

Externalism and Inference*

Paul A. Boghossian

1 Introduction

Recent work in the theory of mental content has persuaded a large number of philosophers that many of the propositional attitudes attributed in ordinary speech and in psychology cannot be individuated wholly individualistically, without regard to the physical and/or social environment in which those attitudes are tokened. Even a Cartesian conception of thought could, of course, accommodate the claim that a person's *de re* attitudes might vary while his behavior and behavioral dispositions, his physical acts and states, and his qualitative feels, remained fixed. But *de dicto* attitudes, attitudes attributed with that-clauses forming oblique contexts, were traditionally held to be determined by facts wholly internal to the thinker whose attitudes they were. No more:

*This is a slightly revised text of a talk given at the Third SOFIA Conference, on Epistemology, held in Campinas, Brazil in August 1990. It presents, perforce very informally, a small section of a much larger piece of work entitled "Inferential Rationality and the Transparency of Content".

externalism—the view that even commonly ascribed *de dicto* attitudes are individuated in part by reference to external facts and events—is now widely accepted.

Externalism is supported by a series of now-famous thought experiments. Their common strategy is to show that two individuals who are molecular duplicates of each other, may nevertheless think different thoughts if their environments differ from each other in certain specified ways. Thus, Putnam has argued that part of what makes it true that some of my thoughts involve the concept *water*, is that it is typically *in re* H₂O that I token those thoughts; a duplicate of mine who grew up in an indistinguishably similar environment except that in it the liquid that filled the lakes and swimming pools consisted of XYZ and not H₂O would not have the concept *water* but some other concept, *twater*. Similarly, Tyler Burge has argued that part of what makes it true that some of my thoughts involve the concept *arthritis* is that I live in, and defer to, a community in which the word ‘arthritis’ is used in a certain way; a duplicate of mine, who grew up in an indistinguishably similar environment, except that in it the word ‘arthritis’ was so used as to cover all rheumatoid ailments, would not have the concept *arthritis* but some other concept, *tharthritis*.

Finally, and most radically, Donald Davidson has argued that part of what makes it true that I have thoughts *at all* is that my thoughts have a causal history rooted in my external environment; a molecular duplicate of mine who lacked such a causal history would *a fortiori* lack thoughts. Here is Davidson’s example:

Suppose lightning strikes a dead tree in a swamp; I am standing nearby. My body is reduced to its elements, while entirely by coincidence (and out of different molecules) the tree is turned into my physical replica. My replica, The Swampman, moves exactly as I did; according to its nature it departs the swamp, encounters and seems to recognize my friends, and appears to return their greetings in English. It moves into my house and seems to write articles on radical interpretation. No one can tell the difference.

But there *is* a difference. My replica can't recognize my friends; it can't recognize anything, since it never cognized anything in the first place. It can't know my friends' names (though of course it seems to), it can't remember my house. It can't mean what I do by the word 'house', for example, since the sound 'house' it makes was not learned in a context that would give it the right meaning—or any meaning at all. Indeed, I don't see how my replica can be said to mean anything by the sounds it makes, nor to have any thoughts.²

The question I want to look at in this paper is this: To what extent does an externalist conception of mental content threaten our ability to *know* the contents of our thoughts? I shall argue that, in an important sense, externalism is inconsistent with the thesis that we have authoritative first-person knowledge of thought content: in particular, I shall argue, it is inconsistent with the thesis that our thought contents are *epistemically transparent* to us (more on what that means below). I shall further argue that this is true in a sense that falsifies another important and traditionally fundamental view—that we can detect *a priori* whether our inferences are logically valid or not. I shall leave for another occasion the question whether these results reflect badly on epistemic transparency or on externalism.³

2 Reconciling Content–Infallibility with Externalism

Many philosophers have asserted recently that there is no substantial conflict between externalism and first-person knowledge of thought content. The best way to understand why these philosophers fear no conflict between these two doctrines would represent them, I think, as believing a thought that is succinctly expressed by John Heil. Consider

²“Knowing One's Own Mind”, Presidential Address, *Proceedings of the APA*, 1986.

