

of the two senses of "meaning" to use, we can simply introduce two notions of synonymy to

correspond to the two senses of "meaning" for indexicals. For any indexical expression t, what t

expresses in a particular context c differs from what it expresses in suitably different contexts.

When I say that an expression t' is synonymous with t-in-c, I mean that t' means exactly what t

meant in c, and further that t' means what t meant in c *on every occasion of its use*. That is,

whenever t' is used, it refers to the same object that t-in-c referred to, and further, t' presents this

referent in the same way that t-in-c presented it.

If either [Stip] or [Stip*] is true, then indexical belief will turn out to satisfy [Sent]. For [Sent] just requires that for any expressible proposition there is, there must be some sentence that eternally expresses that proposition. Both [Stip] and [Stip*] entail that there will be such sentences for every proposition that can be expressed with an indexical sentence. The central question for the defender of the argument from [Sent] is: is either [Stip] or [Stip*] true? In what follows, I consider reason for rejecting [Stip] and [Stip*], and argue that this reason is equally convincing against the Causal Theory of Naming. If I am right, those who embrace that theory will not want to reject either [Stip] or [Stip*] for this reason, as it provides equally good reason for rejecting the Causal Theory.

What, then, is the objection to these two principles? Typically a stipulation has a limited scope; only those aware of the stipulation can invoke it, and perhaps only in certain situations. Perhaps only those who heard Fred say, "Let's take 'S' to be a sentence that anyone can use to express what I just said," or "Let's take 'A' to be a singular term that means just what the word 'I' meant when I just used it," can use S or A in this way, and even then only in contexts in which the speaker can safely assume that his or her audience is party to the stipulation as well. Someone might utter S or A outside of such a context, but this person could not be said to have expressed what these sentences mean according to the stipulation. But [Sent] requires that there must be a sentence that expresses that proposition on every occasion of its use, not just in the highly restricted set of conversational contexts in which a stipulation can be invoked.

If this objection gives us reason to reject [Stip] or [Stip*], it must also give us reason to reject the Causal Theory of Naming. According to that theory, anyone can use a proper name to refer to its bearer, so long as he or she is appropriately causally connected to the original dubbing of the object with its name. And what is a dubbing if not a performative utterance through which the speaker establishes that some expression X will be a name of a particular object o? Of course, dubbings are not strictly speaking stipulations, since dubbings are embedded in a social practice of

using names to refer to their bearers which extends far beyond the circumstances of the dubbing itself. Someone can use a name competently without being present at the dubbing, without being party to anything like a stipulation, and without having any independent idea of which object the name is a name of. But an expression only means what it has been stipulated to mean in a conversational context in which the speaker and at least some members of the audience are aware of the stipulation (and are aware that they are all aware). Outside of such contexts, the term is meaningless, or else just has its conventional meaning. The central difference between the sorts of stipulation introduced by [Stip] and [Stip*] and the introduction of a name through dubbing is *not* the initial association of an expression with a meaning of some sort. Rather, the difference lies in the way the association of an expression with a meaning arising from a dubbing or a stipulation is incorporated into the language, a difference which clearly rests in contingent facts about our linguistic practice. Once we grant this, nothing remains on which to base an objection to [Stip] or [Stip*] that would not also be an objection to the Causal Theory. If I can stipulate how the sentence s is to be used, by saying "Let s be a sentence that eternally expresses the proposition I express with this utterance of 'I am here now,'" then surely it is also possible that this stipulation could become widely known, so that any competent speaker of the language would be aware of them. Then clearly everyone could use s to express the proposition I expressed while making the stipulation.

The argument from [Sent] at first appeared to be a promising candidate to demonstrate the uniqueness claim about indexical belief. If either [Stip] or [Stip*] is true, then the argument is blocked. It would not establish that there is a problem about indexical belief at all, hence it would not establish the uniqueness claim. If neither [Stip] nor [Stip*] is true, then the conclusion is not blocked, but in this case it is hard to see how one could reject either principle without also either rejecting to the Causal Theory of Naming, or rejecting the claim that sentences containing proper names express propositions.

6.   Those who are convinced that the case against the traditional picture of belief is overdetermined might wonder why we should so much as stop to ask whether indexical belief presents us with *yet another* reason to reject it, or whether the arguments typically offered are simply a specialized application of the more general arguments against this picture. But the vehemence of argument against the traditional picture testifies, I think, to the resilience of the intuitions that motivate it. Doubtless it is the *prima facie* plausible account of belief, and we only reveal what is philosophically problematic about indexical belief by examining the problems it poses for an intuitively attractive account of belief. What such an examination reveals is that it is not entirely clear what the problem about indexical belief is. And a clearer grasp on what this problem is should be of value in assessing the solutions that have been offered to it.

One response to the evident extendibility of the arguments I have considered is to maintain that they *do* isolate unique problems about indexical belief, and that the class of indexical expressions is much wider that it was first perceived to be. The category of indexical expressions is typically demarcated by example; such terms as "I," "now," and "that" are offered as paradigm cases of indexicals. But more recently, we find an "indexical theory" of almost every kind of philosophically problematic natural language expression. Mark Richard identifies propositional attitude verbs as indexicals, Tyler Burge offers us an indexical theory of proper names as well as an indexical theory for the word "true," David Lewis proposes that vague predicates are indexical, and Hilary Putnam, though he stops short of saying that natural kind terms *are* indexicals, asserts that his theory of such terms "can be summarized as saying that words like 'water' have an unnoticed indexical component."[57]

This profusion of indexical theories of various kinds of expressions should at least lead us question the view that indexicals in the traditional and narrow sense constitute a special class. If

---

[57]See Richard [1990] p. 107, Burge [1973], Burge [1979a], Lewis [1970], and Putnam [1975].

most natural language expressions effectively function as indexicals, we should not expect to find

any reason to think that terms like "I," "you," and "here" present any problems about belief that

terms like "water," "bald," and "John Smith" do not present. The failure of the traditional

arguments to isolate any problem might be thought to provide all the more reason to think

indexicality far more pervasive than the traditional and narrow picture of it suggests.

None of this suggests, however, that there are no significant semantic differences between

the traditional indexicals and the other sorts of expressions of which natural languages consist, or

more importantly, that there is nothing of particular epistemological significance about indexical

belief, narrowly construed. In what follows, I sketch some positive reasons for thinking that there

is a special problem about indexical belief.

7.   The second argument of Perry's that I considered[58] aims to establish that indexicals do not

have descriptive senses, hence that beliefs expressible with indexicals do not conform to the

Fregean theory. This argument can be extended to show that beliefs expressible with proper

names do not conform to the descriptive sense theory, either. My point about the limitations of the

descriptive sense theory is a familiar one. The failure of this theory to provide a semantically or

epistemologically adequate account of proper names and natural kind terms as well as indexicals is

widely appreciated. Yet different conclusions are drawn from the failure of this theory for

different kinds of expressions. In the case of proper names and natural kind terms, the conclusion

typically drawn is that, although the bearers of truth value and the objects of intentional states are

propositions, the descriptive sense theory gives the wrong account of what propositions must be.

But in the case of indexicals and demonstratives, a conclusion often drawn is that there is no

account of propositions on which they can be the bearers of truth value and the objects of indexical

---

[58]See sect. 4 of this chapter.

or demonstrative belief. This asymmetry gives rise to a question: what is the special feature of indexicals and demonstratives that warrants the stronger conclusion?

An attractive strategy in answering this question is to first consider why it should be of particular concern to anyone that certain kinds of terms, indexicals included, do not have descriptive senses. Rather than pointing out that indexicals do not have descriptive senses, we might equally well have noted that descriptions do not have indexical or demonstrative senses, unless of course they contain indexicals or demonstratives. It should not be surprising that different sorts of natural language expressions secure their referents in different ways. And it is not clear why all mechanisms of singular reference must conform to the same general constraints if intentional states are to be understood as relations between a person and a proposition. So we should take Perry's point to be not simply that indexicals and demonstratives *differ* from definite descriptions in their semantical and epistemological roles, rather it must be that there is something *special* about the epistemological and semantical role of definite descriptions. The traditional picture of propositions must somehow *require* that all referring expressions secure their referents descriptively.

What presented a particular difficulty for indexicals was the descriptive sense thesis [Descriptive Sense], which I restate here.

> [Descriptive Sense] For any natural language sentence "a is F," where a is a singular term, and any context of utterance c, there is a non-indexical definite description "The G" such that "The G is F" has the same sense as "a is F" has when uttered in c.

Of course, if the traditional picture only makes this requirement on all referring expressions by *fiat*, by adopting something like [Descriptive Sense] as a constitutive thesis, then it is not at all interesting that this picture cannot accommodate expressions whose semantical and epistemological role is unlike that of a description. If, however, something like [Descriptive Sense] was a *consequence* of other requirements that the roles of bearers of truth value and objects of

propositional attitudes make on propositions, then the failure of the objects of certain kinds of belief, including indexical belief, to meet this requirement would point to a substantive problem. So how might these other roles of proposition alone require something like [Descriptive Sense]?

One of the roles of propositions in theories of belief is to be what is communicated. If I understand and believe what you sincerely and accurately assert, then it is natural to say that you and I both believe the very same proposition. Thus, the traditional picture offers us at least a schematic account of what it is for two different people to have the same belief. According to the traditional picture, people who share a belief will stand in the belief relation to the very same proposition. The evident shareability of belief sets certain constraints on what propositions can be. For instance, if two people in different places in a room can have the same belief about a table in the room while having different spatial perspectives on the table, it follows that their perspectives on the table cannot be a part of the proposition they both believe. Thus, a proposition must be divorced, in some way, from the believer's spatial perspective on the objects the proposition is about.

When seen to require this perspective-neutrality, the shareability of propositions sits oddly with the further requirement imposed by theses like [Frege's Criterion] or [Psychological Criterion], which individuate propositions according to their psychological or cognitive roles. A thinker's perspective on an object, for instance whether he or she perceives it from the top, the left or the right, would seem to figure centrally in an account of how that object is presented to the thinker, and what he or she would do to act towards it. And on at least the second version of the traditional picture I considered, the way in which an object is presented to a thinker is itself a part of the proposition the thinker believes about that object. So how can propositions enter in to psychological explanation if they do not include the thinker's perspective on the objects the proposition is about?

[Descriptive Sense] offers a solution to this problem. It accepts the Fregean dictum that objects are not directly given to the mind, but they are always presented to the mind in some way, while denying that the way an object is presented to a thinker depends on the thinker's spatial perspective on the object. According to [Descriptive Sense], we succeed in thinking about an object by thinking of it as the object picked out by a certain description. One can identify an object descriptively regardless of the nature of one's causal interaction with it, and regardless of one's spatial position relative to it. Communication of indexical belief is obviously possible even when the interlocutors have different spatial perspectives on the objects under discussion. If parties to the communication all come to have the same beliefs about these objects, their different spatial perspectives on them obviously cannot be part of the common content of their beliefs. But it is at least possible that they all might think of the object using the same description. So the observation about the role of our spatial perspective on the objects our beliefs are about is particularly appropriate to the case of indexical belief.

