In this article I offer a resolution of Kripke's belief puzzle. The puzzle runs as follows, in Kripke's own words:

Suppose Pierre is a normal French speaker who lives in France and speaks not a word of English or of any other language except French. Of course he has heard of that famous distant city, London (which he of course calls 'Londres') though he himself has never left France. On the basis of what he has heard of London, he is inclined to think that it is pretty. So he says, in French, "Londres est jolie."

On the basis of his sincere French utterance, we will conclude:

(4) Pierre believes that London is pretty.

I am supposing that Pierre satisfies all criteria for being a normal French speaker, in particular, that he satisfies whatever criteria we usually use to judge that a Frenchman (correctly) uses 'est jolie' to attribute pulchritude and uses 'Londres'--standardly--as a name of London.

Later, Pierre, through fortunate or unfortunate vicissitudes, moves to England, in fact to London itself, though to an unattractive part of the city with fairly uneducated inhabitants. He, like most of his neighbors, rarely even leaves this part of the city. None of his neighbors know any French, so he must learn English by 'direct method', with-
out using any translation of English into French: by talking and mixing with the people he eventually begins to pick up English. In particular, everyone speaks of the city, 'London', where they all live. Let us suppose for the moment—though we will see below that this is not crucial—that the local population are so uneducated that they know few of the facts that Pierre heard about London in France. Pierre learns from them everything they know about London, but there is little overlap with what he heard before. He learns, of course—speaking English—to call the city he lives in 'London'. Pierre's surroundings are, as I said, unattractive, and he is unimpressed with most of the rest of what he happens to see. So he is inclined to assent to the English sentence:

(5) London is not pretty.

He has no inclination to assent to:

(6) London is pretty.

Of course he does not for a moment withdraw his assent from the French sentence, "Londres est jolie": he merely takes it for granted that the ugly city in which he is now stuck is distinct from the enchanting city he heard about in France. But he has no inclination to change his mind for a moment about the city he still calls 'Londres'.

This, then, is the puzzle. If we consider Pierre's past background as a French speaker, his entire linguistic behavior, on the same basis as we would draw such a conclusion about many of his countrymen, supports the conclusion ((4) above) that he believes that London is pretty. On the other hand, after Pierre lived in London for some time, he did not differ from his neighbors—his French background aside—either in his knowledge of English or in his command of the relevant facts of local geography. His English vocabulary differs little from that of his neighbors. He, like them, rarely ventures from the dismal quarter of the city in which they all live. He, like them, knows that the city he lives in is called 'London' and knows a few other facts. Now Pierre's neighbors would surely be said to use 'London' as a name for London and to speak English. Since, as an English speaker, he does not differ at all from them, we should say the same of him. But then, on the basis of his sincere assent to (5), we should conclude:

(7) Pierre believes that London is not pret-
KRIPKE'S BELIEF PUZZLE

Of course, (4) and (7) seem to involve Pierre in a straightforward inconsistency. As Kripke puts it: "So we must say that Pierre has contradictory beliefs, that he believes that London is pretty and he believes that London is not pretty". Yet, given the set-up of the story, it seems clear that Pierre is not being inconsistent at all.

Let us call this form of the puzzle the inconsistency version. Kripke now presents the puzzle in an even stronger form:

Again, we may emphasize Pierre's lack of belief instead of his belief. Pierre, as I said, has no disposition to assent to (6). Let us concentrate on this, ignoring his disposition to assent to (5). In fact, if we wish we may change the case: Suppose Pierre's neighbors think that since they rarely venture outside their own ugly section, they have no right to any opinion as to the pulchritude of the whole city. Suppose Pierre shares their attitude. Then, judging by his failure to respond affirmatively to "London is pretty", we may judge, from Pierre's behavior as an English speaker, that he lacks the belief that London is pretty: never mind whether he disbelieves it, as before, or whether, as in the modified story, he insists that he has no firm opinion on the matter.

Now (using the strengthened disquotational principle), we can derive a contradiction, not merely in Pierre's judgments, but in our own. For on the basis of his behavior as an English speaker, we concluded that he does not believe that London is pretty (that is, that it is not the case that he believes that London is pretty). But on the basis of his behavior as a French speaker, we must conclude that he does believe that London is pretty. This is a contradiction.

Let us call this version of the puzzle, in which we seem to be driven to a contradiction in describing a case which seems perfectly possible, the contradiction version of the puzzle.

To buttress the reasoning that lies behind his analysis of the case, Kripke provides the following plausible principles:

A. The Disquotational Principle: "If a normal English speaker, on reflection, sincerely assents to 'P', then he believes that p" (where "... 'P'" is to be replaced inside and outside all quotation marks by an appropriate
B. The Biconditional Form of the Disquotational Principle: "A normal English speaker who is not reticent will be disposed to sincere as­sent to "p'' if and only if he believes that p" (where "... any appropriate English sentence may replace 'p' throughout").[6]  

C. The Principle of Translation: "If a sentence of one language expresses a truth in that lan­guage, then any translation of it into any other language also expresses a truth (in that other language)."[7]  

These three principles are indeed necessary for Krip­ke's reasoning. He needs the Disquotational Principle in order to move to (7) from Pierre's assent to (5), and also to move from Pierre's assent to 'Londres est jolie' to (8):  

(8) Pierre crois que Londres est jolie,  

and he needs his principle of translation to move from (8) to (4). Thus, in the inconsistency version of the puzzle he needs the Disquotational and the Translation Principles. In the contradiction version of the puzzle, however, Kripke needs the Biconditional Form of the Disquotational Principle in order to move from Pierre's non-assent to 'London is pretty' to the conclusion that Pierre doesn't believe that London is pretty. Thus, since by this principle, in our case, 'Pierre believes that p' implies 'Pierre would be disposed to sincere reflective assent to 'p''', his not being disposed to sincere reflective assent to (6) would entail that he doesn't believe that London is pretty. This would contradict (4), which, again, resulted from (8) (via the Principle of Translation), which resulted (via the Disquota­tional Principle) from his assent to 'Londres est jolie'.  

In this paper I shall treat Kripke's puzzle by drawing a sharp distinction between beliefs of a certain believer, which are constitutive of his doxastic state, on the one hand, and belief-sentences on the other; in other words, a distinction between 'p' being a belief of a cognizer r and the established ordinary-language construction of 'r believes that p'. I will treat beliefs as linguistic (sym­bolic) representations (at least where limited to cognizers with minimal logical, conceptual and linguistic acumen),[8] and analyze what I shall henceforth call belief-sentences, i.e., sentences of the form 'r believes that p' (read de dicto) in terms of beliefs. It will be argued that such belief sentences do their reporting job not by displaying beliefs but rather by describing them via the mechanism of adequate paraphrasing, while sentences to the effect that one is disposed to assent or not to assent do display one's beliefs. I will argue that the question of whether a be­
Kripke's Belief Puzzle

The consistency of a believer lies squarely within the level of his beliefs. That is, the consistency of a believer is a function of the logical relations between his beliefs and of the structure of his representational system. It is not determined by the linguistically entrenched manner in which we report his beliefs, nor by the manner in which his beliefs hook up with the world (as by the reference relation). These contentions will allow us to propose a solution to both the consistency and the contradiction problems.

Section II: Beliefs

In the discussion below the notion of 'p' being a belief of r (for a believer r) will play a major role. My main working hypothesis involves taking beliefs as linguistic representations. At a given time t a cognizer can be said to be in a certain belief-state, reflecting whatever beliefs he has at the time. Such belief-states are classifiable via particular beliefs: if 'p' is a belief of our cognizer r (as I shall henceforth call our believer) at time t, then r can be said to be in a 'p'-belief-state. (I remind the reader that I limit my discussion in this paper to cognizers with minimal logical, linguistic and conceptual acumen.[9])

Of course, having particular beliefs is a constituent of the causal order. The acquisition of beliefs can be caused by various stimuli,[10] and r's possession of certain beliefs can cause modes and dispositions of behavior, not the least of which would be his verbal dispositions to assert, or to assent to, various sentences, or to refrain from such assents. Beliefs, being linguistic representations, that is, sentences, are in one language or another: a cognizer r may have 'the king is bald', but not 'le roi est chauve', count as one of his beliefs.[11] This would be attested to by his disposition to assent to the first but not to the second in 'appropriate' circumstances (which would obtain, for instance, when he knows no French). If he assents to both in 'appropriate' circumstances, then they constitute two distinct beliefs of his. The perspective underlying the discussion here should be contrasted with the theory which takes beliefs to be propositions.[12] On the line taken here, a belief of r is a linguistic representation of his in a particular language, which need not (and normally would not) be context-independent: among one's beliefs there could be such as 'I am here' or the 'the President is eating now'. A multi-lingual believer possesses beliefs in more than one language, and thus may possess various distinct beliefs in different languages he speaks which convey the same message.

I am thus operating within a framework in which beliefs are taken as linguistic representations, i.e., sentences. That a cognizer r has a certain sentence as a belief of his
Beliefs belong, therefore, to the level of the internal representations of a believer; they fully reside, so to speak, within his belief-world. To specify the beliefs of a cognizer is to characterize his belief-world purely internally, that is, without making any essential use of, or having to assume the existence of, any particular object outside the believer himself. Beliefs, being linguistic representations of a cognizer, can, of course, be characterized and classified in various ways, e.g., via the causal relations that entities outside the cognizer's belief world bear to their being possessed by the cognizer. Such characterizations are involved in analyzing belief sentences read de re,[13] but can be ignored for my purposes in this essay.

In what follows I shall disregard most of the causal roles played by one's having certain beliefs, roles such as being an effect of various events, or as being a cause of modes of behavior. Rather, I shall concentrate on the relation between having certain beliefs and the cognizer's dispositions to verbal behavior. A disposition to assent to 'p' by r at t under 'appropriate' circumstances will indicate that 'p' is a belief of r (at t); similarly his disposition not to assent to 'p' in the appropriate circumstances will indicate that 'p' is not a belief of r.[14] Such 'appropriate' circumstances will involve, of course, r's sincerity,[15] consciousness, reflectiveness etc. Without attempting here to characterize these 'appropriate' circumstances under which a disposition to assent to 'p' indicates that 'p' is a belief of r, I shall call them Belief Indicating Circumstances (in short: BIC). Again, saying that r's being disposed to assent to 'p' in BIC indicates that 'p' is a belief of r is not intended as an analysis of the notion of belief. Rather, the relation between the two is taken as an empirical connection (or a bridge-rule) associated with the notion of belief so conceived. The above remarks are thus to be taken as indicating sketchily the type of cognitive model I presuppose and work with here in assessing Kripke's claim that there is a puzzle about the belief construction 'r believes that p' in ordinary language. The notion of belief presupposed here can thus be taken as a theoretical notion in an appropriate model of a cognitive system.

Notice that taking beliefs as linguistic representa-
KRIPKE'S BELIEF PUZZLE

...would suggest that the possession of different beliefs (differing, of course, as linguistic representations, i.e., as sentences, regardless of whether they differ in 'content') can play different causal roles in r's interaction with the world. If r, who knows no French, has 'today is Monday' as a belief of his (at a given time), and is now reliably informed that if 'aujourd'hui est Lundi' is true then there are $100 in a barrel in the next street, he need not display the behavior of looking for that barrel in a way in which he would if he had 'aujourd'hui est Lundi' as a belief of his (e.g., had he known French). Of course, having a verbal disposition to assent to 'aujourd'hui est Lundi' is a primary behavioral indication of 'aujourd'hui est Lundi' being a belief of r, regardless of whether or not 'today is Monday' is a belief of his.