³That it reflects badly on externalism is argued in “Inferential Rationality and the Transparency of Content”.

a mental state M of mine which, according to externalism, has the content that p by virtue of standing in an appropriate relation to some external state of affairs A ; and consider my second-order mental state M' in which I introspect on M . Now, says Heil:

We are supposing that externalism is correct, hence that the content of M' is determined by some state of affairs, A' , that is at least partly distinct from M' . What, now, is to prevent A' from determining an intentional content for M' that *includes* the content of M ? What, for instance, keeps our simplified theory from allowing that a causal relation of a certain sort endows my introspective thought with a content encompassing the content of the thought on which I am introspecting?⁴

How should we understand Heil's thought here? Let's assume, without loss of generality, that we think in a language of thought. (This will make the exposition easier, but is not required for anything that's to follow: similar arguments can be constructed without assuming that the states which possess content properties have syntactic structure.) With this convenient assumption in place, we may say that Peter believes that p just in case Peter has a token of a sentence S , which means that p , in his belief box; and Peter desires that q , just in case Peter has a token of the sentence S' , which means that q , in his desire box; and so on for the other attitudes.

⁴ "Privileged Access", *Mind*, 1988, p.245. See also Donald Davidson:

Showing that there is no conflict [between externalism and first-person knowledge] is basically simple. It depends on realizing that whatever is responsible for the contents of our thoughts, whether known or not, is also responsible for the content of the thought that we have the thought. ("Reply to Burge", typescript, p. 1.)

For some similar ideas see the paper by Burge to which Davidson's is a response — "Individualism and Self-Knowledge", *Journal of Philosophy*, December 1989; and Barry Loewer and Ernest LePore, "Solipsistic Semantics", *Midwest Studies in Philosophy*, X, 1986. There are important differences between these various ideas for reconciling self-knowledge and externalism, but they won't matter for what follows.

It is now easy to explain Heil's thought. For, regardless of whether externalism is true, we can surely count on there being *syntactic* mechanisms which ensure that if a token of the sentence *S* enters the belief box, then *ceteris paribus*, so does a token of the sentence 'I believe *S*'. And so, provided that the embedded and unembedded tokens of *S* may always be counted upon to express the same content, the second order belief will always report correctly on the content of the first-order belief, the externalist nature of belief content individuation notwithstanding. Hence, externalism is fully consistent with content infallibility.

3 Self-Knowledge and Content Transparency

Now, I don't really know that we can always count on the embedded and unembedded occurrences of *S* to express the same content. In fact, as I propose to argue below, there may be special reasons for doubting the reliability of this covariation under externalist assumptions. Secondly, this proposal would appear at best to ensure reliable true belief, not knowledge; and there are tricky questions about justification that need answering.⁵ But I will not be primarily concerned to pursue either of these lines of thought here. What I *will* be concerned to argue is that the assurance that this sort of proposal provides, about the compatibility of externalism with authoritative self-knowledge, is, in a sense to be explained, *hollow*: it carries with it none of the usual *consequences* of first-person authority about thought content.

To see what I mean, let me begin by noting what epistemic abilities *don't* follow from the imputation of the sort of knowledge that the Heil-style proposal ensures. Suppose that, in addition to *S*, there is another sentence *S'* in my language of thought that is synonymous with *S*. Accepting the proposal implies accepting the claim that there is a sense

⁵For a discussion of some of these issues see Anthony Brueckner, "Skepticism about Knowledge of Content", *Mind*, 1989.

in which I may be said to know the content of *S* and to know the content of *S'*; but it doesn't at all follow, from the attribution of the sort of knowledge of content that's on offer, that I would know that the two sentences are synonymous. Suppose, further, that there is another sentence *S''* that is distinct in meaning from *S*. Similarly, accepting the proposal implies accepting the claim that there is a sense in which I may be said to know the meaning of *S* and to know the meaning of *S''*; but it doesn't follow from the sort of knowledge of content that's on offer, that I would know that the two sentences express distinct meanings. In fact, it doesn't even follow, from the sort of knowledge of content that the proposal reconstructs, that I would know whether *two distinct tokens of S* express the same or different meanings. All that the proposal guarantees is that, for a given meaningful token of a given sentence type in my language of thought, I can state its content correctly.