What Perry's argument shows is that the apparent fix that [Descriptive Sense] offers doesn't work for indexical belief. Our spatial and temporal perspectives on the objects our indexical beliefs are about are too closely tied to the psychological roles of these beliefs. It is our perspective on these objects that explains our actions directed towards them. The competing commitments of shareability of belief, on the one hand, and explanation of behavior on the other, are most sharply at odds in the case of indexical belief.

Things are different, at least in this respect, for beliefs expressed with proper names or natural kind terms. Although Kripke's arguments make clear, at minimum, that proper names do not have descriptive senses, there is no analogous conflict between the requirements of shareability and psychological role for beliefs expressed with proper names. Suppose you and I are talking on the phone. I tell you that Sammy Sosa leads the Cubs in home runs in the 1998 season. You believe me, so we come to share a belief about Sosa. Nothing about this case requires that these

beliefs have different psychological roles for us. Neither of us need think of Sosa descriptively, but this does not suggest that we could not think of him in exactly the same way.

If the reason for adopting [Descriptive Sense] is simply that it provides some way of squaring the shareability of belief with a psychological role individuation thesis like [Psychological Criterion] or [Frege's Criterion], then indexical belief presents a special problem. For it is not simply that [Descriptive Sense] cannot do this work, nothing could, as the two roles set competing constraints on the objects of indexical belief. While [Descriptive Sense] gives the wrong account of the content of our shared belief about Sammy Sosa, this is not because the two roles of the content of this belief exclude the possibility of any shared content at all. [Descriptive Sense] does not do the required work, but this does not suggest that nothing could.

Do the competing requirements of the psychological role and the shareability of indexical belief constitute a unique problem that indexical belief presents for the traditional picture? There is some reason to think not. Our spatial and temporal perspectives on objects are only one kind of contingency that must be filtered out of the contents of our thoughts if shared belief is to be quite generally possible. Inevitably the particularities of a person's experiences will contribute to the ways in which the objects his or her thoughts are about are presented to him or her, and in general to the psychological roles of his or her beliefs. Yet, people whose particular experiences with certain kinds of objects vary quite widely can still share beliefs about these objects. For instance, a zoologist of the African savanna and an investment banker living in Chicago have quite different particular experiences with lions. The former might see them in the course of her field work every day, while the latter only sees them when his securities firm takes a trip to the zoo. Yet it still seems right to say that both the zoologist and the investment banker can have the same beliefs about lions. For the same reasons that one's ability to grasp a proposition could not depend on one's spatial perspective on the objects it is about, it could not depend on the contingent facts of that person's particular experiences, either.

At the same time that the shareability of belief requires that belief contents must be abstracted away from certain contingencies of the believer's particular experiences and circumstances, we also find that belief contents seem to be *tethered to* certain other contingencies. We saw, in section 3 of this chapter, that there is a reason for thinking that there are no such things as purely internal belief states that are immediately available to the person who has them, and that can be re-identified in different possible worlds. This reason is that there is no saying what contents these belief states would determine in worlds that do not look, to us, exactly or very nearly like the actual world. Cousin-Oscar on Cousin-Earth, for instance, might be in the internal belief state that we are in when we think about water, but there is no saying which of the various water-like substances on Cousin-Earth his beliefs are about. Considering the change in the contents of these beliefs under radical changes in the facts suggests that the belief states themselves, like their propositional contents, are also sensitive to fact. Our natural kind concepts seem to reach out into logical space and support a certain degree of counterfactual speculation, but they don't reach out *that* far.[59]

---

[59]There is some reason for thinking that indexical belief exhibits a much looser conceptual dependence on the facts than our other beliefs, for instance our beliefs about water, do. "I"-beliefs like "I'm cold" present the clearest case for conceptual independence from the situations in which they are entertained. What makes something an instance of the "I'm cold" belief state guarantees a referent for the "I", even if the rest of the content is indeterminate in that world. As far as the indexical component of this belief state is concerned, it has a determinate content in every world in which it exists. "Here"-beliefs and "now"-beliefs display a conceptual independence from the facts to almost the same degree. Even in worlds radically unlike the actual world, "here"-beliefs like "It's cold here" will be about the place the thinker is located, and "now"-beliefs will be about the time at which they are entertained. There are situations in which these belief states do not have determinate contents, but they push on the boundaries of conceptual coherence. Gareth Evans offers an example of someone hurtling through deep space at top speed while thinking to himself "It's cold here," and suggests that this person fails to have a belief with a determinate content. (See Evans [1980] p.206 fn.3) And if the person is thinking slowly enough, and moving quickly enough, Evans' observation about the case seems intuitively correct. But it is difficult to imagine cases in which a thinker is not located in space or in time while entertaining a thought. Other sorts of indexicals, such as "today," or "tomorrow," more clearly depend on the contingencies of the actual world, in this case the contingent fact that there are days, and that there is typically a clear distinction between the start of one day and the beginning of the next. For instance, a "today" belief entertained in deep intergalactic space clearly would not have determinate truth conditions; trans-Pacific flights present a difficult case. There may be reason within these considerations to draw a principled distinction between the beliefs expressible with natural kind terms and indexical beliefs. I cannot claim to have settled the matter here.

8.   It has been my aim in this chapter to identify the unique problem that indexical belief

presents for the traditional picture of belief. If there is such a problem, the three arguments that I

have considered here do not conclusively isolate it. Yet, I do find that there is something quite

right about the uniqueness claim. Above, I called attention to the competing constraints that

shareability and psychological role of belief place on the objects of indexical belief. Of course,

there is no real conflict between these two features of indexical belief. We need not abandon the

truism of belief-desire psychology that beliefs and desires explain behavior to preserve the claim

that objects of belief, including indexical belief, are publicly available. Yet, the "conflict," such as

it is, requires a revision to a simple and attractive explanation of what it is for two people to share a

belief. In particular, it requires us to deny that if two people share a belief, they invariably stand in

the belief relation to the very same proposition. In the case of shared indexical belief, for

instance, A might stand in the belief relation to the proposition p, and B might stand in the belief

relation to a distinct proposition q, yet nonetheless it may be right to say that this is what

constitutes A's and B's sharing a belief.

The phenomenon is perhaps more familiar in the case of tensed thoughts. It is not easy to

say what it is for a person to retain a tensed thought over time. I might look out the window

during a thunderstorm and think to myself "it's really raining now." What would it be for me to

retain that very thought over time? In so doing, I would necessarily think of the time that my belief

is about in different ways. As time passes, I can no longer think of this time as "now," rather I

must think of it as "then" or "a while ago." If these ways of thinking are a part of the proposition I

believe, then it appears that retaining my belief requires that I stand in the belief relation to different

propositions at different times. Yet it seems also that there are some constraints on how I can think

of this time if I am to be said to have retained this belief, rather than coming to have a different

belief with the same truth-conditional content.[60] We find a similar problem about the retention of

---

[60]David Kaplan raises the problem in "Demonstratives, "Suppose that yesterday you said, and believed it, "It is a
nice day today." What does it mean to say, today, that you have retained *that* belief? It seems unsatisfactory to just

demonstrative belief, i.e. the sort of belief expressible with "this" or "that". One's perspective on the objects in one's immediate environment changes as one moves around within it, and these differences in perspective are properly understood as *intentional* differences which are responsible for differences in behavior. Thus, these differences in perspective ought to correspond to differences in the proposition which is the object of the demonstrative belief. And as above in the case of tensed thought, it is clear that retaining a demonstrative belief involves thinking of the object demonstrated in different ways when one is in different places with respect to it. And similarly, it is not clear how retention of such a belief differs from the acquisition of a new belief with the same truth conditional content.

The analogous problem arises in the interpersonal case for "you" and "I". Suppose you say to me, "I'm from Chicago." I believe what you say, so I come to have a belief I would express by saying "You're from Chicago". In this case, unlike the two cases considered above, we are not concerned with what it would be for you or me to retain these beliefs over time. In those cases, we were concerned mainly with what it is for a person to retain an indexical belief under relevant changes in the context. Which changes are relevant depends on which features of the context the indexicals are sensitive to. So for instance, temporal indexicals like "now" are sensitive to the time of the context, so a relevantly different context is one in which the time of the context is different. Here, the indexicals are sensitive to the agent of the context, so relevantly different contexts are ones whose agents differ. Asking what it is for you and me to share the belief you expressed by saying "I'm from Chicago," is thus relevantly like asking what it is for me, today, to retain the belief I expressed yesterday by saying "It's really raining now."

No doubt, there are important disanalogies between the three sorts of cases I have mentioned here, which could provide reason for treating them asymmetrically. However, it seems

---

believe the same content under any old character--where is the *retention*? You *can't* believe that content under the same character" (Kaplan [1989] p. 537). He also notes the interpersonal analog of this problem for the indexicals "you" and "I".

not unreasonable at the outset to suppose that we ought to treat the cases symmetrically unless we have a clear reason not to. It seems right to say that retaining a tensed belief from t to t+ is not the same as abandoning the tensed belief believed at t and coming to have a new belief at t+ which is truth conditionally equivalent to the abandoned belief. If we want to say this, we ought also to say that A and B might share the belief that A expresses by saying "I am F", and that this is not the same as A and B having truth conditionally equivalent beliefs that A is F.

9.   Is this problem really uniquely raised by indexical belief? I should first note that there is a relatively trivial sense in which the problem I raise here only arises for indexical belief. Contexts shift constantly for terms like "I," "you," and "now,", but not at all for "covertly" indexical terms such as "water". So it is easy to imagine a conversation in which the context shifts for the purposes of determining indexical reference, but there is no natural situation in which the context that determines the reference for terms like "water" will shift in the course of a conversation, or will differ for different participants in it. Unless there is communication between people whose "water" thoughts are about different substances, the special problem I characterized above simply does not arise.

But suppose we were to consider a cooked up case. Ocsar, a Twin-Earth native, has just been transported to Earth, and he knows this, and further he knows what the difference is between Earth and Twin-Earth, namely that the substance on Earth that Earth-natives call "water" is chemically quite different from the substance that he calls "water" back on Twin-Earth, even though it looks, tastes, and feels exactly the same. Suppose Fred, and earth-native, and Oscar are standing next to a pond, and Fred says that the water in the pond is warm. What is it for Oscar believe what Fred tells him?