In what follows we shall, on many occasions, need the notion of 'a disposition to assent to p at BIC'. Instead of introducing a new symbol for that locution, I shall just use the phrase 'assent to p'. So, when I say 'r' assents to p I shall assume that BIC prevail, and that 'assent' is understood as 'disposition to assent'. (I shall call the believer under discussion r throughout, after Quine's famous Ralph, rather than use Kripke's 'Pierre' and 'Jones').

SECTION III: BELIEF SENTENCES

So far we have discussed beliefs, which pertain to the realm of facts concerning r's doxastic state, with no regard for how such facts are described in a constructionally entrenched way in ordinary language. At this point I wish to draw attention to the ordinary language construction 'r believes that p' in its de dicto reading. (Henceforth, unless otherwise stated, I shall confine myself to the de dicto reading of this sentence, even when I do not specifically say so). Surely, if p is an English sentence,[16] and r assents to p in the 'right' circumstances, thus indicating that p is a belief of his, then indeed r believes that p. To be on the safe side, we shall limit ourselves to cases where r understands the sentence p to which he assents and uses the linguistic items in p in a standard way (relative to the linguistic community to which he belongs, which uses the dialect to which p belongs).[17] Indeed, we shall conduct the discussion in the next few sections assuming this restriction, as does Kripke,[18] which is perfectly congruent with the puzzle under discussion.

Often, however, we claim that r believes that p on the basis of utterances of his to the effect that p, even though we have never heard him assert or assent to p. We might make such a claim to indicate that we would expect r to assent to p, i.e., that p is indeed a belief of his. But if r's assertions were expressed in another language,
while r knows no English, such an expectation would no longer be defensible. It is clear, though, that this claim could still be made if r's assertion to 'q' in another language is an adequate paraphrase of 'p'. Thus, even though it is true that r believes that p, it need not be the case the 'p' is a belief of r; but then there should be some 'q', which is a belief of r, which is an adequate paraphrase of 'p'. If r asssents to 'q' in another language, it would be a matter of an inductive inference to conclude that 'p', which is an adequate paraphrase of 'q' in English, is a belief of r, given that he knows English. But it would be a matter of a valid logical inference to claim that r believes that p. This would hold as well when 'q' and 'p' belong to the same language, or to different dialects in the same language (e.g., when scientific jargon is involved: if r asserts 'John is blue', a psychiatrist may validly infer that r believes that John is depressed, regardless of whether r is familiar with the predicate 'is depressed'.) Let us therefore introduce the predicate 'PBr' as in 'PBr', as follows (the 'P' in 'PBr' is to serve as reminder of 'paraphrase'): [19]

\[ PBr'p' \text{ def. : } \text{"PBr'p'" is true iff for some 'q' such that 'p' is an adequate paraphrase of 'q', 'q' is a belief of r.} \]

This, of course, is introduced for the purpose of claiming that (for an English sentence 'p'):

Analysis of 'r believes that p': \[ r \text{ believes that } p \text{ is true iff } PBr'p'. \]

Of course, the idempotent paraphrase (of 'p' by 'p') is an adequate paraphrase, and thus, if 'p' is a belief of r, then r believes that p. [20] More generally, if 'a' is a sentence of L, and 'Br' is an adequate paraphrase of 'r believes that' in L, so that 'Br'a' is a well formed sentence of L, then:

\[ Br'a' \text{ is true (in } L) \text{ iff } PBr'a'. \]

The construction 'r believes that p' therefore involves essentially a relation such as adequate paraphrase, which has the form ap('a', 'p'), being read as 'a is an adequate paraphrase of p', where 'a' and 'p' are two linguistic items. (Of course, in full form it should involve the languages as well: it should have the form ap('a', L1, 'p', L2), where 'a' is a phrase in L1 and 'p' in L2. For simplicity, however, we shall assume that the phrases carry the languages they belong to 'on their sleeves', as is so often the case, and use the two-place form.) What we therefore get from the above analysis is the validity of the following inference pattern:

I 'q' is a belief of r
KRIPKE'S BELIEF PUZZLE

\[ \text{ap('q', 'p')} \]

\[ r \text{ believes that p.} \]

Adequate paraphrasing is a cornerstone of the indirect speech construction. It is often used within the boundaries of a given language (common to the believer and the speaker); but it figures critically in adequate translation. Now it will not be my business in this paper to analyze this adequate paraphrase relation, which lies at the heart of the belief construction 'r believes that p' (and, for that matter, of other constructions such as 'says that', 'denies that', and other propositional attitudes). The use I shall make of this notion of adequate paraphrasing for the purposes at hand will rely primarily on its intuitive force, and further elaborations will be made along the way as the need arises. My purpose here is to defend the ordinary language belief construction against Kripke's charge that it is afflicted with paradox. For that purpose a general and thorough analysis of the notion of adequate paraphrase will not be necessary. What I claim is that with it stands or falls the ordinary-language belief construction, which is free of the paradoxes imputed to it. The ordinary-language belief construction 'r believes that p' depends upon the notion of adequate paraphrase in a way in which the notion of 'p' being a belief of r does not depend. My point will be that the sources of Kripke's puzzle can be brought to light, given this analysis of belief sentences which uses this notion of adequate paraphrase, in a way that does not make this notion paradoxical or even more suspect. Thus, the intuitive use of this notion (under the elaborations to be discussed below) can be maintained without providing a full-fledged analysis for it. Objections to this notion or requests for its analysis will thus have to be made independently of Kripke's puzzle. In attempting to free ourselves, while working with the belief construction which is based on the notion of adequate paraphrase, from the suspicion of paradox which Kripke has imputed to this construction, we need not worry, for this purpose, about possible inadequacies or unclarities in the notion of adequate paraphrase which are unrelated to the problem at hand, which is Kripke's puzzle.[21]

The implications of the foregoing analysis of belief sentences (which will henceforth be taken to be sentences of the form 'r believes that p') are in the main part congruent with the spirit of various observations of Kripke's. Thus, we get from this analysis that if r assents to 'p' in 'appropriate' circumstances, and therefore 'p' is a belief of r, then, if \[ \text{ap('p', 'q')} \], r believes that q, where 'p' is, say, a French sentence (but 'q' of course is English). Kripke's Disquotational Principle will lead him from r's assent to 'p' to 'r crois que p' and, via his Translation Principle, to 'r believes that q'.

377
It is important for the spirit of my general direction here that I should emphasize the distinction between, on the one hand, what I take to be the (primarily doxastic) facts of the matter—r’s having particular beliefs (as well as various features of those beliefs), and, on the other hand, the level of the ordinary-language constructions which serve the communicative role of describing these facts. To say in our semi-technical jargon that ‘p’ is a belief of r is to display the belief in question; but to say that r believes that p is merely to report such a belief. This reporting function has the force of an existential claim to the effect that there is a belief with such-and-such features (i.e., which is an adequate paraphrase of a given sentence; see the ‘PB(r,’p)’ def. above), and thus serves to report that r has a belief of a certain kind. The same holds for the de re reading of belief sentences, though this will not be discussed here.[22]

Linguistic constructions designed to serve a certain communicative role can play their role better or worse in comparison with potential alternatives; and communicative roles of ordinary-language constructions scarcely coincide with the roles of locutions which are to serve a particular scientific investigatory function.[23] It is therefore quite important, when elements relating to the doxastic-level are discussed, to focus attention on the beliefs of the cognizer rather than on the linguistic construction which serves to report them in ordinary language and on the particular semantical features of that construction. Fit to fulfill its communicative role, a construction in ordinary language possesses semantical features which are adept in this role, but need not be useful for specialized investigatory purposes. There is little doubt that the belief construction is quite useful for reporting through one formulation (in a certain language, or dialect) beliefs formulated in another. This is made possible thanks to the adequate paraphrasing device. But since the belief construction doesn’t display the beliefs in question but rather describes them in a certain way (or rather makes an existential claim under certain classificatory specifications), dealing with epistemological questions via the belief construction is comparable to watching a subject through a veil: it keeps us from a closer touch with the doxastic phenomena themselves. It would be a mistake then to attribute features which reside at the level of our doxastic phenomena to the level of the linguistic constructions which report those phenomena, and vice versa. Not keeping these two levels clearly separated is a mistake that can, and often does, easily lead us astray.

Thus, for instance, whether a cognizer is consistent or not is a question which lies wholly at the level of his beliefs. Reporting the beliefs of a cognizer via the belief construction amounts to presenting the subject’s belief-world from the perspective of the reporter’s belief-world;
and when the two differ—in particular, when the reporter's information is wider in certain ways than that of the subject—the resulting report can have awkward features. But we should be quite careful not to confuse that with inconsistency of the subject himself. If the subject were to acquire extra information which he does not possess, he may become inconsistent (and may have to give up certain beliefs). Reporting his beliefs from a framework equipped with such extra information by describing them rather than displaying them can thus be easily misleading regarding the actual doxastic state of the believer, and can thus fail to reflect the specific features of the subject's doxastic state which determine whether he is consistent. These reflections will be seen to bear on the consistency sub-problem of Kripke's puzzle[24] since, as I will claim, consistency is a purely internal feature of a cognizer's belief-world. The belief construction "r believes that p", however, instead of displaying a belief of r, only yields that there is a certain belief classified modulo the adequate paraphrase relation.

**SECTION IV: INFERENCE**

In sections II and III we have presented the notion of belief and an analysis of belief sentences. Our discussion here, it will be recalled, has been limited to believers with minimal logical and linguistic acumen and to cases of stable dispositions to assent or not to assent to 'p', that is, dispositions that are uniform under varying circumstances evoking assent or non-assent to 'p'.[25] We have also assumed that dispositions to assent or not to assent are taken in Belief Indicating Circumstances, and that the cognizer under discussion uses his terms in a standard way, i.e., he understands correctly the items in the sentences under discussion to which he is disposed to assent or not to assent.[26] I shall now use our notion of belief, our analysis of belief sentences and the notion of one's disposition to assent or not to assent, to quickly go through some valid inferences which connect them (and also some invalid ones) which will be useful in the sequel. The reader who wishes to move directly to the main thrust of our treatment of Kripke's puzzle may skim though this section and move directly to sections V and VI, where the puzzle is handled (he can then go back to the inferences here as the need arises, via the back-references in the following sections.

In section III we presented inference I, which we derived from the analysis of the belief sentence 'r believes that p' as 'Pr p'. Since, obviously, adequate paraphrase is a reflexive relation, we get as an immediate consequence:

\[ \text{II } 'p' \text{ is a belief of } r \]
\[ r \text{ believes that } p. \]
Under the limitations specified above (which we shall take to be operative throughout the following discussion unless and until specified otherwise), a disposition to assent to 'p' in Belief Indicating Circumstances (for short: BIC) indicates that 'p' is a belief of r, and vice versa: 'p' being a belief of r would give rise to his disposition to assent to 'p' (in BIC). (Recall that we abbreviate 'r is disposed to assent to 'p' under BIC' by 'r assents to 'p'' and 'r is disposed not to assent to 'p' under BIC' by 'r does not assent to 'p''). We thus have as valid:

III r assents to 'p'

'p' is a belief of r.

and likewise:

IV 'p' is a belief of r

r assents to 'p'.

From II and III we get:

V r assents to 'p'

r believes that p.

From III and I we get:

VI r assents to 'p'

ap('p', 'q')

r believes that q.