Michael Dummett once wrote:

It is an undeniable feature of the notion of meaning — obscure as that notion is — that meaning is *transparent* in the sense that, if someone attaches a meaning to each of two words, he must know whether these meanings are the same.⁶

I can summarize the points I've been making, then, by saying that, whatever else may be true, the Heil-style proposal does not provide for the *transparency* of thought content, for it does nothing to ensure that if I knew, in the sense it provides for, the contents of two thought tokens, that I would know after sufficient reflection whether those tokens have the same or different contents.

Indeed, the first principal thesis of this paper is that nothing *can* be done to provide for the transparency of content on externalist assumptions, for externalism really is inconsistent with knowledge of content so understood.

Now, the literature already contains a number of examples which are either designed, or which may be adapted, to show

⁶ "Frege's Distinction between Sense and Reference", reprinted in *Truth and Other Enigmas*, Harvard University Press, 1978.

that transparency fails given externalism. They all have the following form:

- [A] A situation in which two tokens of mentaleses belonging to *distinct* syntactic types have the *same* meaning, but the subject is in principle not in a position to notice that they do.⁷

So far as I can tell, however, no one seems to have noticed that externalism allows for transparency to fail in the converse case as well. Thus:

- [B] A situation in which two tokens of mentaleses belonging to the *same* syntactic type have *different* meanings, but the subject is in principle not in a position to notice that they do.

The existence of such converse cases is, I think, very important, even for issues beyond the ones presently at hand. I will use the existence of such cases to argue for the second principal thesis of his paper: namely, that externalism is inconsistent with a very important aspect of our intuitive conception of the mind —namely, with the *a priority of our logical abilities*.

It is common to think that the logical properties of the inferences we engage in must be judgeable purely *a priori*. But I will argue that this may well be false on externalist assumptions, that judging the validity of our inferences may well require —potentially absurdly— empirical investigation.⁸

In what follows, I will first describe how we can get such converse cases on externalist assumptions and, second, how they might be used to threaten the *a priority* of our logical abilities.

⁷See, for example, Brian Loar, “Social Content and Psychological Content”, *Contents of Thought*, ed. Grimm and Merrill, University of Arizona Press, 1986.

⁸That it is absurd is, again, argued in “Inferential Rationality and the Transparency of Content”.

4 The Semantics of Travel

To see how situations of type [B] might arise we need to think—harder perhaps than is common in the literature—about the semantics of so-called “one-world traveling cases”. Suppose that earth and twin earth are part of the actual world and that Peter, a normal, competent adult earthling, is suddenly and unwittingly transported to twin earth. He goes to sleep one night at home and wakes up in twin home in twin bed. He suffers no discernible disruption in the continuity of his mental life. Here on twin earth Peter happily lives out the rest of his days, never discovering the relocation that he has been forced to undergo. How should we think about the semantics of Peter’s thoughts?

Well, one intuition that is shared by practically everyone who has thought about these cases is that, after a while (how long is unclear), tokens of ‘water’ in Peter’s mentalese will cease to mean *water* and will come to mean *twater*. Thus, to quote just one example, Tyler Burge writes:

The thoughts would not switch as one is [quickly] switched from one actual situation to another twin actual situation. The thoughts would switch only if one remained long enough in the other situation to establish environmental relations necessary for new thoughts. So quick switching would not be a case in which thoughts switched. . .

But slow switching could be such a case. [p. 652]

It’s hard to deny, from a pre-theoretical perspective, the force of the intuition that Burge is giving voice to here. You imagine Peter moving to twin earth and staying there. Over time he thinks to himself a variety of thoughts that he would express with sentences like:

Gee, the water is cold today.