Oscar knows that Fred is talking about the stuff in the pond, that he's saying that *that stuff* is warm. But *that stuff* is not just presented directly to Oscar's understanding. He thinks of it in

some way. How does he think of it? Oscar might try to understand what Fred said by thinking of

the stuff in the pond just as he thinks of what he calls "water" on Twin-Earth. But since he knows

the difference between Earth and Twin-Earth, he knows that *that stuff* is not what he would

typically call "water," even though all appearances suggest that it is. The stuff in the pond meets

the descriptive concept of what he calls "water," but it seems to be an important part of this concept

that it is a concept *of* a certain substance, the substance found in the ponds, lakes, and bathtubs of

Twin-Earth. Oscar might be tempted to think of it as "water," and if he did not know that he had

left Twin-Earth he no doubt would think of it this way, albeit mistakenly. But, knowing what he

knows, Oscar would have to think to himself, "It looks like water, but it just *isn't* water."

Oscar might, instead, try to believe what Fred has told him without thinking of the stuff in

the pond as what he would call "water." That is, he might think that Fred has just said, about

some substance whose nature is completely foreign to him, that some of it is warm. Certainly this

is possible, since one need not have a descriptively rich concept of an object to have beliefs about

it. The fact that Oscar can see the substance Fred is talking about is enough to allow him to have

beliefs about it. But Oscar's belief, about the stuff in the pond, that it is warm, would not be a

way of really believing what Fred has told him, either. For it is a part of believing what Fred says

that one thinks of the substance in the pond in a way that broadly resembles the way Fred himself

thinks of it.

3

A Problem About Indexical Belief

1.   Suppose that Victor and Victim are walking together in the forest when a bear appears around a bend. Both men are experienced mountain hikers, so both know that anyone in danger of being attacked by a bear should curl into a ball immediately. Suddenly the bear lunges in Victim's direction and Victim immediately curls into a ball. Victor, knowing that bears almost never move on to a new victim until they are finished with the first, runs to the nearby ranger station for help.

Given certain plausible assumptions, the situation described above, which I call "Bear", raises a problem for a very natural account of beliefs and desires and their relationship to actions. Victor and Victim behave differently, so there ought to be a difference in their beliefs or their desires which explains the difference in their behavior. Yet once we allow that beliefs and desires are shareable, it seems that even the finest individuation of the contents of belief and desire will not yield a difference in either their beliefs or their desires that explains the difference in their behavior. The upshot of this problem is a dilemma about indexical beliefs and desires. Either we deny that these beliefs and desires play the central role in the explanation of behavior that they are typically thought to play, or we complicate the traditional account of the expression and communication of indexical belief and desire to accommodate their evident public availability.

Some may recognize the problem about Bear as the familiar problem of the "essential indexical" as John Perry has famously called it.[61] A puzzling aspect of this familiar problem is the particular difficulty involved in its precise formulation. Arguments that attempt to bring the problem

---

[61]See Perry [1979].

to light are typically arguments that some theory or other is ill-suited to accommodate indexical belief. These presentations can obscure the fact that the problem lies not so much with these various theories as it does with our conflicted intuitions about indexical belief. In suggesting that there is some problematic class of beliefs, indexical beliefs, that the traditional theory cannot accommodate, these presentations immediately invite us to ask which kinds of beliefs are members of this class. But a well-motivated answer to this question requires the background that only a clear grasp on the problem can provide. In relying on intuitions about a particular case rather than a general theory of belief, I hope to redress these two infelicities of presentation in what follows.

I first state three plausible assumptions. Two concern Bear, and the third concerns the relationship between belief, desire, and action. I then show how they generate a problem about belief, and consider the possible solutions to it. I argue that we cannot plausibly deny two of the three assumptions, and focus my attention on the third assumption, that Victor and Victim share all of the relevant beliefs and desires in Bear. Then, I consider a solution to the problem about Bear due to John Perry. His solution seems to me to be almost right, but I argue that it cannot accomodate a special way in which Victor and Victim share their beliefs and desires in Bear. Finally, I suggest that Frege's response to the problem about Bear points the way towards an emendation of Perry's account that will enable it to suitably characterize the special way in which Victor and Victim share their beliefs and desires in Bear.

2.   The first of the three assumptions concerns a general relation between our beliefs and desires, on the one hand, and what we do, on the other. I first state a principle relating these intentional states to actions; but certain problems with this initial formulation require a revision of the principle. Thus the principle I shall assume, [Beliefs and Attempts], replaces the notion of an action in the first principle with that of an attempt.

That actions are to be explained by citing the beliefs and desires of the agent is practically a truism. Actions differ from reflexes or purely physical processes in that they are intentional, and this is just to say that they are susceptible to intentional explanation. I offer [Beliefs and Actions] as an initial formulation of this claim.

> [Beliefs and Actions] If A and B act differently, then there must be a difference in either their beliefs or their desires that explains the difference in what they do.

Several difficulties for [Beliefs and Actions] undermine its intuitive appeal. First, one might, like Buridan's ass, have several courses of action available without having any reason to prefer one over the others. For instance, Fred might want one of the beers in the refrigerator without having a reason to select the one on the left over the one on the right. Fred chooses the beer on the left, but apparently without having any beliefs or desires that explain why he picked that one over the other. One might think that such cases are counterexamples to [Beliefs and Actions], however I am inclined to retain the generalization about intentionality found there. Actions, after all, are intentional, and how could they be intentional if they are not subject to belief-desire explanations? And how could there be a difference in what is to be explained without a difference in what does the explaining? But since my view on this point is perhaps contentious, I shall note that the problem about Bear in no way depends on my claim that such cases are compatible with [Beliefs and Actions]. In Bear, Victim has a reason for acting as he does which is not also a reason for acting as Victor does.

The second difficulty is more serious. What we do is determined in part by what the world is really like, not just by what we believe or desire it to be. Suppose that Uta Pippig and I both run the Boston Marathon. While she finishes the race in record time, I pass out near Wellesley Hills. Clearly Pippig and I have done different things, as she won the race, and I did not even finish it. But the difference in what we did is not intentional, there need be no difference in our beliefs and desires that explains why she won and I did not. Rather, the difference is a difference in the world,

specifically in our physical capabilities. But according to [Beliefs and Actions], what we do is

completely determined by our beliefs and desires. So this case is a counterexample to [Beliefs and

Actions].[62]

> For this reason, I qualify [Beliefs and Actions] to [Beliefs and Attempts].

> [Beliefs and Attempts] If A and B attempt to do different things, then there is a
> difference either between A's and B's beliefs or between A's and B's desires that
> explains the difference in what they attempt to do.

> How can [Beliefs and Attempts] circumvent these problems? Suppose that Pippig and I

both start the Boston Marathon with the same beliefs and desires about it. Although we did

different things, in that she won the race and I did not, we both attempted to win the race. So the

case that presented a problem for [Beliefs and Actions] is perfectly compatible with [Beliefs and

Attempts].[63]

> The second and third theses I urge as intuitively plausible specifically concern our example,

Bear. They are [Different Attempts Thesis] and [Common Attitude Thesis].

> [Different Attempts Thesis] In Bear, Victor and Victim attempt to do different
> things.

> [Common Attitude Thesis] In Bear, Victor and Victim share all of the beliefs and
> desires that are relevant to explaining what they do.

> Concerning [Different Attempts Thesis], I shall note only what it does not say. It does not

say that there is nothing that Victor and Victim both attempt to do. Clearly there are many things

---

[62]This difficulty for [Beliefs and Actions] is the flip side of a familiar point about the semantics of attitude verbs and action verbs. Verbs like "to believe" are intensional. Even if Fred's yelling "Boo!" at the woman in the navy suit just is Fred's offending the Secretary of State, Fred might believe that he has yelled "Boo!" at the woman in the navy suit without believing that he has offended the Secretary of State. But certainly if the actions are one and the same, then Fred cannot help but do the one if he does the other. (Here, It is important that we understand the expressions "Fred's yelling 'Boo!' at the woman in the navy suit" and "Fred's offending the Secretary of State" are terms referring to particular actions, not action-types.) If the objects of the attitudes, presumably propositions, are individuated differently from the objects of action verbs, then it should not be surprising that differences in the latter are not always reflected in differences in the former.

[63]Some will object that even the move to [Beliefs and Attempts] does not adequately address this problem. Just as certain conditions in A's environment must obtain if we are to call what A does "doing x", it might be that certain conditions in A's environment must obtain if we are to call what A does "an *attempt* to do x". I defer further discussion of this difficulty to section 4.

that both attempt to do in Bear. For instance, both attempt to save Victim from death by bear attack, and [Different Attempts Thesis] is perfectly compatible with this evident fact. This thesis claims only that there is something that Victor attempts to do that Victim does not attempt to do. And since Victor attempts to run for help and Victim attempts to curl into a ball, there is a clear difference in what they attempt to do.

[Common Attitude Thesis] seems to me both intuitively compelling and theoretically well-motivated. I shall advance only some of the intuitive considerations here; I defer discussion of other arguments in support of it for later discussion. First, we can suppose that Victor and Victim both have full knowledge of the relevant facts about their situation in Bear, that is, they both have full knowledge of the facts that are relevant to explaining why it is rational for them to act as they do. For instance, they both know that they are walking in a forest, that the ranger station is behind them to the left, and that curling into a ball is the best course of action for someone in immediate danger of death by bear attack. After the bear appeared on the scene, both came to know that a bear was nearby, then shortly afterwards both came to know that the bear was attacking Victim.[64]

It is not clear what must be the case for one's belief that p to count as knowledge that p. But the path of analysis from knowledge to belief is comparatively straightforward; if one knows that p, then one believes that p. So if Victim and Victor both have full knowledge of the relevant facts about their situation, that is, the facts that are relevant to an explanation of why it is rational for them to act as they do, they must also share all relevant beliefs about their situation as well. Of course, it is important here that both have *full* knowledge of the relevant facts, otherwise we could not infer the claim that they share all relevant belief from the fact of their shared knowledge.

---

[64]Of course, these are not all of the facts that both Victim and Victor know, not even all of the facts that they both know that explain why they act as they do in this situation. It is impossible, and fortunately unnecessary, to list all of these facts. There may be some interesting philosophical question about exactly which of the facts that both Victor and Victim know are relevant to an explanation of what they do, as typically there are such questions lurking behind the use of the word "relevant." But even if it is not clear exactly which facts are the relevant facts and which are the irrelevant facts, or how we could in general distinguish those facts in the former category from those in the latter, we can suppose that whichever facts these are, Victor and Victim both have full knowledge of them.

We can also suppose that Victor and Victim are both rational, in that they act in such a way as to lead to the satisfaction of their desires, were the beliefs on which they act true beliefs. To say that Victor and Victim are both rational in this sense is to say nothing about the content of their beliefs or desires. This procedural conception of rationality is only concerned with what courses of action are rational for a person to take, given what that person believes and desires; it says nothing about whether desires or beliefs, *per se*, are rational or irrational. Yet we can also suppose that Victim and Victor are rational in the quite different sense that there are certain outcomes that they both see to be obviously bad, obviously to be avoided if possible. For instance, they both take the outcome of Victim's death by bear attack to fall into this category. Perhaps it is more pressing for Victim that this outcome be avoided than it is for Victor, but both have a strong desire that it not be the case that Victim dies from this bear attack. It is difficult to say exactly which of Victor's and Victim's desires are operative in an explanation of what they both do in this case. But if we suppose that a desire is a desire *that* something or other is the case, we can then assert that Victor and Victim share all of the desires that are relevant to explaining what they do.