Now for 'p' and '-p'[27] to be beliefs of r, in our present setup,[28] is for r to be inconsistent. For r to have neither 'p' nor '-p' as beliefs is for him to suspend judgment concerning 'p': Thus, 'p' not being a belief of his is compatible with '-p' being a belief of his and is also compatible with '-p' not being a belief of his (even when he is consistent). Hence the inference:

VII x -( 'p' is a belief of r)

'-p' is a belief of r

is invalid (we shall use a lower-case 'x' on the side of the inference to indicate invalidity). Again, if '-p' is a belief of r, 'p' may not be a belief of his on pain of his being inconsistent; but since inconsistency of a believer is a prevalent phenomenon not to be disregarded, the inference:

VIII x '-p' is a belief of r
KRIPKE'S BELIEF PUZZLE

\((-('p' is a belief of r))\)

is invalid, even though

\[IX\]
\[
\begin{align*}
\text{r is consistent} \\
\text{\('p' is a belief of r)}
\end{align*}
\]

\(-('p' is a belief of r)\)

is valid. Thus, due to VII, since 'p' not being a belief of r is compatible with 'p' not being a belief of r, the inference

\[X\]
\[
\begin{align*}
\text{r believes that \(-p\)}
\end{align*}
\]

is also invalid. This is so since the conclusion will be true only if (by the analysis of 'r believes that p') there is a belief of r which is an adequate paraphrase of 'p'. But the premise, though allowing for 'p' to be a belief of r without an inconsistency on r's part, does not require that 'p' be a belief of r, much less that an adequate paraphrase of it be a belief of r. Thus, X would be invalid even with the added premise that r is consistent, i.e.:

\[XI\]
\[
\begin{align*}
\text{r is consistent} \\
\text{\('p' is a belief of r)}
\end{align*}
\]

\[r\text{ believes that } -p.\]

The inference:

\[XII\]
\[
\begin{align*}
\text{r assents to 'p')} \\
x
\text{\('p')}
\end{align*}
\]

\[r\text{ believes that } p.\]

is also invalid, since 'p' not being a belief of r is consistent of course with all adequate paraphrases of 'p' not being beliefs of r. (Notice, of course, that VII and XII would remain invalid even if we added to them a premise to the effect that r is consistent).

'r believes that p' can be true in case some 'q' (in some other language), which is an adequate paraphrase of 'p', is a belief of r, without 'p' being a belief of r. (The case in which r knows no English will be an obvious special case here; 'p', of course, is in English, as we assume throughout when the construction 'r believes that p' is concerned, to ensure its being well-formed.) It is therefore clear that the set:

\[a\{('q' is a belief of r) \mapsto (('q','p')\}\}

is consistent.

381
is consistent, even when 'p' is an English sentence due to cases in which r knows no English. To become convinced that α can be consistent even when r knows English and uses the items involved in a standard way, we may make use of the 'Londres' case in which

\[
\alpha' \begin{cases} 
\text{'Londres est jolie' is a belief of } r \\
\neg (\text{'London is pretty' is a belief of } r)
\end{cases}
\]

was a set of true sentences, thus consistent. That α is consistent when r knows no English, and yet is a consistent believer, is quite obvious. That α can be consistent when r knows English and is a consistent believer will depend on how we resolve the problem of r's consistency in the 'Londres' case. If we become convinced that r can be consistent in this case, then his consistency is compatible with α and his knowledge of English. In general, the consistency of α will not be affected if we add to it the extra member 'r is consistent'. In order to make this claim in particular for cases when r knows English (and, of course, uses his terms in a standard way), we will have to await the resolution of the inconsistency version of the puzzle (section V below).

From the consistency of α (by I) it is clear that:

\[
\begin{cases} 
\text{r believes that } p \\
\neg (\text{'p' is a belief of } r)
\end{cases}
\]

is consistent. Hence the following inference is invalid:

\[
\begin{array}{c}
\text{XIII} \\
\hline
p \\
\hline
\text{r believes that } p
\end{array}
\]

Consequently, obviously its converse:

\[
\begin{array}{c}
\text{XIV} \\
\hline
\neg (\text{'p' is a belief of } r) \\
\hline
\neg (\text{r believes that } p)
\end{array}
\]

is invalid too, and, due to III, so is the inference:

\[
\begin{array}{c}
\text{XV} \\
\hline
\neg (\text{r assents to 'p'}) \\
\hline
\neg (\text{r believes that } p)
\end{array}
\]

and so is, therefore, its converse:

\[
\begin{array}{c}
\text{XVI} \\
\hline
\text{r believes that } p \\
\hline
\text{r assents to 'p'}
\end{array}
\]
Kripke's Belief Puzzle

From the consistency of \( a \) we get the invalidity of:

\[
\text{XVII} \quad \frac{\text{'q' is a belief of } \mathfrak{r}}{\text{ap('q', 'p')} \quad \text{'p' is a belief of } \mathfrak{r}}
\]

as well as the invalidity of:

\[
\text{XVIII} \quad \frac{\text{-('p' is a belief of } \mathfrak{r}) \quad \text{ap('q', 'p')} \quad \text{-('q' is a belief of } \mathfrak{r}).}
\]

Cases in which \( \mathfrak{r} \) knows no English ('p' being an English sentence) proved \( a \) to be consistent, even when the believer \( \mathfrak{r} \) is consistent, as we have seen above. Similarly, it can be easily seen that, more generally, cases in which 'q' and 'p' belong to different languages, one of which \( \mathfrak{r} \) doesn't know at all, would make \( a \) consistent for a consistent believer, regardless of which of these languages, if any, is English. Such cases would therefore symmetrically show the set

\[
\{-('q' is a belief of } \mathfrak{r}) \quad \text{ap('q', 'p')} \quad \text{p' is a belief of } \mathfrak{r}\}
\]

to be consistent too, when 'p' is an English sentence. Hence also:

\[
\{-('q' is a belief of } \mathfrak{r}) \quad \text{ap('q', 'p')} \quad \text{r believes that } p\}
\]

is consistent. Hence:

\[
\text{XIX} \quad \frac{\text{-('q' is a belief of } \mathfrak{r}) \quad \text{ap('q', 'p')} \quad \text{-(r believes that } p)}{\text{-r asssents to 'p'}}
\]

is invalid as well. And similarly (from XIX and III), so is:

\[
\text{XX} \quad \frac{\text{-(r assents to 'p')} \quad \text{ap('q', 'p')}}{\text{-(r believes that q)}}
\]

Since \( a \) would remain consistent with the addition of 'r is consistent', XII, XIV, XV, XVI, XVII, XVIII, XIX and XX would remain invalid even with an added premise to the effect that r is consistent. Given the remark above (before inference III), the claim that this will remain so in par-
ticular for cases in which r knows English (and uses his terms in a standard way) is conditional upon whether he can be consistent in a case such as the one described in the 'Londres' puzzle.

So far we have explored what we need for the resolution of Kripke's puzzle. But the situation will be clearer if we explore the viability of using a stronger premise in inference XIV; that is, if we explore whether

\[
\text{'}-p'\text{ is a belief of } r \\
\text{-(r believes that } p) 
\]

is valid.

Kripke's case was one in which as Kripke showed:

\[
\begin{align*}
\beta \{ & r \text{ assents to 'London is not pretty'} \\
& \text{ap('Londres est jolie','London is pretty')} \\
& r \text{ assents to 'Londre est jolie'}
\end{align*}
\]

are all true. Hence, the set

\[
\begin{align*}
\lambda \{ & r \text{ assents to '}-p'\} \\
& \text{ap('q','p')} \\
& r \text{ assents to 'q'}
\end{align*}
\]

is a consistent set. As Kripke correctly mentions, this also obtains with kind-terms instead of proper names.

Due to III and the consistency of \( \lambda \),

\[
\mu \{ \begin{align*} & \text{'}-p'\text{ is a belief of } r \\
& \text{ap('q','p')} \\
& \text{'}q'\text{ is a belief of r} \end{align*} \}
\]

is also a consistent set. Hence, with I, also

\[
\rho \{ \begin{align*} & \text{'}-p'\text{ is a belief of } r \\
& r \text{ believes that } p \end{align*} \}
\]

is consistent, and therefore similarly (using \( \rho \) with IV)

\[
\{ r \text{ assents to '}-p'\} \\
\text{r believes that } p
\]

is consistent. Therefore (from \( \rho \)) the inference

\[
\text{XXI } x \quad \text{'}-p'\text{ is a belief of } r \\
\text{-(r believes that } p) 
\]

is invalid, and so is (given XXI and IV)
KRIPKE'S BELIEF PUZZLE

XXII x \( r \) asserts to \(-p\)

\(- (r \text{ believes that } p) \).

Notice that, as Kripke observes, it is intuitively clear in the situation described in his example that \( r \) may very well be consistent. Kripke, however, provides an argument for \( r \) being inconsistent, which generates one aspect of the puzzle. Let us notice, then, that we would be in a position to rely on this intuition if we eliminate this suspicion of inconsistency; in this case we would be able to add \( 'r \text{ is consistent}' \) to \( \beta \) without endangering \( \beta \)’s consistency, and then add the assumption \( 'r \text{ is consistent}' \) to all the inferences which follow, without changing their status as invalid.

From \( \mu \) we thus conclude the invalidity of

XXIII x \( 'p' \text{ is a belief of } r \)

\(- (p', q') \)

\(- (q' \text{ is a belief of } r) \).

From the consistency of \( p \) we get the consistency of \( p' \) (by replacing in \( p \) 'p' by \(-p\)):

\( p' \begin{cases} 'p' \text{ is a belief of } r' \\ r \text{ believes that } -p \end{cases} \)

and thus the invalidity of

XXIV x \( r \text{ believes that } -p \)

\(- (p' \text{ is a belief of } r) \)

and hence also (with IV) the invalidity of:

XXV x \( r \text{ believes that } -p \)

\(- (r \text{ assents to } 'p') \).

Now from the consistency of \( p' \) we get the consistency of

\( p'' \begin{cases} r \text{ believes that } p \\
 r \text{ believes that } -p \end{cases} \)

and hence the invalidity of

XXVI x \( r \text{ believes that } -p \)

\(- (r \text{ believes that } p) \)

Having established the invalidity of the inferences of the last group and the validity of the inferences of the
first group, we can proceed to analyze Kripke's puzzle.

Notice that if we dispose of the argument for $r$'s inconsistency in the situation described in Kripke's puzzle, given our above remark (after inference XXII), inferences XXI, XXII, XXIII, XXIV, XXV, XXVI will remain invalid with an additional premise to the effect that $r$ is consistent.

SECTION V: THE INCONSISTENCY PROBLEM

It will be recalled that, in the set-up of Kripke's puzzle, the believer (Kripke's Pierre, whom I rename $r$) is a cognizer with minimal logical and linguistic acumen who understands the relevant phrases in French and English and uses them in a standard way.[29] This, however, does not protect him ipso facto from a charge of inconsistency. For convenience, we abbreviate the sentences 'London is pretty' and 'Londres est jolie' as 'Pin' and 'Jlr' respectively, and their negations as '-Pin' and '-Jlr' respectively. (Thus, 'P' and 'J' abbreviate the predicates 'pretty' and 'jolie' respectively, and 'In' and 'lr' the names 'London' and 'Londres' respectively.)

As we have seen,

(10) $r$ assents to 'Jlr'

is true (taken as a disposition of $r$ in the 'right' circumstances), as well as

(11) $r$ assents to '-Pin'.

Consequently, under these conditions, given our discussion in the previous sections, we are entitled to infer both:

(12) 'Jlr' is a belief of $r$

(13) '-Pin' is a belief of $r$

(cf. inference III, section IV). Given inference II, we are entitled to infer from (13) that

(14) $r$ believes that London is not pretty.