Or:

I’m so thirsty I could drink that entire pitcher of water.

And so on. It’s very plausible that these expressions of beliefs about one’s present environment, expressions of current desires and current intentions are about *twater* not water.

Indeed, not only are they plausible in themselves, they are especially plausible for an externalist. To see why, consider what principles of content fixation underlie the standard twin earth cases, and the case of Davidson's Swampman. It is this: the contents of thought tokens of a given syntactic type are determined by whatever environmental property is the typical cause of the perceptions that cause and sustain tokens of that type.⁹ Thus, if the perceptions have as yet no *typical* cause (as in Davidson's newly minted Swampman), it is natural to say that the thoughts have no content (and hence are not really thoughts). And if their typical cause is *water* (as is presumably true in our case), the natural thing to say is that the thoughts are about water. And if their typical cause *changes* (as in Peter the traveller's case), the natural thing to say is that tokens of thoughts of that (syntactic) type switch from being about water to being about twater.

So far, so good. The principles underlying the fixation of semantic properties in the standard examples appear to explain well our intuitions about the semantics of travel. We see why a quick visit to twin earth wouldn't suffice for a switch in thought content; why a long one would; and why we would expect many borderline cases.

But reflection should reveal that these very principles would predict that if there were thoughts whose typical perceptual causes continued to be earthly, in spite of their being tokened on twin earth, then those thoughts would retain their earthly interpretations, even on twin earth.

And that there are such thoughts should be, in a sense, obvious: *memories*, and *beliefs about the past* based upon them, provide ready illustrations. Such thoughts, unlike, for instance, beliefs with undated general contents, thoughts about one's present surroundings, are caused and sustained by *previous* perceptions long gone. In the normal case, they owe little, if anything, to current perceptions and cognitive transactions with one's environment. They would be expected,

⁹I am thinking here primarily of beliefs about natural kinds like *water*. It is not important for my purposes here to argue that this is the way *every* type of mental expression acquires its meaning.

therefore, to retain their earthly interpretations even when tokened on twin earth.

Consider an example. While still on earth, Peter goes hiking in the mountains of northern New Zealand. Here he comes across Lake Taupo and is startled to see the famous tenor Luciano Pavarotti floating on its pristine waters. The tenor, it turns out, has snuck off to this remote spot for a peaceful vacation. They talk amiably for a while and Peter goes off flushed with excitement.

Understandably enough for an opera buff, this experience of Peter's gives rise to many subsequent memories on his part, and to beliefs based upon them. Consider the ones that are tokened while he is still on earth, many years prior to his trip to twin earth. There can be no question about their content. They are about Pavarotti and Lake Taupo and water. In fact, I take it, they would be about Pavarotti even in a possible world in which Pavarotti has a twin living in Los Angeles who Peter doesn't know about but whom he subsequently meets. Even after he does so, the memories of the encounter on Lake Taupo, and the beliefs based upon them, would continue to be about Pavarotti, not twin Pavarotti, because it would continue to be true that they are caused and sustained by perceptions of Pavarotti, not by perceptions of his twin.

Well, some years go by and Peter moves to twin earth and becomes happily ensconced there. Of course, he maintains his interest in opera and so continues to read and hear about his favorite performers. Eventually, some of the tokens of his mental names come to refer to the twin counterparts of the familiar earthly performers: tokens of "Domingo" come to refer to twin Domingo, tokens of "Pavarotti" to twin Pavarotti, and so on.

One day, while reading about twin Pavarotti's latest concert, he is moved to reminisce about the occasion when he saw him swimming in the waters of Lake Taupo. He calls up vivid and accurate representations of the scene. Of course, he takes himself to be remembering scenes involving the singer he is reading about now. But he isn't. His memories, intuitively, are about the earthly Pavarotti, the earthly Lake Taupo, and earthly water, previous perceptions of which are

the sustaining cause of his later ability to recall what the scene looked like and how it felt.