Even if Victor and Victim agree about which outcomes are to be pursued, and which outcomes are to be avoided, it is likely that they will disagree about the relative importance of these outcomes. As I mentioned, Victor's death by bear attack is the outcome that is most likely at the top of Victor's list of outcomes to be avoided, while Victim's death by bear attack is likely to be at the top of Victim's list. This appears to be a difference in their desires. Why then do I claim that Victor and Victim share all of their relevant desires? A difference in the relative importance each one places on various outcomes does not explain why Victor ran for help and Victim curled into a ball, and this is evident once we reflect on whether Victor and Victim would have had reason to act differently if their desires had been ordered differently. If Victor's desire that Victim not die from this bear attack had been just as strong as Victim's, would this have given Victor reason to curl into a ball, as Victim did, rather than run for help? Clearly not.

3.   It should be clear how the three theses present a problem. According to [Different Attempts

Thesis], there is a difference in what Victor and Victim attempt to do. Thus, by [Beliefs and

Attempts], there is a difference either in their beliefs or their desires that explains this difference.

Yet according to [Common Attitude Thesis], there is no such difference in their beliefs or desires.

So the three intuitively attractive theses cannot all be true. Our solution to the problem, then, must

be one of the following three:

> A. Deny [Different Attempts Thesis]
>
> B. Deny [Beliefs and Attempts]
>
> C. Deny [Common Attitude Thesis]

I consider the merits of the first two possible solutions in section four, and I conclude that neither

one offers us a plausible way out of the problem. The issues involved in the third possible solution

are more complicated, I discuss these in sections five through seven.[65]

4.   I take the first response to be an obvious non-starter. Here is Victim, on the ground,

curling into a tight ball, and there is Victor, sprinting from the scene. How can we deny that Victor

and Victim are thereby attempting to do different things without stepping far outside the boundary

of the pretheoretical notion of an attempt that is at work in [Beliefs and Attempts]?[66] Denying

[Beliefs and Attempts] is also the wrong way out, but in this case it is not obvious that this is so.

---

[65]An attractive response to the problem will offer some explanation for the initial plausibility of the denied thesis, so it may take the form of a charge of equivocation on some crucial term in the denied thesis. Such a solution, of course, is a way of adopting one of A-C above, but it will be rhetorically useful to consider this sort of solution separately.

[66]We should remind ourselves that [Different Attempts Thesis] is quite compatible with what is evident, that there are some things that Victor and Victim both attempt to do. I shall note also that I have made no attempt to provide an account of what sentences like "A attempts to do y" say. I do not want to *deny* that there are thorny and difficult issues involved in making the logical forms of these sorts of claims precise, I want merely to *sidestep* these issues. My point is that we need not have any particular account of actions and attempts to conclude Victor and Victim attempt to do different things in Bear.

Below I consider two kinds of cases which present a problem for it, and then suggest rejecting [Beliefs and Attempts] on these grounds sheds no light on the source of our problem.

The first difficulty for [Beliefs and Attempts] concerns the kind of problem I initially raised for [Beliefs and Actions], above. Suppose that in world w and w', Fred believes that Gingrich is on the cell phone. Desiring to damage Gingrich's career, he attempts to tape his conversations. In world w, Fred's belief is true, but in w' it is false, since Lott, not Gingrich, is on the cell phone. Some would say that in w', Fred is attempting to tape Lott's conversations, although Fred himself would not put it that way. But even if Fred is attempting to tape Lott's conversations in w', he surely is not attempting to do so in w. Thus, there is a difference between what Fred attempts to do in w and w', but no obvious difference in Fred's beliefs or desires that explains this difference.[67]

I think it is clear what the defender of [Beliefs and Attempts] should say about this sort of case. If A is attempting to do x, then A is in a certain intentional state. And we are quite familiar with the idea that our intentional states may be individuated by the objects that these states are about. For instance, we are quite ready to say that believing of a that it is F is not the same as believing of b that it is F, if a and b are distinct. And we are quite ready to say this even if the believer does not know that a and b are distinct, and could not even tell the difference between a and b. The example above suggests that just as beliefs and desires can be *de re*, attempts can as well. Unless there is a principled reason to think that beliefs and desires are more internal, and less dependent on contingent facts about the agent's environment, than attempts, what motivates us to say that Fred is attempting to tape Lott in w' should also motivate us to say that even though Fred cannot tell the difference between w and w', nonetheless he has different beliefs and desires in these two worlds. In w his beliefs and desires that explain what he attempts to do are *de re* beliefs about Gingrich, but in w' he has different *de re* beliefs, since these beliefs are about Lott. And it is the difference in *de re* beliefs and desires that explains the difference between Fred's attempts in w and w'.

---

[67]I thank Mark Richard for bringing this sort of case to my attention.

The second sort of case presents a more serious difficulty for [Beliefs and Attempts], in that we cannot use this same maneuver to evade the problem. Consider (Paralysis):

> (Paralysis): A sees B across the street and desiresΦto attract her attention. He lifts his hand to wave to her. C has exactly the same beliefs and desires as A, and attempts to raise his arm to wave, butΦunbeknownst to him, his arms are completely paralyzed.

In (Paralysis), there is not even an outside chance that C will succeed in raising his arm, even though he has not the slightest idea that this is so. But if this is the case, can anything C does be correctly called "C's attempt to raise his arm"? If we can only correctly say that C is attempting to do x if there is at least an outside possibility of C'sΦdoing x in the circumstances, then we must admit that [Beliefs and Attempts] is false, as clearly C and A can have the same beliefs and desires, evenΦif they cannot attempt to do the same things. As I would like to maintainΦ[Beliefs and Attempts], I claim that even if someone has no chance ofΦsucceeding at her chosen task, she can still *attempt* toΦperform it, and we make no conceptual error in claiming that she can.

One might want to avoid problems raised by these sorts of cases by avoiding talk of actions or attempts altogether. The second case in particular seems to suggest that whether we call what someone does an attempt to do x depends on the cooperation of the physical world. "But," one might say, "if the problem the two theses raise is supposed to be a problem about *belief,* then the gap between the purely psychological core of belief and its outward manifestation in action is certainly irrelevant. The agent's psychological state alone is what matters, Victor's and Victim's actions only matter to the problem in so far as they indicate their psychological states. The problem in Bear exists before anyone starts running for help or curling into a ball, and this is because the problem arises from the difference between the psychological states of the agents involved."

One might, then, seek to replace the notion of an attempt in [Beliefs and Attempts] with a more strictly psychological notion. However I reject this internalizing move, as I reject its assumption that psychological states can be individuated without reference to the objects in the

world that these states concern, and to the thinker's interactions with these objects. But the choice

to eschew the notion of a psychological state in no way leads to the problem about Bear. The

problem remains for those who are completely comfortable with the notion of an internal

psychological state. The facts are that it is rational for Victor to act in one way, and rational for

Victim to act in another way, given what they know about their situations. Retreating into the

purely psychological core of belief, if such there is, will not change this fact.

But let us leave aside Bear for the moment, and just consider [Beliefs and Attempts]. Can

we make sense of a denial of it? Let us suppose that [Beliefs and Attempts] is false, so that at least

in some cases, two people, equally rational (in the procedural sense), might make different attempts

even though they have exactly the same beliefs and desires. Recall that to explain a difference

between what two people attempt to do, we must cite a difference between at least one of the sorts

of things that explain what people attempt to do. Can we explain what people attempt to do by

citing something other than beliefs and desires? If beliefs and desires were just one sort of thing

among many that explain what people attempt to do, then perhaps we could insist that a difference

in attempts without a difference in beliefs or desires could still have some explanation. But it seems

that nothing else can do the explanatory work. So the denial of [Beliefs and Attempts] just comes to

the claim that there are things people attempt to do which have no intentional explanation at all. It

is, however, paradigmatic of attempts that they are susceptible to intentional explanation. This

distinguishes an attempt from a reflex motion or a sneeze. [Beliefs and Attempts] is criterial of

attempts, and for this reason denying [Beliefs and Attempts] is a particularly unattractive response

to the problem.

5.   In section two, I presented an intuitive appeal for [Common Attitude Thesis]. But intuitions

about belief and desire are notoriously malleable. Perhaps there are some further intuitive

considerations that count against [Common Attitude Thesis]. I now turn to a more complete

appraisal of this matter.

As I noted earlier, Victor and Victim share their beliefs about their situation, in the sense that

they agree about the facts of their situation. But does such agreement constitute shared belief? This

is not obvious. Suppose, for example, that A believes that Hesperus is the brightest star in the night

sky, and B believes that Phosphorus is. Since Hesperus is Phosphorus, A and B would seem to

agree about which planet is the brightest in the night sky. Of course, A uses the name "Hesperus"

to refer to this object, and B uses "Phosphorus," but this is hardly a disagreement about the facts.

If neither A nor B knows that Hesperus is Phosphorus, they might *think* that they disagree about

which planet is the brightest. When they learn that Hesperus is Phosphorus, they will see that their

disagreement was illusory, and they could say that they really agreed all along. Even if they never

learn that Hesperus is Phosphorus, and take themselves to disagree about which star is the

brightest, a third party, describing the situation, could still correctly say that they do not really

disagree. But despite their agreement, there does seem to be a sense of "shared belief" in which A

and B do not share their beliefs about which planet is the brightest in the night sky. In this case, the

temptation to say that A and B do not share their beliefs about which planet is the brightest has to do

with the fact that they use different names to refer to this planet, but this is not the only sort of

cognitive difference that tempts us to distinguish between beliefs. Perhaps A and B identify the

planet Venus in different ways. A identifies it by its covering of gases seen through her high-

powered telescope, and B identifies it in quite a different way, through its position relative to other

planets and stars from her vantage point. In this case, we might equally well be tempted to say that

A and B do not share their beliefs about Venus, despite their agreement about the facts.

Frege offers us a way of individuating propositions so that people who agree about the

facts need not do so in virtue of believing the very same proposition. For Frege, propositions have

as constituents modes of presentation of the objects that these propositions are about. So if A and B

are presented with the planet Venus in different ways, then their beliefs about it will differ in their

propositional content, even if their believing these different propositions constitutes their agreement

about the facts.