Similarly, under the above conditions, (12), together with (15):

(15) $ap('Jlr','Pin')$

yield the truth of (via inference I):

(16) $r$ believes that London is pretty.

We can safely conclude (16), because $r$'s use of the
terms involved in his French belief in (12) is quite a standard use. Thus, if we consider \( r \) before the time of his arrival in London, he is, in our story, a normal French speaker in every relevant respect, in particular in his use of 'Londres' and 'jolie'. Surely his lack of knowledge of English does not in the least subject him to a charge of not using 'Londres' in a standard way; and the information he possesses with respect to 'Londres' is no different from (in particular, not richer than) that possessed by a great many French speakers. The fact that he has been unaware that 'Londres est la capitale d'Angleterre' does not detract from his standard use of the term 'Londres', but only leaves him short of some pertinent information concerning that city, which is not unlike information concerning its population, elevation, climate, etc.\[30\] Notice, in particular, that given our story, by 'Londres' \( r \) indeed refers to the city London, and thereby a necessary condition for a standard use of a proper name is thus fulfilled here.\[31\] I shall not analyze this point here,\[32\] but rather assume it to be obvious on intuitive grounds.\[33\]

But surely \( r \)'s learning another language, i.e., English, did not interfere with his knowledge of French, in particular with his standard use of the word 'Londres', since that use of his has not changed one bit: he still holds, after having learned English, the same French beliefs in which 'Londres' occurs that he held before (since he would not take himself to have learned anything concerning 'Londres' while in England, only concerning what he takes to be another city, called 'London'). Hence, since it is obvious that his use of 'Londres' (and, of course, of 'jolie') was standard before he came to England, it remained so thereafter too. Thus, (14) and (16) help us vindicate Kripke's conclusion that, for 'p' taken as 'Pin', it is the case that

\[
(17) \ r \ \text{believes that } p, \text{ and } r \ \text{believes that not } \neg p.
\]

In section III we emphasized the need to separate the level of \( r \)'s beliefs, on the one hand, from the reporting level of the construction 'r believes that p' on the other. We have seen that what such a reporting sentence conveys is that there is a belief of \( r \) which is an adequate paraphrase of 'p'.

Now the notion of the consistency of a believer resides in whether his beliefs yield a contradiction. It resides in whether the use of logical tools alone can allow him to derive a contradiction from the beliefs he possesses, which is tantamount to whether he has contradictory beliefs. Thus, the subject matter of his consistency lies at the level of his beliefs,\[34\] in whether there is a set of beliefs of his such as 'p', '-p', or, more generally, a set of beliefs of his 'p1', ..., 'pn', which is self-contradictory.
Whether, and if so, to what extent, this is reflected by conditions formulated at the reporting level must be carefully examined, and not just taken for granted. Thus, while obviously

\[(18) \ 'p' \text{ is a belief of } r, \text{ and } '¬p' \text{ is a belief of } r\]

will make \(r\) inconsistent by virtue of having contradictory beliefs, \[35\] it remains to be seen whether the condition which Kripke takes to make him inconsistent, i.e.,

\[(17) r \text{ believes that } p, \text{ and } r \text{ believes that } ¬p\]

indeed does so as well.

In so far as the 'Londres' case is concerned, it is clear that \(r\)'s relevant beliefs involve

\[
a^* = \{ 'Jlr', '¬Pln', 'ap('J'(in \ F),'P'(in E))' \} \ [36]
\]

(where the last item merely represents any one of various other possible beliefs to the same effect, \[37\] e.g., '\(J\)' (in \(F\)) is an adequate translation of '\(P\)' (in \(E\)), or: 'J' (in \(E\)) means '\(P\)' (in \(E\)), or: 'J' (in \(F\)) and 'P' (in \(E\)) stand for the same thing, or even: whatever 'J' (in \(F\)) is true of, 'P' (in \(E\)) must be true of too, etc.). Is this set of beliefs self-contradictory? Can a contradiction be derived from it? Of course not. If someone wants to charge self-contradiction, he must include the belief 'ap('lr'(in \(F\)),'ln'(in \(E\)'))' (or some other belief to the same effect), and thus focus on the set of beliefs \(β^*:\)

\[
β^* = \{ 'Jlr', '¬Pln', 'ap('J'(in \(F\))','P'(in \(E\))')' \}
\]

But certainly no one could charge that \(a^*\), which does not include the last item of \(β^*\), is self-contradictory. To realize that, assume that \(r\)'s use is non-standard, and that he takes 'Londres' to be a name of a cat, while 'London' for him is the name of the city. Surely \(a^*\) yields no contradiction here. But this interpretation is not possible for \(β^*\), due to its extra condition.

Thus, surely the set \(a^*\) is compatible with 'ap('lr'(in \(F\))','ln'(in \(E\)))'. Since 'ap('lr'(in \(F\))','ln'(in \(E\)))' is not a belief of \(r\) in our case (nor is any other sentence to the same effect), clearly \(a^*\) in itself is not self-contradictory. Above we have argued that \(r\) uses the items here in a normal way, and in this we are in agreement with Kripke. No logical acumen would enable \(r\) to derive a contradiction from the set \(a^*\), which is not self-contradictory, given his normal use of the items there. We thus see that 'ap('lr'(in \(F\))','ln'(in \(E\)))' (or another sentence to the same effect) is a crucial item missing from \(r\)'s beliefs for a charge of self-contra-
KRIPKE'S BELIEF PUZ ZLE

diction to take off the ground. r is thus consistent in our case, and sustains at the same time our assumption of standard use.

Yet, of course, as Kripke realized, (14) and (16) are true in our case, thus making true:

(17) r believes that p, and r believes that -p.

We therefore see that (17) cannot be a sufficient condition for r's being inconsistent. But how can that be? By our analysis of 'r believes that p', (17) would be true in case for some 's' and 'q', which are beliefs of r, ap(\'s\', p) and ap(\'q\', \tilde{r}-p) (when r uses the items in \'q\' and \'s\' in a normal way). Thus, to check (17) vis-a-vis r's consistency, we must consider whether for r to have \'q\' and \'s\' as beliefs, while ap(\'q\', -s') (where r uses \'q\' and \'s\' normally), would render him inconsistent.

As before, the crucial question is whether 'ap(\'q\', -s')' (or another sentence to the same effect) is a belief of r. It must be, in order for a charge of inconsistency to become serious. If it is not a belief of r, surely no logical acumen of r would help him derive a contradiction from his beliefs \'q\' and \'s\'. If r doesn't realize, therefore, that ap(\'q\', -s'), he cannot be branded inconsistent. But can r fail to realize that ap(\'q\', -s') while \'q\' and \'s\' are beliefs of his, if his use of the linguistic items involved is standard? Here lies the catch: normally, he couldn't. Normally, a believer with a normal understanding of the linguistic items he uses, whether in one language or more, would realize whether or not two items are adequate paraphrases of each other. In many cases, a failure to realize that would detract from his normal use of these items. But not always. We have seen in our example that r can have a perfectly normal use of 'London' and 'Londres', despite his not realizing their being adequate paraphrases of each other. What can happen with 'Londres' and 'London' can happen with other proper names, and with kind terms, too. Even though (to use an example of Putnam and Kripke) 'hetres' and 'beech' are adequate paraphrases of each other (in French and English respectively), r might use them in a normal way and yet feel that they somehow differ from each other in a way he can't specify; he may feel that 'hetres' is a species which grows only on the continent, while beech grow only in North America. Such a mistake would not be grave enough to detract from his normal use of the terms involved. If you will, Kripke's example, once we have shown it to involve no paradox, can be taken to show how one can use two items in a standard way without realizing they are adequate paraphrases of each other.

We thus realize that (17) is not a sufficient condition for r being inconsistent, since it does not entail that r has contradictory beliefs, even in cases with standard use.
of the items involved. We thus see that the characterization of the inconsistency of a believer, belonging properly to the level of his beliefs, cannot be relegated to the reporting level. (17) won't do for determining inconsistency since it does not yield that \('p' and \('-p' are beliefs of \(r\), nor does it yield that \('q' and \('-q' are beliefs of \(r\) for any other \('q'\). To conclude from (17) that \(r\) is inconsistent is thus a mistake, and this mistake underlies the seeming paradox of \(r\)'s inconsistency in Kripke's puzzle.

Rather, inconsistency would occur (under the assumptions of the present discussion) in case

\[(18') \quad \text{'}q' \text{ is a belief of } r \text{ and } \text{'}-q' \text{ is a belief of } r.\]

Of course, \((18') \Rightarrow (17) \) (via inference I, when \('q' is a sentence in \(L\) such that \(ap('q' (in \(L\)), \(p'(in \(E\)), \text{ and hence } ap('-q' (in \(L\)), \(\neg p'(in \(E\))))\); but not vice versa: \((17) \text{ may be true without (18') being true for any } 'q'. \text{ (Of course (17) } \not\Rightarrow \text{ (18) } \text{ in (18') involves } 'p' \text{ as in (17) rather than a possibly different } 'q' \text{ as in (18')) since } r \text{ need not know English (and 'p' is an English sentence). As we saw above, (17) \text{ can be true in the case } r \text{ has two beliefs } 'q' \text{ and } 's', \text{ such that } ap('q', 'p') \text{ and } ap('s', '-p'), \text{ without } ap('q', '-s') \text{ being a belief for } r, \text{ and without a violation of his normal use. In such a case, 'q' and '-s' need not, of course, be the same sentence, and may well belong to different languages. Hence (18') need not be true for any } 'q' \text{ even when (17) is true. Thus, (17) does not logically imply that } r \text{ is inconsistent, and, in our case, it is true without } r \text{ being inconsistent. It is here that Kripke went wrong in the inconsistency problem, in assuming (17) to be a sufficient condition for } r\text{'s being inconsistent. The separation of the level of } r\text{'s belief-world, that is, the level of his beliefs, from the reporting level of belief sentences in ordinary language--this separation being made possible by the conception of beliefs as linguistic representations--allows for the realization that consistency lies primarily at the level of beliefs and not at the reporting level of belief sentences, and also for the realization that condition (17), which resides at the reporting level of belief sentences, is inadequate as a sufficient condition for inconsistency. Indeed, one cannot, it seems, do better than (17) when only the reporting level is available. One must, therefore, move to define the notion of beliefs, and to introduce (18) as a better sufficient condition for inconsistency. Not having the notion of beliefs presented here, Kripke was unable to clearly separate the two levels. Consequently, he uses 'r believes that . . .' and 'that . . . is a belief of r' almost interchangeably in various places.[38] Kripke had therefore to remain at the reporting level, and the result is the misleading conception that (17) suffices for inconsistency.
Krippke's Belief Puzzle

And yet, of course, (17) is not wholly irrelevant to questions of inconsistency. Its seductiveness lies in the fact that normally r would indeed be inconsistent when (17) is true. This is so since the cases where (17) would fail to yield an inconsistency are cases in which two items which occur in r's beliefs, and which he uses standardly, are adequate paraphrases of each other without r realizing it. But this phenomenon, though possible, as we have seen, is relatively uncommon. Unless something like that happens, (17) does indeed yield an inconsistency. Since (17) yields an inconsistency in all cases but one rare sort, (17) therefore constitutes good evidence for inconsistency. Surely if r has two contradictory beliefs 'p' and '-p', (17) would hold. Thus, (17) is indeed a consequence of inconsistency of a certain kind (e.g., which involves only two beliefs). Therefore, for practical purposes, (17) is a reliable indicator of inconsistency. But this is so by way of (17) providing the grounds for a good ampliative inference to the effect of r being inconsistent. Yet (17) fails to be a sufficient condition in that it does not logically entail that r is inconsistent. To determine inconsistency when (17) obtains, one must check whether we don't have a phenomenon of the type indicated, which would block an inconsistency despite (17).