5 Externalism and the a priority of our Logical Abilities

In the situation described, Peter's externally individuated thought tokens are not transparent to him. In particular, Peter's language of thought contains token expressions that possess different semantic values, despite being of the same syntactic type. And yet, clearly, Peter does not know that they do. Tokens of 'Pavarotti', 'water', and 'Lake Taupo', in sentences expressing memories and beliefs about that memorable occasion, will refer to Pavarotti, water and Lake Taupo, respectively; whereas other tokens of that type, in sentences expressing beliefs about his current environment, as in

There is a lot of water around here

or current desires, as in

I want to embrace Pavarotti

will refer to twater and twin Pavarotti. From the inside, however, there will be no indication of this: as far as Peter is concerned, they will appear to express precisely the same contents.

As a consequence, it seems easy to show that Peter will not be able to judge *a priori* the logical properties of his inferences. Thus, Peter might think to himself, reminiscing about his memorable experience:

— Pavarotti once swam in Lake Taupo.

Given that this is about an experience had while on earth, it is about the earthly Pavarotti and the earthly Lake Taupo. He then calls up an experience of yesterday's concert, and thinks:

— The singer I heard yesterday is Pavarotti.

Given that this is a memory of a twearthly experience, it is, as I have argued, about twin Pavarotti. But it should be perfectly clear that Peter will not shrink from using these premises to infer the conclusion (we may imagine that he had not previously put these two items of information together):

- Therefore: the singer I heard yesterday once swam in Lake Taupo.

For to him the inference will appear blindingly valid. But it isn't. The first premise is intuitively about our Luciano. The entirely independent second premise expresses a memory of yesterday's concert and is consequently about twin Luciano. True premises conspire, through a fallacy of equivocation that Peter is in principle not in a position to notice, to produce a false conclusion.

Analogous cases involving practical reasoning can also be devised. Thus, consider the following:

- I want to embrace anyone who once swam in Lake Taupo
- Pavarotti once swam in Lake Taupo
- So, I shall set off to embrace Pavarotti

The first premise, by virtue of expressing a present want, is about twearth and twin Lake Taupo. Whereas the second premise, by virtue of expressing an earthly memory is, as I have argued, intuitively about earthly Pavarotti and earthly Lake Taupo. Finally, the conclusion, by virtue of expressing a present intention, is about twin Pavarotti; but, however it is taken, it doesn't follow from the believed premises.

In travellers like Peter, both the relationship between derivability and validity and the transparency of thought content break down, with the result that inferences that look to be, "from the inside", valid, aren't. And the thesis of the *a priori* of logical abilities is shown, thereby, to be inconsistent with externalist assumptions.¹⁰

¹⁰The terminology here should not mislead. Obviously, I am not saying that externalism undermines our ability to tell *a priori* what form an argument would have to have in order to be logically valid; I am arguing that externalism undermines our ability to tell *a priori* whether any particular inference of ours satisfies one of those forms.

6 Peter and Pierre

The case of Peter the traveller is just the converse of a case made famous by Kripke: that of Pierre. Pierre, as you may recall, has two token expressions in his language of thought —‘Londres’ and ‘London’— which, despite belonging to two distinct syntactic types, express the same meaning; however, Pierre is in principle not in a position to introspect that they do. Peter has two tokens in his language of thought —‘Pavarotti’ and ‘Pavarotti’, if you will— which despite being of the same syntactic type, express different meanings; however, Peter is in principle not in a position to introspect that they do.

I have just claimed that Peter illustrates the inconsistency of externalism with the intuitive *a priori* of our logical abilities, by instantiating a case in which external circumstances beyond his knowledge or control frustrate the logical validity of his inferences. What is the relation between Peter and Pierre?

The details of the case of Pierre are well enough known that I need not rehearse them here. A plausible story culminates in Pierre having two expression tokens —‘Londres’ and ‘London’— belonging to two distinct syntactic types, but referring to precisely the same city, without his realizing that they do. In one scenario, he exhibits no reluctance to believe both that:

Londres is pretty

and that

London is not pretty

Yet he is unable to infer an inconsistency from his separate beliefs that London is pretty and that London is not pretty.