One might think, then, that a Fregean view will offer the required difference in Victor's

and Victim's beliefs or desires. I shall argue, however, that this impression is mistaken. No matter

how we understand modes of presentation, they will not provide the required difference between

Victor and Victim's beliefs or desires in the clear and simple way that the Fregean theorist might

have hoped.[68]

My argument, which I call "the argument from belief shareability," does not proceed from

difficulties for any particular account of Fregean modes of presentation, I shall not linger presenting

a detailed account of what they are. A few words should be sufficiently suggestive of the sort of

view I have in mind, which many have thought adequate to provide the required difference in

beliefs or desires. Although Victor and Victim need not think of the objects their beliefs are about in

Bear using different uniquely identifying descriptions, they must perceive these objects in different

ways, just given where they are in relation to them. For instance, they are presented with the bear

in different ways, in that their perceptual experiences of it are quite different. Victim identifies it,

among other things, as a large furry mass looming over him, and Victor identifies it, among other

things, as a large furry mass looming over someone else. And obviously, each one thinks of, and

perceives, the other quite differently from the way he thinks or and perceives *himself*. When Victim

thinks to himself "I'm being attacked by a bear," he thinks of himself in a certain self-conscious

---

[68]I should note that my argument is not intended to point out an obvious difference between Victor's and Victim's beliefs that the Fregean view fails to capture. Rather, I consider the possible application of the Fregean view to this problem precisely because it is explicitly sensitive to the differences in the ways in which believers think of the objects their beliefs are about, and takes these differences to yield differences in the propositions believed. If anything, the Fregean approach individuates propositions too finely to permit us to say that two people share their beliefs when it seems intuitively obvious that they do. Many have criticized the Fregean approach on precisely these grounds. That the difference between Victor's and Victim's beliefs in Bear resists characterization by even the finely individuated products of the Fregean account of propositions suggests a significant cognitive similarity between their beliefs.

way, as *himself*. His own body is presented as the object whose movements he has a special sort

of direct control over, his mind is the one whose thoughts and sensations he has non-inferential

access to. Victor only has access to Victim's thoughts through what Victim says, although it might

be reasonable for Victor to suppose that Victim has certain beliefs given the situation he knows

Victim to be in. Victim's body is presented to Victor as a physical object over which he, Victor, has

no direct control.[69] Thus we might understand modes of presentation as heavily perceptual rather

than descriptive. I do not mean to suggest that I think it is clear what it would be to so-understand

them. In fact I think it is far from clear what this would be. However, given the inadequacy of the

descriptive account of modes of presentation in accounting for indexical belief, if modes of

presentation are to be a part of the contents of indexical belief at all, then they must be heavily

perceptual rather than descriptive.

   The argument from belief shareability begins with two observations about the distinction

between phenomenal states, like having a headache, and intentional states, like believing that the

earth is round. First, phenomenal states are essentially *someone's* phenomenal states, there is no

such thing as a headache that is not someone's headache. And although different people can have

qualitatively similar headaches, they cannot both have the very same headache. Phenomenal states

must belong to exactly one sentient being, they cannot be shared. Unlike phenomenal states,

intentional states are shareable. Two people can have the very same belief. Second, phenomenal

states are commonly said to be ineffable. I take this to mean that there is nothing that A can say to B

to make B have any reasonably accurate beliefs what it is like to be in a particular phenomenal state,

unless B has already been in a qualitatively similar phenomenal state. Intentional states, on the

---

[69]These descriptions of the differences between the first- and third-person modes of presentation could no doubt be refined. I am only concerned that there is some difference between the way each person is presented to himself and the way he is presented to others. Sydney Shoemaker has argued that we err even in saying that we are presented to ourselves in any way at all, as he thinks that the analogy with perception that the term "mode of presentation" invokes offers an inappropriate characterization of the way we get information about ourselves. If Shoemaker is correct, then the Fregean view as I have characterized it does not even get off the ground. See Shoemaker [1968] for these arguments.

other hand, are completely expressible in language. So as long as a thinker knows what she

believes, she can in principle explain to someone else exactly what it is. That is, so long as she is

not subject to constraints of time, or a particular inability to articulate her thought, such as the

inability to find the word she needs to express herself, and so long as her audience is sufficiently

patient, charitable, understands the language she speaks and so on, she can express her beliefs to

the audience, in a way that makes it possible for them to understand exactly what it is that she

believes.[70]

On the Fregean view outlined above, modes of presentation, whatever we take them to be,

are themselves parts of propositions. So on this view it would be natural to say that two people

have the very same belief if and only if they both stand in the belief relation to the same proposition.

Thus, these intuitions about intentional states support [Belief Shareability Thesis].

> [Belief Shareability Thesis] If A is sincere, and tries to get B to believe something,
> and is successful both in completely articulating what he wants B to believe and in
> getting B to adopt it, then A and B both believe the proposition that A tried to get B
> to believe.

The antecedent of [Belief Shareability Thesis] gives a neutral description of successful

communication between a speaker and audience, which results in the audience's coming to believe

what the speaker asserted. There are, of course, a wide range of defeating conditions that might

undermine successful communication. I listed some of these defeating conditions above, but surely

there are others I did not specify. I understand the antecedent of [Belief Shareability Thesis] to

imply that none of these defeating conditions holds. The consequent of [Belief Shareability Thesis]

gives a more theoretical description of this exchange, in using the notion of a proposition. But the

constraints it sets on how successful communication that results in shared belief should be

represented in a theory or belief are minimal.

---

[70]I invoke the contrast between phenomenal and intentional states merely to highlight certain features of the latter.
My main point is not that phenomenal states are private and ineffable, rather it is that intentional states are *not*.

[Belief Shareability Thesis] is, I think, extremely plausible. It is no doubt the natural account of what it is for one person to understand and believe what another person asserts. If the audience and speaker do not come to believe the very same proposition, one is inclined to think that this must be the result of some mistake. Either the speaker did not say what he or she meant, or the audience misunderstood. Although [Belief Shareability Thesis] should appeal to Fregean theorists, and indeed Frege does endorse it, it should also be attractive to those who take propositions to be the objects of belief, but who reject the Fregean account of what propositions are. I have specifically considered the Fregean account of propositions, as it is often thought that this view takes seriously a thinker's cognitive access to the objects his or her beliefs are about, thus allowing fine discriminations between beliefs that other accounts of the proposition leave no room for.[71] But anyone who accepts [Belief Shareability Thesis] faces the problem I will raise shortly.

The argument requires the additional assumption that for any belief that is relevant to explaining the difference between what Victor and Victim attempt to do in Bear, they both can communicate this belief to the other in Bear. Of course, not every belief is expressible in every possible circumstance. For instance, a demonstrative thought cannot be expressed to someone who does not perceive the object demonstrated. For similar reasons, people alive in 1790 could not have grasped the proposition that Bill Clinton is a Democrat. But the argument does not require the stronger, and obviously false, claim that propositions are always and everywhere expressible.

Why should we accept the additional assumption about Bear? The same considerations about expressibility of belief that motivate [Belief Shareability Thesis] also motivate the additional assumption that all beliefs are expressible, at least in some circumstance or other. What gives us reason to think that a belief is, as we might say, *in principle* inexpressible, also gives us reason to

---

[71]Examples of such other accounts of the proposition are what is often called the Russellian view, on which propositions are structured entities composed of the referents of the expressions in the sentences that canonically express them, and the possible worlds account, on which a proposition is a set of possible worlds, intuitively the set in which the proposition is true. Both accounts are thought to individuate propositions too coarsely to respect certain intuitively natural judgements about the difference between beliefs.

think that the state in question is not a belief. So there is at least some circumstance in which the relevant beliefs in Bear are expressible. Is Bear itself such a circumstance? In Bear, we find that none of the defeating conditions that make communication impossible in particular cases holds. Victor and Victim speak the same language. There is nothing terribly subtle or difficult to express about the beliefs that explain why it is rational for them to act as they do. There is no particularly difficult concept to be communicated, rather the relevant beliefs would seem to be those concerning their spatial relations to the physical objects in their immediate environment. It is not Victor's failure to understand a concept, or an expression, that explains why he runs for help rather than curling into a ball. Victor is not making a mistake in running for help, rather he is doing what is rational for him to do given full knowledge of the relevant facts, which he has. Also, Victor is ideally placed to know all of the information about the context thatΦcould possibly be relevant to interpreting what Victim says. He is, after all, standing right next to Victim, as he has been for some time. So it seems clear that for any beliefs that might be relevant to explaining the difference in what Victor and Victim attempt to do, a circumstance in which these beliefs are successfully communicated would be very much like Bear.

So it is possible for Victor and Victim to communicate to each other the beliefs relevant to an explanation of what they attempt to do in Bear, and such communication would result in shared belief. Thus, for any candidate belief that purportedly explains the difference in what Victor and Victim attempt to do, there will be some possible circumstance, very like Bear, in which this belief is successfully communicated, and as a result Victor and Victim come to share that belief. But to show that [Common Attitude Thesis] is true, or at least that the part that deals with beliefs is true, it is not enough to show that for any belief that plausibly explains the difference in what Victor and Victim attempt to do in Bear, there is some *other* situation that resembles Bear in which they share this belief. Rather, we must show that they share this belief *in Bear*.

How then do we reach the stronger conclusion? The communication between Victor and Victim that occurs in these other possible situations would not change their epistemic situations one iota from their original epistemic situations in Bear. Any beliefs that are communicated in these other possible situations are beliefs that they already have in Bear. Of course, there are things that Victor and Victim might say to one another that would result in a change in their beliefs, none of these beliefs could be relevant to an explanation of the difference in what they attempt to do.

I suggested above that the sort of belief that is likely to explain the difference in what Victor and Victim attempt to do is a belief expressible with an indexical or a demonstrative. Let us focus on one such belief, the belief Victim would express to Victor if he were to say, "I'm being attacked by a bear." I take it as obvious that whether or not Victim in fact utters this sentence, he has this belief. If Victim were to express this belief to Victor by saying "I'm being attacked by a bear," Victor would no doubt understand and believe what he says. And what is more, Victor would not thereby come to realize that Victim had this belief, surely Victor knew this from the very moment the bear lunged at Victim. Let us say that this belief, the belief that Victim would express by saying "I'm being attacked by a bear," is *communicatively transparent* to Victor in Bear. That is, in Bear, Victor already believes what Victim would express in uttering this sentence. The same holds for any of Victim's beliefs that is plausibly relevant to the explanation of the difference in what Victor and Victim attempt to do in Bear is also communicatively transparent to Victor. Similarly, all of Victor's beliefs that could plausibly do the analogous explanatory work are communicatively transparent to Victim. Let us call this claim [Communicative Transparency Thesis].

> [Communicative Transparency Thesis] All of Victim's beliefs that are relevant to explaining the difference between what Victor and Victim attempt to do in Bear are communicatively transparent to Victor in Bear, and similarly, all of Victor's beliefs that are relevant to explaining the difference between what Victor and Victim attempt to do in Bear are communicatively transparent to Victim in Bear.

If [Communicative Transparency Thesis] is correct, Victor and Victim already believe everything they would believe at the end of an exhaustive discussion of all of their relevant beliefs about the

situation. But we already know that none of these beliefs explains the difference in what they attempt to do. Thus it appears that there is no communicable belief that could explain the difference, either.