Yet, surely there is a certain oddity in r's having the beliefs in a*, even if this does not amount to an inconsistency. r is in a doxastic state in which, were he to acquire and retain the belief 'lr'(in E) is 'ln'(in E)', or 'ap('lr'(in E)', 'ln'(in E))', or something else to this effect, he would have to give up one of his beliefs in a* (even though it directly contradicts none of them severally). One might say that having the beliefs in a* would make r have incongruent beliefs. But to have incongruent beliefs is not to be inconsistent.

Hopefully, we have now established that in the situation in Kripke's puzzle r's use of his terms may well be normal, and that r can hold his beliefs without being inconsistent, since condition (18), the appropriate indication of inconsistency for such a case, is not satisfied. We can now go back to the last item left open in section IV. In discussing the set β there, describing the facts of the 'London's' case, we left it open whether the addition of 'r is consistent' would leave β consistent. Now we are in a position to answer this question affirmatively. Our conditional conclusion in section IV was that, if the answer to this question is affirmative, inferences XXI, XXII, XXIII, XXIV, XXV and XXVI there would remain invalid with the additional premise 'r is consistent'. We can now assert this conclusion unconditionally.

Of course, since β in section IV would now remain consistent after the addition of the sentence 'r is consistent', so would λ there. It is therefore not surprising now
that \( \lambda \), with this added sentence, yields the consistency of 
(via I and II):

\[
\begin{align*}
\{ & \text{r believes that } -p \\
& \text{r believes that } p \\
& \text{r is consistent}
\end{align*}
\]

reflecting again the inadequacy of (17) as a sufficient condition for r's inconsistency (not to mention the inadequacy of (17) as a necessary condition, which Kripke most likely never held it to be).

Let us illustrate how Kripke goes astray by not separating the reporting of beliefs via belief sentences on the one hand, and the beliefs themselves, on the other, in his discussion of the inconsistency problem. Thus he says:

Suppose that, in France, Pierre, instead of affirming "Londres est jolie," has affirmed, more cautiously, "si New York est jolie, Londres est jolie aussi," so that he believed that if New York is pretty, so is London. Later Pierre moves to London, learns English as before and says (in English) "London is not pretty." So he now believes, further, that London is not pretty. Now from the two premises, both of which appear to be among his beliefs (a) If New York is pretty, London is, and (b) London is not pretty, Pierre should be able to deduce by modus tollens that New York is not pretty. But no matter how great Pierre's logical acumen may be, he cannot in fact make any such deduction, as long as he supposes that 'Londres' and 'London' may name two different cities . . .

Yet, if we follow our normal practice of reporting the beliefs of French and English speakers, Pierre has available to him (among his beliefs) both the premises of a modus tollens argument that New York is not pretty.[40]

The apparent problem is that he should be able to draw the conclusion of a modus tollens argument whose premises he possesses among his beliefs, yet in fact he cannot. But this problem is generated by a failure to realize that the reporting belief construction does not display his beliefs. Thus, in light of our analysis, the last underlined sentence in the quotation comes out as strictly false: his beliefs are 'Si New York est jolie, Londres est jolie aussi' and 'London est not pretty', and these are not premises of a modus tollens argument: (a) in the quotation is not a belief of Pierre. Since he has no premises of a modus tollens argument among his beliefs, he cannot be expected to draw any such conclusion. The mistake here lies again in examining, when exploring logical relations which the believer should
Kripke's Belief Puzzle

realize, sentences in the scope of belief idioms in the belief-sentence construction rather than the beliefs themselves.

Section VI: The Contradiction Problem

We now come to the contradiction version of the puzzle. We still have as true:

(10) r assents to 'Jlr'

but, instead of (11), we have a suspension of judgment on r's part with respect to 'Pln': r has no disposition to assent to or dissent from it; i.e., we have the truth of both

(19) -(r assents to 'Pln')
(20) -(r assents to '-Pln').

Under the assumed circumstances, (10) yields, as before (inference III):

(12) 'Jlr' is a belief of r.

But (19) and (20) now yield (inference IV) both:

(21) -( 'Pln' is a belief of r)
(22) -( 'Pln' is a belief of r).

Again, as above, in the assumed circumstances, (12), together with (15):

(15) ap('Jlr','Pln')

yield (via inference I):

(16) r believes that London is pretty.

Thus, (16) is reached via the analysis of 'r believes that p', according to which 'r believes that London is pretty' is true iff there is some 'q' in some language L, such that ap('q'(in L),'Pln'(in E)), and 'q' is a belief of r. And indeed, (12) and (15) secure that 'Jlr' in French is just such a requisite 'q'.

Can we infer the negation of (16) via (21)? Such an inference would be fallacious, given our analysis of 'r believes that p'. What is required in this analysis for the truth of 'r believes that p' is that, under the given circumstances of standard use, some adequate paraphrase of 'p' be a belief of r. If this requirement is fulfilled, the question of whether some paraphrase of p is not a belief of
r, or whether some adequate paraphrase of \( -p \) is a belief of \( r \) (as was the case in the inconsistency version), is immaterial to whether \( r \) believes that \( p \) according to this analysis, so long as it doesn't undermine the assumption of understanding and normal use under which the above requirement is to be fulfilled. But we have concluded in the previous section that having 'Jir' as a belief while having '-Pln' as a belief as well, in the circumstances of our case, does not reflect negatively on \( r \)'s standard use of the terms involved. A fortiori, this is the case when \( r \) has 'Jir' with neither 'Pln' nor '-Pln' as beliefs.

In particular, the move from (21) (or from (19)) to the negation of (16) will be an instance of inference XIV (or inference XV) which, we have seen in section IV, is invalid.

Let us buttress this point. The question is whether (19), or (21), imply the negation of (16), which is

\((-16) \quad r \text{ does not believe that London is pretty.}\)

Now (-16) is true (via the analysis of belief sentences, section III), provided there is no sentence 'q' (in whatever language), such that \( ap('q', 'Pln') \) and such that 'q' is a belief of \( r \) (again, under our general assumption of normal use by \( r \) of the items in any pertinent belief of his in the language to which this belief belongs). But all that (19), or (21), provides us with is one sentence 'Pln' which does not constitute a belief in \( r \). Given, however, what it takes for (-16) to be true, the information that some sentence is not a belief of \( r \) tells us nothing about whether \( r \) believes that \( p \), even if that sentence is an adequate paraphrase of 'p', or even if that sentence is 'p' itself. This is so since the fact that any given sentence, in particular 'p', is not a belief of \( r \), does not imply that there is no other sentence 'q' such that \( ap('q', 'p') \), and such that 'q' is a belief of \( r \) (while satisfying the requirement of standard use by \( r \)). There may indeed be such a sentence 'q', that would yet not endanger \( r \)'s consistency or standard use (in general, and in particular of 'p', if \( r \) happens to understand 'p'), as we have seen above, as long as \( r \) is not aware that \( ap('q', 'p') \). Even if \( r \) understands 'p' (e.g., if 'p' is 'Pln', in our case) as well as 'q', while both 'p' and 'q' are in the same language, e.g., English, \( r \) need not violate the standard use of 'p' and 'q' while failing to realize that \( ap('p', 'q') \), as Kripke's kind-terms example shows—a fortiori if 'p' and 'q' belong to different languages. So, if an English 'p' is not a belief of \( r \), \( r \)'s linguistic competence and consistency need not rule out even an English 'q' as a belief of \( r \), in case \( ap('q', 'p') \); a fortiori they need not rule out such a 'q' in another language or dialect being a belief of his. Thus, (19) or (21) do not imply (-16).

We thus conclude that we have no contradiction here:
KRIPKE'S BELIEF PUZZLE

(16) is true, and there is no valid argument which leads to its negation from the premises of our case.

But this is so provided, of course, that we reject Kripke's Biconditional Form of the Disquotational Principle. According to this principle (quoted in section I), under the appropriate circumstances r is disposed to assent to 'p' iff r believes that p. What this principle adds to the Disquotational Principle is the validation of (in the appropriate circumstances, when 'p', of course, is an English sentence):

\[
\begin{align*}
\text{r believes that p} & \quad \text{r assents to 'p'} \\
\hline
\neg (\text{r assents to 'p'}) & \quad \neg (\text{r believes that p})
\end{align*}
\]

i.e., the validation of:

\[
\neg (\text{r assents to 'p'}) \quad \neg (\text{r believes that p})
\]

which is inference XV in section IV, shown there to be invalid. And it is indeed this Biconditional Form of the Disquotational Principle which Kripke needs[41] for his derivation of the contradiction version of the puzzle. But if our analysis is correct, this principle is not valid. Hence Kripke's argument for the negation of (16), which is based on this principle, and thus Kripke's argument for there being a contradiction, are undermined. This principle, it will be recalled, is rendered false, given our analysis of belief sentences and the inferences based on this analysis (for an English sentence 'p', and under our usual assumptions), by cases in which some 'q', such that \( \text{ap('q','p')} \) is a belief of r while 'p' is not a belief of r, where 'q' and 'p' are sentences whose items are used by r in a standard way. (That standard use can be retained in such a case has been observed before, in agreement with Kripke; we also showed that r's consistency is not necessarily impaired in such a case.) According to our analysis, in such a case r will not assent to 'p', yet he will believe that p (due to inference I), contrary to Kripke's principle.[42] It is worth noticing that, as Kripke's kind-terms example shows, the failure of this principle does not require knowledge of different languages or dialects on the part of the believer: r might not assent to 'furzes have shallow roots' while assenting to 'gorses have shallow roots', thus making true 'r believes that furzes have shallow roots' (due to the truth of '\( \text{ap('gorse','furze')} \)') without engaging in a non-standard use of the terms involved (or in any inconsistency). Recall that the invalidity of XV, hence of Kripke's Biconditional Form of the Disquotational Principle, was a consequence of our analysis of belief sentences in terms of beliefs. According to this analysis, 'r does not believe that p' is true in case none of a variety of a certain kind
of sentences is a belief of \( r \), while non-assent to \( 'p' \) by \( r \) indicates only that a particular sentence, i.e., \( 'p' \), is not a belief of \( r \). Non-assent to \( 'p' \) leaves it open for other sentences to be beliefs of \( r \), and to give rise to the truth of \( 'r' \) believes that \( p' \).

So far we have examined the suspension of judgment case, and we have seen that one cannot derive a contradiction in this case on the basis of (10) and (19) (or (12) and (21)); see the beginning of this section). But when we discussed the inconsistency problem (cf. the beginning of section V) we had a stronger premise than (19) (or (21)): we had (11) (and thus (13)). (Recall, (11) was: \( r \) assents to \( '-Pln' \). (13) was: \( '-Pln' \) is a belief of \( r \).) Are we likely to generate a contradiction with (11) and (19), or with (12) and (13)? Again, as in the previous discussion, (10) (via (15)) yields (16). Can we also generate (16), the negation of (16), in this stronger version?