In a somewhat modified scenario, Pierre believes more cautiously only that

Londres is pretty, if New York is

and

London is not pretty

And yet he seems unable to infer that New York is not pretty from his separate beliefs that London is not pretty and that Londres is pretty, if New York is.

Now Kripke, of course, intended this case to show that

when we enter into the area exemplified by . . . Pierre, we enter into an area where our normal practices of interpretation and attribution of belief are subjected to the greatest possible strain, perhaps to the point of breakdown.(pp. 269-70)

The overall polemical idea was to protect Millian theories of names from the charge that they generate absurd belief ascriptions, by showing that parallel absurdities can be generated simply by relying on ordinary principles of belief reportage and without relying on a principle of substitutivity. I think in fact that a careful examination of the argument will reveal that *externalism* is also being crucially presupposed. This is most readily visible in the discussion of general terms:

In "Naming and Necessity" I suggested that there are important analogies between proper names and natural kind terms, and it seems to me that the present puzzle is one instance where the analogy will hold. . . . Not that the puzzle extends to all translations from English to French. At the moment, at least, it seems to me that Pierre, if he learns English and French separately, without learning a translation manual between them, must conclude, if he reflects enough, that 'doctor' and 'medicin' and 'heureux' and 'happy' are synonymous, or at any rate, coextensive; any potential paradox of the present kind for these word pairs is thus blocked. But what about 'lapin' and 'rabbit' or 'beech' and 'hetre'? We may suppose that Pierre is neither a zoologist nor a botanist. He has learned each language in its own country and the examples he has been shown to illustrate 'les lapins' and 'rabbits' 'beeches' and 'les hetres' are distinct. (This is especially plausible if, as Putnam supposes, an English speaker. . . who is not a botanist may use 'beech' and 'elm' with their normal (distinct) meanings, even though he cannot himself distinguish the two trees. Pierre may quite plausibly be supposed to wonder whether the trees which in France he called 'les hetres' were beeches or elms, even

though as a speaker of French he satisfies all normal criteria for using 'les hetres' normally....) Once Pierre is in such a situation, paradoxes analogous to the one about London obviously can arise for rabbits and beeches. Pierre could affirm a French statement with 'lapin', but deny its English translation with 'rabbit'. As above, we are hard pressed to say what Pierre believes. (pp. 264-265)

For pairs of non-natural kind terms, says Kripke —for example, for 'heureux' and 'happy'— and even if Pierre learns them separately, still he must conclude, if he reflects enough, that they are synonymous. If he attaches the same meaning to them, then he could not fail to know that he does, if he gives it enough thought.

This need not be true, however —so the argument proceeds— in the case of natural kind terms. In the case of such terms, it is easy to describe word pairs which in a given person's idiolect express exactly the same natural kind, even though he is not necessarily in a position to discern that fact.

Well, what explains the asymmetry? Why should facts about the synonymy of non-natural kind terms be transparent, when facts about the synonymy of natural kind terms aren't?

The answer seems plain. Kripke is assuming that, although we may think of the meaning of a non-natural kind term as determined by facts that are purely internal to a speaker's psychology, and hence as introspectively accessible to him, we have to think of the meaning of natural kind terms as determined in part by causal facts completely external to the speaker's psychology, and, hence, as not necessarily available to introspection.¹¹ In such a case, consequently, two expressions may end up picking out the same natural kind in a given speaker's idiolect; plausibly, however, he need not know that they do.

¹¹Tyler Burge, of course, has subsequently argued that the meaning even of non-natural kind terms is determined in part by facts outside the head. Kripke appears not to be accepting this further thesis. As we shall see, it would have made no difference to any point presently at issue if he had.

The correct moral of all this would appear to be, however, not that there is an irresolvable puzzle about belief, but rather that externalism is inconsistent with the transparency of thought content.