It is worth noting that [Communicative Transparency Thesis] does not follow from the claim that initially motivated [Common Attitude Thesis], that Victor and Victim agree about all of the relevant facts in Bear. Suppose that A believes that Hesperus is visible and night, and B believes that Phosphorus is visible at night, and that B does not know that Hesperus is Phosphorus. It seems correct, as I suggested above, to say that A and B agree about this fact. If A were to express his belief about Hesperus to B, and if B were to believe what A says, then B would learn something new. Certainly B comes to have a new belief. What B learns depends on how B incorporates what A says into her system of belief. Suppose B comes to believe that A is talking about the very planet she knows as Phosphorus. If she had never before heard the name "Hesperus," she would acquire a metalinguistic belief, that "Hesperus" is another name for the planet she had been calling "Phosphorus." If she is in the situation of the ancient Babylonian, thinking that Hesperus and Phosphorus are distinct, then the new information that they are identical will effect a more radical change in her beliefs. If she had conflicting beliefs about Hesperus and Phosphorus, she would now be in a position to realize, merely by reflecting on her beliefs, that they could not both be true. Where she might have had two different beliefs, that Hesperus is visible at night at Phosphorus is, too, she is now in a position to realize that these beliefs are one, again merely by reflecting on her beliefs.

The preceding arguments only concern Victor's and Victim's beliefs. A difference in Victor's and Victim's desires could equally well explain the difference in what they attempt to do. But the argument from belief shareability can be extended to show that the difference cannot be a difference in desires, either. The argument of this section was motivated by quite general considerations about intentional states and their objects, propositions. It is widely thought, and for

good reason, that the objects of intentional states must all be of the same sort. How else would it be possible for Fred to desire exactly what Ed fears, or for Jones to believe the very thing Smith doubts? I noted that what distinguishes intentional states from phenomenal states is their shareability and complete expressibility in language, so surely this applies to the objects of desire as well as belief, as evidently they are one and the same.

Of course, the process by which one comes to believe p is quite different from the process by which one comes to desire p. But at issue in the argument from belief shareability is not the process by which one comes to believe a proposition, but rather the nature of our access to the proposition believed. Thus, since the argument appeals only to features of belief and desire that these two propositional attitudes have in common, viz., the nature of propositions, then surely if it gives us reason to think that Victor and Victim share all of their relevant beliefs in Bear, then it gives us equally good reason to think they share all of the relevant desires as well.

Let us summarize our progress. Our intuitions about Bear suggest that for any of Victor's beliefs that is relevant to the explanation of the difference between what Victor and Victim attempt to do, there is a possible circumstance in which Victor communicates this belief to his audience, and his audience comes to believe what he says, and similarly for all of Victim's relevant beliefs. Examination of the case suggests that these possible situations would be very like Bear, since Victor and Victim are ideally situated to interpret and understand each other's utterances. Thus, we have, by [Belief Shareability Thesis], that in those possible situations, like Bear except that the relevant communicative exchange takes place, Victor and Victim both stand in the belief relation to the same proposition, the one that the speaker (whichever one he was) expressed in that situation. But then it is clear that Victor and Victim learn nothing new about the situation, or about each other's beliefs, that could be relevant to explaining the difference between what they attempt to do in Bear. Victor and Victim are in the very same epistemic situation before and after the communication takes place. This is the observation motivating [Communicative Transparency Thesis], and from this we infer

that, in Bear, Victor and Victim share all of the beliefs that are relevant to explaining what they attempt to do. A parallel argument concerning desires rather than beliefs then entails [Common Attitude Thesis].[72]

7.   If the argument from belief shareability is valid, we face yet another difficulty in finding a solution to our original problem about Bear. The original problem arises from three intuitively plausible theses, [Different Attempts Thesis], [Beliefs and Attempts], and [Common Attitude Thesis], not all of which can be true. I suggested that rejecting either of the first two is highly unpalatable and brings us no closer to an explanation of our competing intuitions on these cases. The only option remaining, then, is to reject [Common Attitude Thesis]. Yet, not only does this thesis have the force of intuition on its side, there is also, as we have just seen, an argument for it from [Belief Shareability Thesis], and several other reasonable assumptions. If [Common Attitude Thesis] is false, this argument must be shown to be unsound.

I think it is clear that the argument from belief shareability is valid, so I take it that our only option in response to it is to reject one of its premises. The argument is quite complex, but assumes very little in the way of theory. For the most part, it arises from a simple observation about Bear. Victor and Victim both agree about the facts of their situation that are relevant to an explanation of what they attempt to do. Further, there is nothing that either one can sincerely say to the other that will change their epistemic situations in the slightest, at least nothing that will change their epistemic situation with regard to how they should act in Bear. So although it might be possible, in certain

---

[72]If the argument from belief shareability is sound, then no version of the Fregean view as I have characterized it above will provide the difference in Victor's and Victim's beliefs that [Beliefs and Attempts] requires, no matter what we take modes of presentation to be. As I understand the Fregean view, it includes [Belief Shareability Thesis]. Many of those who endorse a Fregean view of indexical belief are quite aware of the difficulty that the possibility of communication of indexical belief presents for the view Frege explicitly adopts, and they circumvent these difficulties by rejecting [Belief Shareability Thesis]. I should note that my objections to what I have called the "Fregean view" do not apply to these views, in fact I think such a view of indexical thought is quite likely to be right. See Evans [1981], Evans [1980], Bezuidenhout [1997] and McDowell [1984] for Fregean approaches to indexical belief that deny [Belief Shareability Thesis].

circumstances, to convey a Fregean mode of presentation of an object by means of pointing or describing, and thereby put one's audience in a position to acquire a new belief about that object, it is clear that Bear is not such a circumstance. These observations about Bear cannot plausibly be denied. Surely there is no conceptual incoherence in the supposition that a circumstance might have all of the features that I have claimed for Bear.

Thus the only place to object to the argument is its sole theoretical commitment, contained in [Belief Shareability Thesis]. According to this thesis, when one person understands and believes what another person has asserted, the two people stand in the belief relation to the very same proposition, the one the speaker asserted. Obviously, this is hardly a descriptive stipulation about Bear, rather it is a claim about what is the best theoretical representation of what the parties to the conversation have in common.

And this makes our last possible solution to the problem about Bear altogether less attractive. Intuitions about belief and desire are often conflicted, and much as there are intuitions that support [Common Attitude Thesis], there are equally compelling intuitions against it. The intuition that supports [Beliefs and Attempts] is a good example. What motivates that principle is what might be called the "internalist" conception of belief and desire, on which our beliefs and desires are immediately available to us, so they are sensitive to all and only the distinctions that the thinker is sensitive to. Attention to this internalist conception would suggest that [Common Attitude Thesis] must be false, since given the truth of [Beliefs and Attempts] there simply *must* be a difference between Victor's and Victim's beliefs or desires. On the other hand, attention to the more impersonal and objective conception of belief and desire pushes us in a different direction. On the impersonal and objective conception, the objects of our beliefs and desires are publicly available entities, bearers of truth value which must be sensitive to how things really are, not how a thinker might perceive them to be.

There is clearly something right about both of these conceptions of belief and desire. A plausible account of belief and desire should accommodate both of them. Of course, accommodating both of them requires us to understand them so that both can be true, and our problem is that it is not so clear how this is possible. But what should be clear is that the choice of three possible solutions to the problem about Bear was somehow false. Rejecting [Beliefs and Attempts] seems to require us to reject the internalist conception of belief and desire, and rejecting [Common Attitude Thesis] seems to require us to reject the objective conception. Of course, we cannot plausibly reject either conception.

In surveying the three possible solutions, I have suggested that [Beliefs and Attempts] is somehow immovable, that we cannot reject it without thereby rejecting what is obvious, that actions are not mere reflexive or random motions, rather they are *intentional*, and *chosen*. I then concluded, by what might have looked like a process of elimination, that we must reject [Common Attitude Thesis]. Thus, one might object that my presentation of the problem has been biased in favor of the internalist conception of belief and desire. I might, of course, have taken a quite different tack in presenting the problem, so as to make Bear look like a counterexample to [Beliefs and Attempts]. Since it is clear that Victor and Victim attempt to do different things in Bear, [Beliefs and Attempts] entails that there is a difference between either their beliefs or their desires. Thus, one might suppose that it ought to be possible to identify this difference. Such identifications canonically take the form of belief or desire reports. So if there is a difference between Victor's and Victim's beliefs or desires, there ought also to be a true sentence of the form "Victor believes (desires) that ..., and Victim does not believe (desire) that ..." where what goes in the "..." is the same in both places. But what the preceding discussion makes clear is that there is no true sentence of that form.[73] If there were such a sentence, Victor, for instance, could utter the "...", and Victim could then come to believe or desire that ... as well. Any sentence that could go in the "..." falls under the scope of

---

[73]Strictly, there is no such true sentence that reports a *relevant* difference in beliefs or desires.

[Belief Shareability Thesis]. One might insist in this case that the requirement that questions like "What is the difference between A's and B's beliefs?" must have answers of the form "A believes that ... and B does not believe that ..." But one might equally well insist that if there is no true sentence of this form, then A and B must have all of their beliefs in common.

I hope that my presentation of the problem suggests that the malleability of the concepts that [Common Attitude Thesis] and [Beliefs and Attempts] employ make the proper solution to the problem about Bear to some degree a matter of stipulation. If we understand "belief" and "desire" in one way, the proper response to the problem is to reject [Beliefs and Attempts], but if we understand these notions in a different but equally reasonable way, we ought to retain [Beliefs and Attempts] and reject [Common Attitude Thesis]. Either way the problem of accommodating the objective and the internalist conceptions into our solution remains.

Thus the decision to reject [Common Attitude Thesis] rather than [Beliefs and Attempts] is attractive in that it gives us greater leeway to accommodate these two conceptions of belief and desire. If we deny [Belief and Attempts], we are left with nothing intentional to explain the difference between what Victor and Victim attempt to do. But we might respect the intuition behind [Common Attitude Thesis] without claiming that Victor and Victim have identical beliefs and desires. We might claim instead that their beliefs and desires are similar in some particular way, and that this similarity is enough to accommodate our intuition that Victim and Victor share all of the relevant beliefs and desires. Thus an attractive solution to the problem about Bear could resemble a charge of equivocation on "belief" and "desire" between [Common Attitude Thesis] and [Beliefs and Attempts].

In what follows, I will consider two solutions to the problem along these lines, one due to Perry, and the other due to Frege. Perry's solution is attractively simple, and I think almost entirely correct, yet it is ill-equipped to handle the secondary problem presented by the argument from belief shareability. Perry denies [Belief Shareability Thesis], as I suggest is the only plausible response to

this argument, but in doing so he allows too much to count as shared belief, thereby leaving himself

no room to account for the somewhat special way in which Victor and Victim share their beliefs in

Bear. I then turn to Frege's discussion of the problem. Frege's solution to the problem about Bear

is hardly palatable. However, his remarks point to an emendation of Perry's solution that could

provide a response to the secondary problem that I think causes trouble for it.