Again, however, this is not the case. What we have here, that we did not have in the suspension of judgment case, is the truth of (11), and thus of (13). But it is a mistake to conclude (16) from (11): this is an instance of inference XXII which we have shown to be invalid (section IV). Similarly, it is a mistake to conclude (16) from (13): this would be an instance of inference XXI, which we have also shown to be invalid. Also, in this case we have (14) as true, which we did not in the suspension of judgment case. (Recall, (14) was: \( r \) believes that London is not pretty.) But it would be fallacious to infer (16) from (14): this is an instance of inference XXVI, shown there to be invalid too. And, furthermore, as indicated at the end of Section V, the invalidity of XXI, XXII and XXVI would hold even for a consistent believer.

Let us make things a bit clearer: since (13) is equivalent to (11) (under our assumptions), and since (14) follows from (13), we can limit ourselves to showing that (13) does not imply (16) even for a consistent believer. But now we are back to the case discussed above: (13) (i.e.: \( '-Pln' \) is a belief of \( r' \) does not exclude \( r' \)'s having a belief \( 'q' \), such that \( ap('q','Pln') \), with \( r' \) not being aware that this is the case (to protect his standard use and consistency). Is there such a belief of \( r' \) in our case? Of course: \( 'Jlr' \) can serve as the belief \( 'q' \) which \( r' \) can have together with \( '-Pln' \) without endangering his consistency and standard use, which will make (16) true, thus (16) false.

Thus, it is the invalid inferences XIV, XV, XXI, XXII and XXVI which reflect the mistakes in Kripke's arguments for (16), and thus for the contradiction. Nevertheless, it is easy to see the temptation to follow these inferences. They become invalid, for a consistent believer, only when there are two items, \( q \) and \( p' \), which \( r' \) uses in a standard
way, such that \(\text{ap}(q,p)\), of which \(r\) is not aware. As we pointed out at the end of section V, this is an unusual situation. Normally, when this is not the case, the above inferences would not fail us. Thus, normally, barring unusual cases of this sort, \(r\)'s non-assent to \(p\), or his assent to \(-p\), or his believing that not \(p\), would indeed guide us safely to the conclusion that, being consistent, he does not believe that \(p\). Thus, these are good indicators, and provide good evidence, for \(r\) not believing that \(p\), which will fail us only in the unusual cases where \(r\) fails to be aware of the adequate paraphrase relation between two items he uses in a standard way. They would therefore provide for an adequate ampliative inference to \(r\) not believing that \(p\), even though they do not logically entail it. Such an inference fails us only in fairly recondite cases; but the case under discussion is just such a case.

We thus conclude that we resolved Kripke's two versions of the puzzle in his own terminology. We have pushed the notion of the consistency of a believer back to the level of beliefs, and concluded that the level of belief-sentences does not allow for a formulation of sufficient conditions for inconsistency; in particular, (17) was not a sufficient condition for inconsistency. Thus, concerning the consistency problem, we have seen that when there are two sentences \(p\) and \(q\) which are adequate paraphrases of each other, even though \(r\) is not aware of it, with his standard use intact,

\[ \begin{align*} 
  'p' \text{ is a belief of } r \text{ and } '-q' \text{ is a belief of } r \\
  \text{may be true, thus yielding:} \\
  '-q' \text{ is a belief of } r \text{ and } r \text{ believes that } p, \\
  \text{as well as:} \\
  r \text{ believes that } p \text{ and } r \text{ believes that } -p, \\
  \text{which is not sufficient for inconsistency, thus not yielding:} \\
  'p' \text{ is a belief of } r \text{ and } '-p' \text{ is a belief of } r, \\
  \text{which would reflect an inconsistency. In the contradiction version we have confirmed (16), as Kripke does, but rejected its negation, and rejected the argument Kripke presented for its negation by rejecting Kripke's Biconditional Form of the Disquotational Principle, on the basis of our analysis of belief sentences in terms of beliefs (acknowledging that Kripke's arguments based on this principle were valid, and without rejecting his other two principles of Disquotation and Translation).}
\end{align*} \]
Kripke issues the following warning:

It is no solution in itself to observe that some other terminology, which evades the question whether Pierre believes that London is pretty, may be sufficient to state all the relevant facts.

But none of this answers the original question. Does Pierre, or does he not, believe that London is pretty? I know no answer to this question that seems satisfactory. It is no answer to protest that, in some other terminology, one can state 'all the relevant facts'.

I believe that my account heeds Kripke’s warning and provides an answer to his puzzle as he presents it in his own terminology.

Facing what seems to him to be an insurmountable difficulty ("Since our backs, however, are against the wall."). Kripke considers a radical suggestion to this effect: "decree that no sentence containing a name can be translated except by a sentence containing the phonetically identical name." But he finds this suggestion both contrary to our normal practice of translation and very implausible on its face. He therefore asks: "What is it about sentences containing names that makes them—a substantial class—intrinsically untranslatable, express beliefs that cannot be reported in any other language?"

Our handling of the 'Londres' case finds no fault in translating proper names from one language to another in the standard way—it finds no fault in Kripke's Principle of Translation. Furthermore, it surely does not endorse the conclusion that beliefs which include proper names can't be reported in other languages. They certainly can—and the mechanism for that is indeed supplied via the 'r believes that London is pretty' construction, based on the notion of adequate paraphrase, which allows 'Londres' and 'London' to be adequate paraphrases of each other. According to my analysis, 'r believes that London is pretty' is indeed a correct reporting of r's having the belief 'Londres est jolie'. Therefore, such radical moves as disallowing standard translation of proper names are indeed avoided in my analysis.

Kripke delineates four possibilities for resolving the puzzle. The one closest to our approach is his "(b) that we do not respect his English utterance (or lack of utterance)". Of course, this formulation is indicative of Kripke's failure to distinguish between beliefs and belief sentences. For him, either we apply the Biconditional Form of the Disquotational Principle to 'r does not assent to 'Pln'', or else we don't consider this non-assen as reflecting r's doxastic state, in which case we 'don't respect it'. According to the approach we have taken, we very much
Kripke's Belief Puzzle

respect this non-assent, and do not disregard it at all: we take it as fully reliable in indicating that 'PIn' is not a belief of r. We part ways with Kripke in that we do not have to choose between concluding that r does not believe that London is pretty, and between disregarding this non-assent. It is simply that in the case in question this non-assent, though reflecting the lack of the corresponding belief, does not entail that r does not believe that London is pretty. The implausibility of disregarding, or not respecting, this non-assent is thus avoided, without affecting the viability of rejecting the conclusion that r does not believe that London is pretty.

However, by using the reporting construction 'r believes that p' we seem to be giving an incomplete description of the situation involved in the contradiction version. 'r believes that p' is true in case some adequate paraphrase of 'p' is a belief of r, and its negation is true in case no such adequate paraphrase is a belief of r. But in cases such as the ones brought to light by the contradiction version he suspends judgment with respect to 'PIn' (in addition to assenting to 'Jlr'), although he understands it in a standard way. We lack a way of reporting this because the issue of whether r believes that London is pretty is being pre-empted by r having another belief (which is an adequate paraphrase of 'PIn'). It seems that our only way of reporting a situation in which r suspends judgment regarding 'PIn' in the indirect mode is to say that r does not believe that London is pretty. But this would be false in case there is an adequate paraphrase of 'PIn' which is a belief of r, as in our case. But even when there is no such paraphrase, saying that r does not believe that London is pretty would not adequately report the fact that r here suspends judgment regarding 'PIn', rather than, by contrast, has '¬PIn' as a belief, which he indeed does not have here. In reporting the case in the inconsistency version of the puzzle we were able to take into account r's assent to '¬PIn' by resorting to (17). But no such comparable reporting mode is available here. Nor would we be helped by correctly saying that r does not believe that London is not pretty. Even though by saying this we would be excluding a situation of the type in the inconsistency version, we would still fail to report suspension of judgment (as contrasted with a case in which an adequate paraphrase of 'PIn' is a belief of r in a language in which he understands it correctly). To say that r does not believe that London is pretty and does not believe that London is not pretty would be true only in case of suspension of judgment in all the languages r understands, and will not be true in the situation envisaged in the contradiction version, in which there is a suspension of judgment in one language but not in another. All this is not to say that we should reconsider our analysis of belief sentences: they have, I believe, been analyzed here correctly. It is just that the belief-sentence construction is not versatile enough to allow for a sufficiently full de-
scription of certain situations.

It therefore seems warranted to introduce another propositional attitude—r suspends-judgment that p.[49] Now we do not need a new construction to say that r suspends judgment with respect to (the sentence) 'p'—i.e., that neither 'p' nor '~p' is a belief of r; but we may want a reporting construction, not a displaying construction.[50] We can thus define:

\[ \text{Suspend-Judgment} \]
\[ \text{Def.:} \quad r \text{ suspends-judgment that } p \text{ iff for some } q', \text{ such that } \text{ap}(q', p'), \text{ neither } q' \text{ nor } ~q' \text{ are beliefs of } r, \text{ though } r \text{ uses the items in } q' \text{ in a normal way in the language to which } q' \text{ belongs.} \]

(Of course, this construction appears here in a de dicto form.)

This will allow us to describe the situation under discussion in the following way. In the contradiction version of the puzzle, when r suspends judgment with respect to 'Pln', we shall be able to say:

(23) r believes that London is pretty and r suspends-judgment that London is pretty.

(23), of course, is not at all contradictory, even though:

(24) 'Pln' is a belief of r and r suspends-judgment with respect to 'Pln'

is a contradiction. Thus, (23) is true in a case of two adequate paraphrases, 'p' and 'q', of 'London is pretty' (such that r is not aware that ap(p', q'), to preserve his consistency and standard use), where 'p' is a belief of r, but neither 'q' nor '~q' are beliefs of r (though r uses the items in 'q' as well as in 'p' in a standard way in the languages to which they belong). Nor, of course, does (23) imply that r is inconsistent: (17) does not, and so, a fortiori neither does (23).

Thus, prior to the introduction of this new construction, it could be argued against us that in the situation which gives rise to the contradiction version (suspension of judgment regarding 'Pln' and assent to 'Jlr') all we could do to describe the situation on the reporting level is to say that r believes that London is pretty, thereby respecting r's assent to 'Jlr', but ignoring his non-assent to 'Pln'. But this fact does not tell against our analysis of belief-sentences: it simply calls for the introduction of a new construction that will enable us to take account of the situation on the reporting level. The suspend-judgment propositional attitude is such a new construction. It allows us now to describe r's suspension of judgment regarding
KRIPKE'S BELIEF PUZZLE

'Pln' together with his assent to 'Jlr' by (23).

It should be noticed that the move made here in no way ignores Kripke's request that we answer the question he posed in the terminology in which he posed it. We have done so. We have fully answered the question of whether \( r \) believes that London is pretty in the case described, and have answered it affirmatively, claiming that our answer is anchored in the nature of the belief construction, as reflected in our analysis of that construction. We have only added another construction in a way which is independent of our treatment of Kripke's puzzle. This addition is not required for the resolution of the puzzle. The use of this added terminology is not at the expense of providing an answer for the puzzle in Kripke's terms. The suspend-judgment construction is merely a convenient and useful construction to have in addition to the belief construction in order to report such a situation more accurately.