However, the case of Pierre does not allow us to demonstrate the inconsistency of externalism with the intuitive *a priority* of our logical abilities, in the way that Peter does. The reason is that the belief that ‘Londres is pretty’ does not *logically* contradict the belief that ‘London is not pretty’; and the belief that ‘Londres is pretty, if New York is’ along with the belief that ‘New York is pretty’ do not *logically* entail the belief that ‘London is pretty’. Some loose claims to the contrary notwithstanding, Pierre isn’t *logically* deficient.

7 The *a priority* of our Logical Abilities and the idea of Rationalization

What do we need the thesis of the *a priority* of logic for? There are, I think, several answers, none of which I can go into here.¹² Here let me very briefly indicate an intuitive consideration: the thesis plays an important role in fixing what it is for our propositional attitudes to rationalize our practical and theoretical conclusions.

Suppose Peter looks at his watch and we wish to know why. What we typically want in such a case is a *rationalizing* explanation of his action: an explanation that displays Peter’s *reasons* for looking at his watch. Such an explanation would show why, in light of the way Peter conceives the world to be, and in light of his wants, the performance of that action made sense for him.

Now, consider Peter the traveller again, musing to himself just prior to a recital he plans to attend at (twin) Carnegie Hall:

— Pavarotti once swam in Lake Taupo

¹²For lengthy discussion see “Inferential Rationality and the Transparency of Content”.

- I want to embrace anyone who swam in Lake Taupo
- So, I want to embrace Pavarotti

And off he goes.

It seems to me that if I understand the idea of reasons rationalizing intentions at all —of propositional attitudes explaining why a certain action looks to an agent to make sense from his point of view— that *this* is a case in which reasons rationalize intentions. His intention to go off to embrace twin Pavarotti *does* make sense from his point of view. It certainly seems wrong to say that his inference provides him with *no reason* to go and embrace twin Pavarotti.

In fact, I am inclined to go further. Consider a possible world in which Peter's life unfolds just as it does in my traveling story (however exactly that is) —except minus the overnight transportation to twin earth. In this counterfactual world, Peter's thoughts would undergo no shift and, hence, would contain no ambiguous mental expressions. The practical inference he performs vis a vis Pavarotti would be perfectly valid and would fully rationalize his behavior, if anything does. Now, it seems to me that there is an immediately recognizable sense in which there can be no difference *in respect of rationality* between this counterfactual Peter and my traveler. It seems implausible in the extreme to say that they differ in their capacity *to reason*.

However, I won't bother to defend this stronger claim at greater length just now. All I need for present purposes is agreement that it would be wrong to say that Peter the traveler *has no cogent reason whatever* for his Pavarotti-directed behavior. Yet, if all we had to work with were his mental states widely individuated, that is what we would have to conclude. For given only the wide contents, Peter's inference reads like this:

- *Pavarotti(1) once swam in Lake Taupo(1)*
- *I want to embrace anyone who once swam in Lake Taupo(2)*
- *So, I want to embrace Pavarotti(2)*

Come again?!

Now, I don't pretend to have disposed of all the lines of resistance that might occur to someone to the conclusion I've been urging, so what follows has the status of a mere bold conjecture; still, I guess I think it's pretty plausible as bold conjectures go. It is this: Once the possibility of unambiguous mental tokens of ambiguous mental expression types is conceded, there will be no avoiding the imputation of subjectively undetectable logically invalid inferences, at least if we are confined to using wide contents. If, then, it is also true that there is an important sense in which Peter the traveler's behavior *makes sense from his point of view*, we would appear to have here an argument for the existence of a level of intentional description which conserves that sense.¹³

¹³I am grateful to the organizers of the conference for their hospitality and especially to Enrique Villanueva and Lourdes Valdivia. For helpful reactions I wish to thank Paul Horwich, John Biro, Jonathan Dancy, Jennifer Church, Jim Higginbotham, Richard Foley, Stephen Stich, Richard Fumerton, Bill Ramsey, Stephen Yablo and David Velleman.

Special thanks to Stephen Schiffer for his very useful "Comments" and for numerous conversations on the topics covered in this paper.