8.   I characterized Perry's response as a charge of equivocation on "belief" and "desire"

between [Common Attitude Thesis] and [Beliefs and Attempts]. On one understanding of the

charge of equivocation, this characterization is unfair. Perry does not claim that there is one sort of

belief and desire that explains behavior, and another sort which is communicated from one person

to another, which might be shared between Victor and Victim in Bear. According to the person

who makes this kind of charge of equivocation, it would be a simple confusion to ask how the two

senses of "belief" or of "desire" are related, akin to asking how the two senses of "bank" are

related. The problem in Bear clearly is not driven by this sort of confusion, and Perry does not take

it to be so-driven.

         But there are other ways of making the charge of equivocation. For instance, Aristotle

suggests that we say different things in characterizing different sorts of things as "medical," yet

there may still be one meaning which is primary, and which explains the meanings of the other uses

of "medical."[74] Thus, we can fruitfully ask how the various senses of "medical" are related. We

might understand Perry as saying something broadly similar about "belief" and "desire." He thinks

that we might use the phrase "A has the same belief as B" to call attention to different similarities

between A's and B's beliefs. Our intuitions about sameness and difference of belief answer to

---

[74]In the Metaphysics Z, Aristotle notes, "we apply the word 'medical' when there is a reference to one and the same thing, not meaning of and the same thing, nor yet speaking homonymously; for a patient and an operation and an instrument are called medical neither homonymously not in virtue of one thing, but with reference to one thing."(1030a33-1030b3) The interpretation of this sort of passage, found commonly in the Metaphysics as well as the Nicomachean and Eudemian Ethics, that I offer in the text is due to GEL Owen [1957]. I am grateful to Mitzi Lee for her assistance on these points.

many different concerns. If we are concerned about what is preserved in communication, then we individuate beliefs in one way, but if we are concerned to explain what people do, then we individuate them in a different way. Cases like Bear demonstrate that the various ways of individuating beliefs for different purposes do not neatly converge. If we are to explain all of the intuitively plausible claims about sameness and difference of beliefs, Perry thinks we must introduce objects of belief with greater representational complexity, which will allow many different kinds of comparisons between beliefs.

Perry proposes that we replace the notion of a proposition with a pair of objects, a belief content, and a belief *state*. Belief contents have a sentence-like structure, and are individuated by something like their truth conditions. They contain none of the cognitive overlays that Frege builds into propositions. Someone has a belief with a particular content in virtue of being in a certain belief state under certain conditions. A belief state, then, is a mental property of a believer. These states play two roles in Perry's account. Together with facts about the believer's context, they determine the content of the belief. Belief states also play an epistemological role, in that they are intended to be individuated finely enough to explain the agent's behavior.[75] That is, if there is a difference in what two agents attempt to do, there must also be a difference in either their belief states or their desire states.

How does Perry address the problem in Bear? Even though it is clear what Perry's position is, it is not clear which of the three possible solutions we should best understand him as taking. If we say that two people share a belief if and only if their beliefs have the same content, then we would characterize Perry's response as a denial of [Beliefs and Attempts]. If we say that two people share a belief if and only if they are in the same belief state, and this state uses

---

[75]In both of their roles, their function parallels that of David Kaplan's "characters," which are the functions from contexts to contents which provide the semantic value for indexical and demonstrative pronouns. See Kaplan [1989]

contextual factors to determine the same content, then we would characterize Perry's view as a denial of [Common Attitude Thesis].[76]

The inadequacy of Perry's account concerns its inability to account for the exact sense in which Victor and Victim share their beliefs in Bear. Perry would explain their agreement about the facts of their situation by citing the fact that on his view, their beliefs about their situation have the same content. But presumably almost any two people who agree about the facts also will have beliefs with the same content. Consider a third person, Alice, who reads about the events in Bear in the newspaper a few days after the event. Alice believes that the account in the newspaper is true, and comes to have beliefs about the situation with the very same content as Victor's and Victim's beliefs about it. Yet there is a sense of "shared belief" in which Victor and Victim share their beliefs about their situation in Bear, a sense in which they do not share their beliefs with Alice. If we think that the difference here is one an account of belief should accommodate, then we should find fault with Perry's solution, as it leaves no resources for locating it. Victor, Victim, and Alice all have a belief with the same content, but they all have this belief by being in different belief states. Perry's account cannot, as it stands, explain the intuition that Victor and Victim share their beliefs with each other in a way that they do not share their beliefs with Alice.

9.   Frege dedicates a few brief passages in the late essay "Thoughts" to a discussion of the particular problem indexicals present for his account of propositions. What he says there is far from a complete theory of indexical thought, and I do not pretend to do complete exegetical justice to what he says in these passages. He first considers the case of tensed thought.

> If someone wants to say the same today as he expressed yesterday using the word "today," he must replace this word with "yesterday." Although the thought is the same its verbal expression must be difference so that the sense, which would otherwise be affected by the differing times of utterance, is readjusted.[77]

---

[76]I give a more complete characterization of Perry's view in Chapter 2, secs. 2-3 of this dissertation.
[77]Frege [1956] p. 358

Here, Frege says that the propositions expressed with the word "today" can still be expressed on the next day, so long as the word "yesterday" is used in place of "today" in the original sentence. It is not clear how to extrapolate from this example. Would he also think that the very same proposition could be expressed a week later if a tensed expression like "a week ago today" was used in place of "today"? Is there a point at which the day is no longer "indexically accessible" in Perry's phrase, at which that proposition is no longer expressible? Perhaps Frege would reject the claim that there is such a point, and we should read him as tacitly endorsing a direct reference theory of indexicals, on which all an indexical contributes to a proposition expressed by a sentence containing it is the object to which it refers. This would be one explanation of Frege's remarks about "today" and "yesterday," but it is certainly not the only one the text allows. What we *can* glean from Frege's remark is that the proposition expressed by a sentence need not change simply because the words used to express it change. The words by themselves do not force a change in the proposition expressed.

What presents a difficulty for any interpretation of Frege's remarks about "today" and "yesterday" is that he explicitly denies what would be the analogous view about the pronoun "I". The troubling passage reads:

> Now everyone is presented to himself in a special and primitive way, in which he is presented to no one else. So, when Dr. Lauben has the thought that he was wounded, he will probably be basing it on this primitive way in which he is presented to himself. And only Dr. Lauben himself can grasp thoughts specified in this way.[78]

This passage points to the special way in which each person is presented to him- or herself, and Frege claims that this special and primitive way in which a person thinks of himself is something on which a proposition can be "based". Frege gives an example of a proposition that is "based on" this special and primitive way of presentation, which is the proposition that Dr. Lauben thinks to

---

[78]ibid. p. 359

himself when thinking that he was wounded. Frege's language here suggests that this special and primitive way of being presented to oneself is what Frege alternatively calls a "sense" or a "mode of presentation." On Frege's view, a sense, in addition to being part of a proposition, is also a sense *of* an expression. The obvious candidate expression for this special and primitive sense to be a sense of is "I". But this straightforward understanding of Frege contradicts what he says in the very next sentence.

> But now he [Dr. Lauben] may want to communicate with others. He cannot communicate a thought he alone can grasp. Therefore, if he now says "I was wounded", he must use "I" in a sense which can be grasped by others, perhaps in the sense of "he who is speaking to you at this moment"; by doing this he makes the conditions accompanying his utterance serve towards the expression of a thought.[79]

If we are to take this seriously, the special and primitive way of thinking of oneself does not, after all, provide the sense of "I". While others cannot grasp this special way of thinking of oneself, others can grasp the thoughts expressible with "I". Therefore the special mode of presentation cannot be the sense of "I". The difficulty, of course, is that it seems natural to say the thought that Dr. Lauben expresses when he says "I was wounded" *just is* the thought he thinks to himself, the one Frege takes to be based on the special mode of presentation. We are not told what thought Dr. Lauben's audience would express by saying "you were wounded," so it is difficult to assess the potential analogy with the case of "today" and "yesterday." To parallel the view about these tensed indexicals, Frege would have to maintain here that "I was wounded" from Dr. Lauben's mouth expresses the same thought as "you were wounded" expresses from his audience's mouth. Frege's only indication about the sense of the expression "you" occurs in his suggestion that it might be a part of the "public" sense of "I". Yet it is unclear why the same gap between the private and the public senses that Frege outlines for thoughts about oneself is not also present for the thoughts about others expressed with "you". For it seems that if the audience is going to think of Dr. Lauben

---

[79]ibid. p. 359-60

in the right way, they must think of themselves in a certain way, as "you" is a part of the suggested sense of "I". But if this way of thinking of themselves is the special and primitive way in which each person is presented to himself and only himself, then Dr. Lauben will not be able to grasp this sense, hence he will be unable to use it as part of the sense of "I," in *his* mouth. For the very same reasons that Dr. Lauben's audience will not be able to grasp the special and primitive way in which Dr. Lauben thinks of himself, Dr. Lauben will be unable to grasp the special and primitive way in which members of his audience think of themselves.

To extract a response to the problem in Bear from these few remarks, we need to have some idea of what role the special and primitive way of thinking of oneself plays in the explanation of what one does. Is it because Victim thinks of himself in this special and primitive way that he curls into a ball when thinking that he is under bear attack? If so, then according to Frege this is the belief that Victim has, and Victor lacks, that explains the difference in what they attempt to do. Victor believes that Victim is being attacked by a bear, and he believes Victim when he says "I'm being attacked by a bear," but in saying this Victim does not actually express the belief that explains what he attempts to do. Frege's response to the problem in Bear, then, is to deny [Common Attitude Thesis]. That is, he would deny that, in Bear, Victor and Victim share all of the beliefs and desires relevant to explaining what they do. Is his response to the problem at all plausible?

In section seven I presented several reasons for thinking that rejecting [Common Attitude Thesis] is very unattractive, even more so than is commonly supposed. All of these quite general reasons count against Frege's view. But there are also specific problems with Frege's view. The letter of his view does not require him to deny [Belief Shareability Thesis], the auxiliary principle which presented additional support for [Common Attitude Thesis]. That principle just requires that if communication between two parties is successful, and the second party believes what the first said, then both parties share the belief that the first party expressed. But Frege appears to suggest that what someone expresses when he uses "I" is something that his audience is fully able to

understand. What Frege maintains, rather, is that there are some beliefs, namely the beliefs about ourselves that we think to ourselves, that we simply cannot express to others.