Kripke summarizes the puzzle as follows: "... the present puzzle presents us with ... a challenge to formulate an acceptable set of principles that does not lead to paradox, is intuitively sound, and supports the inferences we usually make."[52] I have attempted in this paper to meet this challenge. At the core of our proposal lies the distinction between beliefs, conceived as symbolic (linguistic) representations, and belief sentences, which are analyzed in terms of beliefs, and which describe beliefs rather than display them. It is primarily the level of beliefs that is attached to verbal dispositions, and it is primarily at that level that the notion of the consistency of the believer belongs.[53]

SECTION VII: ANALOGY WITH THE HESPERUS-PHOSPHORUS CASE

It would be instructive to notice the analogy between Kripke's puzzle and the Hesperus-Phosphorus case (taken as a form of Quine's Orlocut puzzle[54]), and the way they can both be handled. The Hesperus-Phosphorus puzzle runs as follows:

A believer \( r \) was accustomed to see a certain celestial body in the mornings at a certain place in the sky, which he came to know as Phosphorus, and believed it to be large, bright, etc. But he suspends judgment as to the question of whether it was a star (rather than a planet); though he considered the question, he did not believe it to be a star, nor did he believe it not to be a star.

On other occasions he became aware of a certain celestial body, which he came to know as Hesperus, and which appeared regularly in the evenings in the western sky. He came to believe it to be
small, dim, etc., and in particular, came to believe it to be a star. Unbeknownst to $r$, Hesperus and Phosphorus were the same planet (Venus).

But now $r$ is in a strange situation vis-a-vis Venus. It seems that, not having changed his beliefs concerning Phosphorus, which resulted from his experience in the mornings, he still does not believe it to be a star; yet, via his acquaintance in the evening hours with what he takes to be a small and dim celestial body, he does believe it to be a star. This seems to be a straightforward contradiction.

The handling of this case relies primarily on the exportation inference. The construction '$r$ believes $'F'$ of $a'$[55] indicates a belief-relation between the cognizer $r$, the predicative expression $'F'$ and the object $a$, a relation which would make '$r$ believes of $a$ that it is $F'$ true.[56] We shall symbolize this construction as $R_F^r F' a'$ (the superscript $'r'$ indicates $'r$').

The primary working assumption for this case is that a valid exportation inference connects '$Fa'$ is a belief of $r'$ with the de re $R_F^r F' a'$. The connecting premise relies on what I have proposed to call[57] the R-function $R_F^r (a', 'Fa')$, a function whose arguments are the singular term $a'$ and the belief $'Fa'$ in which it occurs (for the believer $r$), and whose value is the object which is to be the referent of the singular term $'a'$ in this belief (for $r$). For instance, consider the famous case in which $r$ stands at a party, believing wrongly that the man in the corner is drinking a martini—in fact he drinks water—and asserts: The man drinking a martini is tall. His referent here by the phrase 'The man drinking a martini' is the man in the corner, not the man who in fact drinks a martini (assuming there is one such man at the party). Thus, in this example:

$R_F^r (\text{the man drinking a martini'}, \ 'tall (the man drinking a martini')} = \text{the man in the corner}.$

Now valid exportation has the following form:

A. $'Fa'$ is a belief of $r$

$R_F^r (a', 'Fa') = a$

$\frac{B_E^r F' a'}{B_E^r F' a'}$.

More generally, it has the form:

B. $'Fa'$ is a belief of $r$

$R_F^r (a', 'Fa') = b$
Kripke's Belief Puzzle

for cases in which the referent for 'a' in r's belief 'Fa' is not (as it need not be) the object a (if there is one).

The problem thus stands as follows: r's suspension of judgment as to whether Phosphorus is a star yields

\( \neg (\neg 'Sp' \text{ is a belief of } r) \)

(\( p = \text{Phosphorus; } S = \text{star} \)). Given that in various beliefs of r's his referent by 'Phosphorus' is Phosphorus, as is clear from the description of the case, it seems that this non-belief of his is also of Phosphorus, and that we should thus be in a position to infer:

\( \neg B^r_F 'S' p. \)

However, his assent to 'Hesperus is a star' yields (h = Hesperus):

\( 'Sh' \text{ is a belief of } r, \)

for which it is the case that

\( B^r_F (h', 'Sh') = h. \)

(27) and (28) together yield, via exportation:

\( B^r_F 'S' h. \)

(29) But since \( h = p \) (= Venus), (26) and (29) yield a contradiction (the position of 'a' in 'B^r_F 'a' being referential).

The catch in this puzzle is that the exportation inference B allows 'B^r_F 'b' to be made true by various beliefs (of r) of the form 'Fa', which fulfill that \( B^r_F (a', 'Fa') = b' \), and not necessarily by 'Fb' being a belief of r. Thus the falsehood of the double-condition ''Fa' is a belief of r and \( B^r_F (a', 'Fa') = b' \), for a given singular term 'a', is not sufficient to establish the falsehood of 'B^r_F 'b'', since the latter could be made true by another double-condition ''Fc' is a belief of r and \( B^r_F (c', 'Fc') = b' \). In our case, such a double-condition is false due to 'Sp' is a belief of r being false. But it is a mistake to infer from that alone that 'B^r_F 'S' p' is false too, which is the transition that leads from (25) to (26), a transition which is therefore unwarranted. Indeed, 'B^r_F 'S' p' is true in our case due to (27) and (28).

The analogy to Kripke's puzzle is quite obvious; just as much as the de re sentence 'r believes 'F' of a', i.e., 'B^r_F 'a'', can be made true via 'Fb' being a belief of r plus the auxiliary condition 'R^r_F (b', 'Fb') = a', in the case of
Kripke's puzzle the de dicto reading of 'r believes that a is F' can be made true by 'Gb' being a belief of r plus the auxiliary condition that ap('Gb','Fa'). Normally, when the referent of r by 'a' is a, then indeed, if 'Fa' is a belief of r' is false, so would be 'r believes F' of a'. Analogously, normally when r's use of the terms involved is standard and 'Fa' is a belief of r' is false, so would be 'r believes that a is F'. But not always. And just as much as 'Bp'Fa' does not logically follow from the falsehood of 'Fa' is a belief of r', even when the referent of 'a' by r is a (although the falsehood of the latter constitutes good evidence for 'Bp'Fa'), so doesn't 'r believes that Fa' logically follow from the falsehood of 'Fa' is a belief of r' (with standard use), although making such an inference would lead one astray only in relatively recondite cases. In both cases the epistemic facts belong primarily to the level of r's beliefs. In both cases we possess linguistic constructions for describing beliefs without displaying them: in one case by specifying the predicate involved and the object which constitutes the referent in question, via the construction 'r believes 'F' of a'[59], and in the other case by describing the belief via adequate paraphrases, by the entrenched construction 'r believes that a is F' (read de dicto). But by so describing beliefs or by so asserting that there are such-and-such beliefs, one does not necessarily specify a unique belief, but rather asserts that there are beliefs of a certain sort. And, therefore, in both cases, the failure of a given candidate to be a belief of the sort specified does not imply that there is no other belief of that sort. Ignoring this point underlies the fallacious reasonings that lead to the falsehood of 'Bp'Fa' and of 'r believes that Fa' (read de dicto) on the basis of the failure of a particular belief-candidate to yield them respectively via exportation or via the belief inferences (inferences I and II. The latter fallacy is exemplified in XIV and XIX, section IV). Thus, in the Hesperus-Phosphorus case, 'r believes 'Star' of Phosphorus' was fallaciously judged false on the basis of the falsehood of 'Sp' is a belief of r' (even though the referent of r by 'p' was indeed Phosphorus), while in the 'Londres' case 'r believes that London is pretty' (read de dicto) was fallaciously judged false on the basis of the falsehood of 'Pn' is a belief of r'. The similarity between the fallacies in the two cases is thus obvious, though one involved a relation between the non-possession by r of a certain belief and a de re construction, and the other--between the non-possesion of a certain belief and a de dicto construction, thus involving distinct sorts of inferences. We have just seen that in both cases the fallacious nature of the inferences involved lies in ignoring that, given the falsehood of 'Fa' is a belief of r', there might still be another belief of r 'Fb', such that 'Fb' is an adequate paraphrase of 'Fa' (in the de dicto case), or such that Rr('b','Fb') = a (in the de re case). A situation of this sort requires that r possesses two singular terms, 'a' and 'b' , such that
KRIPKE'S BELIEF PUZZLE

Although $r$ is not aware of it, in the one case, or that the referents for both '$a$' and '$b$' by '$r$' is the same, though $r$ is not aware of it, in the other. It is fairly easy to realize that the latter (the de re) sort of case is perfectly possible for a consistent believer whose use of the linguistic items involved is standard; it takes some more elaboration to establish the same point in the first (the de dicto) case.

In section VI we have pointed out a certain incompleteness in describing the situation typical to the contradiction version by merely using the de dicto reporting construction and sanctioning '$r$ believes that $Pmn$'. A similar dissatisfaction can be expressed concerning the Hesperus-Phosphorus case against the use of the de re reporting construction '$r$ believes '$F' of $a$'. Having told the whole story as in the beginning of this section, all we seem to be able to do by way of describing the situation in the de re mode is to say that $r$ believes 'Star' of Venus, without, on this level, reflecting the fact that $r$ suspends judgment with respect to 'Hesperus is a star', even though $r$ refers to Venus via '$h$' (in some belief of his). Let us abbreviate the latter (that $R_f('Gh','h') = \text{Venus}$, for some '$G$'), as $R(r,'h',v)$, which basically means that '$h$' serves as a vehicle of $r$ for referring to $v$ ($v = \text{Venus}$). To remedy this aspect of the situation presented in section VI, we have introduced a de dicto reporting construction '$r$ suspends judgment that $P$'. We can similarly help things here if we introduce the de re reading of this construction, as follows:

\[
\text{Suspend-Judgment (de re Reading): } R(r,'b','a').[60]
\]

This construction is of course de re (with '$a$' occurring in a referential position). It conveys that $r$ suspends judgment as to whether '$F$' is true of $a$. Having familiarized ourselves with this construction, we shall now be able to accept that:

\[
(29) \quad r \text{ suspends 'Star' of Hesperus, and } r \text{ suspends 'Star' of Hesperus as correctly reporting the situation under discussion, given our story in its entirety. (29) will be true in case, for some singular term '$b$'. 'Sb' is a belief of $r$ such that $R(r,'b','h)$, while for another singular term '$c$', neither 'Sc' nor '-Sc' is a belief of $r$, although $R(r,'c','h')$. (r's logical acumen could still be faultless as long as he does not take '$c$' and '$b$' to be co-referential).
\]

The above charge, that '$r$ believes 'Star' of Hesperus' does not fully describe the situation at hand, is thus jus-
Prior to the introduction of this new construction it could be argued against us[61] that we cannot distinguish on the reporting de re level between a case such as the Hesperus-Phosphorus case described in the beginning of this section, and a case in which r indeed satisfied the second part of our story concerning 'Hesperus' but not the first part concerning 'Phosphorus'; that is, a case in which r would have 'Hesperus is a star' as a belief, thereby referring to Venus, without ever having heard the name 'Phosphorus', and without having any singular term 'b' by which he refers to Venus without having the belief 'b is Hesperus'. Just to say 'r believes 'Star' of Hesperus' would equally apply to both cases. Of course, even in the case just introduced r would suspend judgment with respect to 'Phosphorus is a Star' (he has never heard, in this case, the word 'Phosphorus'); but, on the reporting de re level, without the introduction of our new construction, we would have no way of distinguishing the two cases, which are worth being distinguished.

But the construction 'r suspends 'F' of a' will now allow us to do just that, in a de re construction. (29) will be true in the Hesperus-Phosphorus case as described above, but will not be true in the modified version alluded to in the previous paragraph, since in this case 'K(r,'p',h)' is false, while being a necessary condition for the truth of 'r suspends 'Star' of Hesperus' to be made true on the basis of both 'Sp' and '-Sp' not being beliefs of r.
FOOTNOTES

1. I wish to thank Gilead Bar-Elli, Yael Cohen, Gideon Makin and David Siegel for stimulating discussions of earlier drafts of this paper.