But Frege retains [Belief Shareability Thesis] at a cost. For what gives us reason to accept [Belief Shareability Thesis] is the thought that beliefs, unlike phenomenal states, are completely expressible in language, and it is this that Frege denies. But the expressibility of thought in language is so central to our understanding of what beliefs, and intentional states in general, are, that we are left without an understanding of how this inexpressible belief of which Frege speaks could be a belief at all. Is it some private analog of the belief Victim expresses to Victor, in that both concern the same things, both have the same truth conditions, but the belief that explains Victim's attempt is just flatly and inexplicably inexpressible? Do all beliefs have a similarly inexpressible analog, or is it just beliefs based on the special and primitive way of thinking of oneself? In so far as it is possible to understand Frege's response to the problem, it invites more difficulties than the problem it was introduced to solve.

Also, Frege posits two propositions, one private and the other public, when it seems that one proposition should do both jobs. For when I understand what someone says about himself using "I", I am fully aware of the sort of thought this person has, and of the way in which this person thinks of himself in virtue of having this thought. And I know, furthermore, that the way this person thinks of himself is not, as Frege suggests, as the person speaking to me at this moment. Frege posits the public proposition to preserve the phenomenon of successful communication using "I", but the public proposition he offers simply cannot be the one that is communicated.

If take Frege to claim that our indexical beliefs about ourselves are simply incommunicable, then his view is puzzling in another way, since it posits an unexplained asymmetry between tensed thoughts, such as those expressed with "today" and "yesterday," and what we might call "egocentric" thoughts, such as those expressed with indexicals referring to persons, such as "I,"

"you," "he," and "she." The passage quoted above suggests that Frege thinks that we can still

entertain the thoughts we expressed yesterday using "today," as long as we replace the word

"today" with "yesterday" in the expression of the thought. But unless we take the proposal of the

public and private propositions associated with indexical belief about oneself seriously, we must

take Frege to deny that the same holds for beliefs expressed with terms like "you" and "I". But the

affinities between the temporal, spatial, and the "egocentric" modalities demand a more symmetrical

treatment. I do not think of Thursday, January 29 in exactly the same way during that day as I do

on the next day or the day after. Yet Frege claims that the thoughts I express on that day using

"today" can still be entertained after that day has passed. Perhaps there is good reason to treat the

two cases differently, but it is not clear what Frege would take this reason to be.

10.   What Frege has to say about the senses of "I" and "you" have the flavor of offhand

suggestions in response to a problem quite remote from his central concerns. Yet his remarks hint

at a solution to the puzzle which respects the special sense in which Victor and Victim share their

beliefs in Bear. I shall attempt to articulate this solution more fully here.

The argument from belief shareability presented in section seven gives us a theoretically-

motivated reason for thinking that [Common Attitude Thesis] is true. This argument shows that if

we wish to retain the natural account of communication found in [Belief Shareability Thesis], we

must understand the epistemic contributions that indexicals make to the propositions expressed by

sentences containing them as not completely articulable in language. But it seems evident that

Victor and Victim do understand those aspects of each other's cognitive states that explain why they

act as they do in Bear. It is just that this understanding does not require them to act in the very same

way. The argument requires us to say that either our interlocutors come to grasp the very same

proposition, or their communication is unsuccessful. When faced with this choice, it seems clear

that theoretical simplicity must give way to the truism about intentionality. We must therefore reject the natural account of communication.

When Sammy Sosa says to me, "I'm at bat," I have more than an indeterminate idea of what he is saying, I grasp his meaning exactly. If I understand what he says, I will of course not think of him in exactly the way he thinks of himself. I might, in some confused way, attempt to do so, as an insane person who believed himself to be Sammy Sosa might do. But clearly this is not the sort of thing I would even attempt to do if I understand what someone says to me in uttering "I'm at bat." It is not so much that I cannot think of Sammy Sosa in the right way, hence cannot grasp exactly the thought he thinks to himself, so I must settle for something less. Rather, understanding him *requires* me to think of him differently from the way he thinks of himself. At least in some cases, successful communication which results in one person believing what the other said requires that the two parties to the conversation grasp different propositions.

It is this observation that motivates Perry's solution, but as we saw, his solution allows too much to count as successful communication. The correct account of communication of indexical belief will set tighter constraints on the ways in which the parties to the communication can grasp the content of the proposition expressed. It is to this end that Frege's comments in "The Thought" suggest a promising strategy. Frege suggests that the "public" sense of "I" is "the person addressing you at this moment," which *contains* the sense of "you". For reasons presented above, this proposal appears to require what Frege acknowledges to be impossible. For if the speaker is to associate the proper sense with "you," the speaker would have to be presented with his or her audience in the special and primitive way in which it is acknowledged that the audience cannot be presented with the speaker. No one can think of others in what Frege calls the "special and primitive way" in which she thinks of herself. But one can understand what it is like for someone else to think of herself in this way, on *an analogy* with what it is like to think of oneself in this way. One understand what someone has said in uttering, for instance "I'm at bat," by taking it that the

speaker is thinking of himself in the same way that you would think of yourself, were you to utter the same sentence.

Of course, I have not presented even the briefest sketch of how the analogical understanding of the egocentric indexicals would work. Clearly the viability of the proposal depends on a more detailed account that I have offered here. It seems, however, that while it allows that successful communication of egocentric indexical belief does not require both parties to the conversation to grasp the very same proposition, it also has the potential to explain the close affinity between Victor's and Victim's beliefs in Bear.

References

Almog, Joseph, John Perry and Howard Wettstein, [1989] *Themes From Kaplan*, New York: Oxford UP

Aristotle, *Metaphysics*

Anderson, C. Anthony and Joseph Owens, [1990] *Propositional Attitudes*, CSLI Lecture notes; no. 20, Stanford University

Bezuidenhout, Anne L. [1997] "The Communication of *De Re* Thoughts" *Noûs* 31:2 197-225

Burge, Tyler [1973] "Reference and Proper Names" *Journal of Philosophy*, 70: 425-39

---, [1979a] "Semantical Paradox" *Journal of Philosophy* 76:169-98

---, [1979b] "Individualism and the Mental," *Midwest Studies in Philosophy IV: Studies in Metaphysics*. Minneapolis: University of Minnesota Press

Cartwright, R.[1979] "Indiscernibility Principles" in Cartwright [1987] 200-215

--- *Philosophical Essays,* [1987] Cambridge, MA: MIT Press

Castaneda, Hector-Neri [1966],"He: A Study in the Logic of Self-Consciousness" *Ratio*, 8, no. 2:130-157

Etchemendy, John [1988] "Tarski on Truth and Logical Consequence," *Journal of Symbolic Logic* 53, no.1: 35-63

Evans, Gareth, [1980] *The Varieties of Reference*, Oxford: Oxford UP

Evans, Gareth [1981] "Understanding Demonstratives" in H. Parret and Jacques Bouveresse, eds. *Meaning and Understanding*, (Berlin: W. de Gruyter)

Frege, Gottlob [1956] "The Thought: A Logical Inquiry" trans. A.M. and Marcelle Quinton *Mind* 65 289-311 reprinted in E. D. Klemke, ed. page numbers cited from this reprint.

Kaplan, David, [1989] "Demonstratives" in Almog, Perry and Wettstein (1989) 481-563

Klemke, E. D. ed. [1968] *Essays on Frege*, Urbana, IL: University of Illinois Press

Kripke, S. [1980] *Naming and Necessity*, Cambridge, MA: Harvard UP (1980)

Lewis, David, [1970] "General Semantics," *Synthese* 22:18-67

Lewis, David [1979] "Attitudes *De Dicto* and *De Se"* The Philosophical Review* 88:513-43

McDowell, John [1984] *"De Re* Senses," in Crispin Wright, ed. *Frege: Tradition and Influence* Oxford: Basil Blackwell 98-109

Owen, G.E.L. [1957] "Logic and Metaphysics in Some Earlier Works of Aristotle" *Aristotle and Plato in the Mid-Fourth Century*, ed. I. Düring and G.E.L. Owen, Papers of the Symposium Aristotelicum held at Oxford in August 1957 (Göteburg: Elanders Boktryckeri Aktiebolag, 1960), 163-190. Reprinted in English in *Articles on Aristotle*, ed. J. Barnes, M. Schofield, and R. Sorabji, vol. 3 (London: Duckworth 1979) 13-32

Perry, John [1977] "Frege on Demonstratives", Philosophical Review 8: 474-97, reprinted in [Perry 1993] 3-26 page references cited from this reprint.

---, [1979] "The Problem of the Essential Indexical" *Nous*, 13: 3-21, reprinted in Perry [1993] 33-50 page references cited from this reprint.

---, [1980] "Belief and Acceptance" *Midwest Studies in Philosophy* 5:533-42 reprinted in [Perry 1993] 53-65 page references cited from this reprint.

---, [1993] *The Problem of the Essential Indexical and Other Essays* New York: OUP

Putnam, Hilary "The Meaning of 'Meaning'" [1975] in K. Gunderson, ed. *Language, Mind and Knowledge,* Minnesota Studies in the Philosophy of Science 7, Minneapolis: University of Minnesota Press. reprinted in *Mind, Language and Reality: Philosophical Papers* Volume 2 ed. Hilary Putnam New York: Cambridge UP (1975) page references from this reprint.

Quine, W.V.O, [1953] "Reference and Modality" in Quine [1953] *From a Logical Point of View,* New York: Harper and Row

---, [1953] "Three Grades of Modal Involvement" *Proceedings of the 11th International Congress of Philosophy*, Brussels, vol. 14 Amsterdam: North-Holland Publishing Co. reprinted in Quine (1966) 159-176

---, [1960], *Word and Object,* Cambridge, MA: MIT Press

---, [1961] "Reply to Professor Marcus" *Synthese* 20, reprinted in Quine (1966) 177-184

---, [1966] *The Ways of Paradox*, Cambridge: Harvard UP

Richard, M. [1983] "Direct Reference and Ascriptions of Belief." *Journal of Philosophical Logic* 12: 425-52

---, [1987] "Quantification and Leibniz's Law" *Philosophical Review* 96:555-78

---, [1990] *Propositional Attitudes*, Cambridge: Cambridge UP

Salmon, N. [1986] *Frege's Puzzle* Cambridge, Mass.: MIT Press

Salmon, Nathan and Scott Soames, *Propositions and Attitudes*, New York: Oxford UP

Shoemaker, Sydney, [1968] "Self-Reference and Self-Awareness," *Journal of Philosophy* 65: 555-567

Soames, S. [1987] "Direct Reference, Propositional Attitudes, and Semantic Content" *Philosophical Topics* 15: 47-87

Stalnaker, Robert, [1990] "Narrow Content" in Anderson and Owens, eds. *Propositional Attitudes.*

---, [1991] "How to do Semantics for the Language of Thought" in Barry Loewer and Georges Rey, eds. [1991] *Meaning in Mind: Fodor and his Critics* Oxford: Basil Blackwell

Tarski, A. [1933] "The Concept of Truth in Formalized Languages" in *Logic, Semantics, Metamathematics* . New York: Oxford, Clarendon Press, 1956 152-210