4. Kripke, 258.


6. Kripke, 249.

7. Kripke, 250.

8. In the course of this essay we shall not need to, nor shall we, consider believers who do not possess these faculties.

9. Thus, I excuse myself from discussing the notion of belief for prelinguistic children, animals, etc., even though I would hold that in every case a belief is a symbolic (if not linguistic) representation. Since our discussion here will be limited to beliefs the possession of which is indicated by verbal dispositions, I can safely ignore in the present discussion cases of beliefs which might be taken to constitute symbolic, but not necessarily linguistic, representations. For more along similar lines, see D.M. Armstrong, Belief, Truth and Knowledge (Cambridge University Press, 1973), Ch. 3; G. Harman, Thought (Princeton University Press, 1973), Ch. 4, Sec. 2; and H. Field, "Mental Representations", Erkenntnis, 13, 1, (1978), 1-61.


11. At time $t$, of course. I shall omit obvious time relativizations in the sequel. For further discussion concerning this notion of belief and the role it plays in exportation, cf. my "Quine and Modalities de re: A Way Out?", the Journal of Philosophy, LXXIX, 6, June 1982, 295-328.

12. This contrast is dealt with at greater length in part II of this essay (unpublished).

13. According to my conception, beliefs de re (to be distinguished from belief sentences read de re) form a subclass of the class of beliefs, whose special feature is the existence of reference to a certain object by a singular term in the belief, which invokes a dependence on facts outside the purely internal realm of $r$'s belief-world;
14. Through the first part of this essay I shall confine myself to cases where dispositions to assent or dispositions not to assent to a given sentence are uniform, i.e., independent of the context of questioning (as long as it is an 'appropriate' context of questioning), which will assure us that a disposition not to assent to \( p \) in one context of questioning implies that there is no 'appropriate' context of questioning in which a disposition to assent would be forthcoming (at the same time, of course), and thus that \( p \) is not a belief of \( r \). Such a lack of dependence on the context of questioning characterizes the 'Londres' case. Due to limitations of space, I do not discuss in this paper what happens when this uniformity breaks down; I do so in part II of this essay (unpublished).

15. Sincerity need not lead us circularly back to the notion of belief (as R. Chisholm held: cf. his "Sentences about Believing", in A. Marras (ed.) Intentionality, Minds and Language, (University of Illinois Press, 1972)) if it is taken, for instance, to be a desire to tell the truth.

16. In the construction \( r \) believes that \( p \) I shall always assume that \( p \) is a well-formed English sentence; otherwise, the construction is not a well-formed sentence.

17. This restriction is dealt with in greater detail in part II of this essay.

18. Kripke, 249.

19. I allow myself a liberal use of regular quotes even when corner-quotes or the like are called for, trusting that this will generate no confusion.

20. Again, when \( p \) is an English sentence, which I will normally not repeat henceforth. Issues related to this analysis will be further discussed below in this section, and in part II of this essay. Let me also repeat that the \( \text{PB}_r.p \)' def. is offered under the proviso, which we have adopted for this part of the essay, that \( r \) uses the items in the sentence \( \delta \), which is a belief of his, in a standard way.

21. Notice though that it seems that adequate translation often connects two items which are somewhat less than synonymous. Thus, it seems that adequate paraphrase is somewhat looser than the notion of synonymy. For that matter, the notion of adequate paraphrase might also be less susceptible to some of the criticisms launched against the notion of synonymy.

22. As mentioned before, in my view, all there is, in so far as beliefs are concerned, are beliefs conceived as linguistic (or symbolic) representations (that is, de dicto beliefs); beliefs de re are a subclass of these, which maintain certain causal and justificatory relations. The de re construction of the form \( r \) believes of \( a \) that it is \( F \) has the force of asserting the existence of a belief of \( r \) of a
certain sort, classified via the object which is the referent of the singular term in the belief, and the class of descriptive phrases which are adequate paraphrases of \( \mathcal{L} \) (cf. fn. 56 below).

According to the conception I develop in my Reference and Knowledge, for each believer there is a special class of definite descriptions, called strict anchors, through which the reference relation is secured. Beliefs which include these terms will thereby be de re, and so will other beliefs involving other singular terms which are co-clus-teral with a strict anchor and maintain certain parasitic connections to a belief which includes a strict anchor. Roughly speaking, to be a strict anchor, a definite description \( \lambda x H x \) must satisfy the condition that the believer (latently) knows (de dicto) that \( \lambda x H x \) being \( H \) is a cause of his having some belief involving \( \lambda x H x \). This condition thus confers both causal and justificational components on the reference relation. The reference relation is thus not a purely causal, or naturalistic, relation, nor does it involve any intentional element whatsoever. That the reference relation involves a justificational component provides a basis for arguing that knowledge must also involve justification (and similarly for the rest of the propositional attitudes).

23. For more on this subject in the context of another type of construction—the counterfactual construction—cf. my book A Theory of Counterfactuals, (Indianapolis: Hackett, 1984), chapter 2, sections VII and VIII.

24. See below, section V.

25. I go beyond this limitation in part II of this essay.

26. This limitation too is discussed further in part II of this essay.

27. \( \sim p \) is a place-holder for an English sentence which is the negation of the English sentence \( p \).

28. In particular, the limitation of the uniformity of verbal dispositions.

29. We shall further defend the feasibility of taking this feature as a part of the set-up of the example below in this section.

30. In a context in which political matters are quite insignificant, one may use 'Londres' in a perfectly standard way, without realizing it names a capital, e.g., if one has no opinion what the capital is. The point is not crucial for our example, however, since the puzzle can be generated with respect to other cities which are not capitals, and therefore there is no need to dwell on it here. Also, one may assume, \( r \) is aware that 'Londres' names a capital, but is not aware that 'London' does. Since he knows 'London' to be the name of the city he lives in, this will clearly not reflect on his normal use of 'London', but only on his ignorance.

31. This point is discussed in greater length in part II of this essay.
32. In this case, r would acquire his reference to London via 'Londres' through the sources from whom he acquired his beliefs concerning 'Londres'; cf. also my (unpublished) manuscript Reference and Knowledge.

33. As Kripke seems to do, thus: "Of course, he has learned of that famous distant city, London (which he of course calls 'Londres') . . ." (242). Notice here the de re use by Kripke, who has previously stated that he exclude de re statements from his concerns in this paper. Kripke must have realized that for r's referring by 'Londres' to London is a necessary condition for his standard use of 'Londres', and it was for that reason that he seems to have felt that he needs the de re mode here.

34. The key issue, of course, is what his beliefs really are. I continue here on the basis of the sketchy account developed in section II. The consequences for the Propositional Theory of Belief are discussed in part II of this essay.

35. Under the assumptions we hold in this part of the essay, i.e., of uniformity of response and standard understanding.

36. 'French' and 'English' are abbreviated as 'F' and 'E' respectively.

37. Of course, it is unrealistic to expect any arbitrary believer to have the notion 'adequate paraphrase' in his vocabulary.

38. Cf., for instance: "According to such a supposition a belief that Hesperus is a planet is a belief that a certain heavenly body, rigidly picked out as seen in the evening . . ." (Kripke, 280); or: "... that he lacks the belief that London is pretty:" (Kripke, 258). Or: "We may give a rough statement of his beliefs. He believes that . . ." (Kripke, 259). Of course, failing to separate the two levels is indicative of the Propositional Theory of Belief, which lurks behind Kripke's formulations. Cf. also part II of this essay.

39. The notion of incongruent beliefs is discussed further in part II of this essay.


41. Kripke, 258; cf. also section I above.

42. Of course, if r knows no English, the invalidity of the principle will be glaring, but Kripke limits himself to an English speaker.

43. Kripke, 259.

44. Kripke, 263.

45. Kripke, 263.
49. Some philosophers have taken the view that in order to provide such a full account we must go beyond the level of propositional attitudes and use constructions which display the beliefs involved by explicitly bringing in a relativization to a mode of presentation (cf. D. Kaplan, "Quantifying In", in *Words and Objections*, D. Davidson and J. Hintikka (eds.) (Dordrecht: Reidel, 1969), 206-42; and S. Schiffer, "Naming and Knowing", in *Contemporary Perspectives in the Philosophy of Language*, P. A. French et al. (eds.), (Minneapolis: University of Minnesota Press, 1977)). If this is taken as a motivation for a different analysis of belief sentences, it is, I believe, too radical a proposal, which breaks the inherent character of propositional attitude constructions, the fundamental function of which is to enable reporting without displaying the beliefs involved. The proposal I make below should, I believe, allow for the descriptive incompleteness noted in the text to be remedied within the sphere of standardly structured propositional attitude idioms. That such a propositional attitude as the one proposed here is not to be found in natural language is all but puzzling: the kind of situations it helps describe more fully is rather rare, and does not therefore justify a separate entrenched construction.

50. Note that my coined phrase 'r suspends-judgment that p' is designed to be akin in form to other propositional-attitude constructions.

51. We still operate, of course, under our general assumption in these sections of uniformity of verbal dispositions. If this assumption is withheld, the above definition should be modified. What happens when this assumption is withheld is more fully discussed in part II of this essay.

52. Kripke, 259.

53. In this paper I have not addressed myself to the Paderewski case, which involves a violation of the uniformity assumptions. I do claim, however, that it can be treated along lines similar to the treatment above. I deal with this and other related issues in part II of this essay.


56. But not vice versa. I take the construction 'r believes of a that it is F' to be true just in case, for some singular form 'b' and
for some predicative expression 'G', such that ap('F', 'G'), 'Gb' is a belief of r, and r refers (by 'b', in his belief 'Gb') to a. Accordingly, r believes 'F' of a iff for some singular term 'b', 'Fb' is a belief of r, and he refers (in it) by 'b' to a. Thus, r may believe of a that it is F without knowing English (in particular, without commanding the expression 'F'); but for him to believe 'F' of a, 'F' must belong to his vocabulary.

57. Cf. my "Quine and Modalities de re: A Way Out". I used there the notation 'IRr('a', 'Fa')', where the 'IR' indicated intended reference. But since the reference relation, as I see it (and as I attempt to analyze it in Reference and Knowledge), is emphatically non-intentional, the notation 'R('a', 'Fa')' seems preferable in preventing confusions and misunderstandings.

58. For further details, see my "Quine and Modalities de re: A Way Out". Obviously, the problem of analyzing the reference of singular terms resides primarily in the analysis of the R-function. Exportation, classically conceived (by Quine), was to connect the de dicto readings and the de re readings of, e.g., belief sentences. In the exportation inferences A. and B., 'Fa' is a belief of r' was used instead of the de dicto reading of 'r believes that p'. But the analysis of belief sentences in section II of this article allows for the extension of inferences A. and B. to the case in which a belief sentence (read de dicto) occurs as the first premise. Similar obvious linkages connect the construction 'r believes 'F' of a' and the de re belief sentence 'r believes of a that it is F'. Due to scope limitations, however, I shall not develop this point further here. Cf. "Quine and Modalities de re: A Way Out?", fn. 24.

59. And more generally in the de re construction 'r believes of a that it is F', for which we need not specify the predicate in a belief which makes this construction true, but only an adequate paraphrase of it.

60. The coining of 'r suspends 'F' of a' may sound somewhat awkward as a semi-English construction—even more so, perhaps, than 'r believes 'F' of a'; but no better candidate occurs to me.

61. This point was made by S. Schiffer in private communication. Cf. also his article "Naming and Knowing".