

On Always Being Wrong

Peter van Inwagen

I believe that truth is radically nonepistemic.¹ In my view, propositions that are not *about* the beliefs (or, more generally, the mental states) of rational beings are true or false quite independently of the beliefs (or mental states) of rational beings. Moreover, whatever precisely ‘about’ may mean, there are, *pace* Berkeley and Hegel, many propositions that are in *no* sense “about” the mental states of rational beings. One such is the proposition that Mount Everest is 8,847.7 meters high. This proposition, as I see matters, is true and would have been true if the metric system had never been devised, if theodolites had never been invented, if Mount Everest had been named Mount Lambton, if Mount Everest had never been named at all, if there had never been such a language as English, if sapient beings had never evolved on the earth, and if life had never existed anywhere in the universe.² (Actually, I am not quite sure about this last statement. For all I know, the existence of life in a universe like this one is so very nearly inevitable that the closest possible worlds in which there is no life whatever are worlds in which the large-scale structure of the universe is different—in which case, all bets about the height of Mount Everest are off. This sort of consideration obviously applies not only to the existence of life, but also to the existence of beliefs and other mental states: The truth-values of various propositions that are not *about* mental states may, nevertheless, be causally connected with the existence of mental states. For example, the proposition that there have been hydrogen fusion explosions near the surface of the earth is not about mental states, but it is true, and would have been false if there had never been any mental states. One who holds that truth is radically nonepistemic, therefore, must say something like this: Although mental states may be among the *efficient* causes of the truth-value of a proposition that is not about mental states, mental states are never among the *formal* causes of the truth-value of a proposition that is not about mental states. That is to say: The existence of mental states has nothing to do with

what it is for a proposition to be true – or not unless that proposition is about mental states.)

If life had never evolved on earth, then, or so it seems reasonable to suppose, no one would ever have thought of the proposition that Mount Everest is 8,847.7 meters high.³ Nevertheless, this proposition would have existed and would have been true. And who can doubt that there are in actuality many propositions that no one *has* ever thought of but which are true? It is, of course, impossible to give an example of one. But we can easily think of some near misses. If I had not just decided to think of a bizarre proposition – if I had sat down to write a half hour earlier or later, with a different set of associations in my mind – no one in the whole history and future of the universe would ever have thought of the proposition that fewer than one hundred and eight Paraguayans have flown to Mercury on broomsticks made of *lignum vitae*. But this proposition would have existed just the same, and it would have been true. The number of such propositions – propositions I might very well have thought of just now but didn't, and which no one else will ever think of, and which are true – is, I imagine, greater than the number of elementary particles. And, of course, there are vast infinities, perhaps there is a proper class, of propositions that no one will ever think of because they are too complicated. And many of these propositions will be true. (Roughly half of them, to borrow a joke of Plantinga's.) For example, for each well-ordering relation on the real numbers, there is the proposition that that relation exists. There are 2-to-the-*c* such propositions, all of them true, and no one will ever think of any given one of them. And there are other propositions that no one could ever think of that are much "nearer misses" than these. If the mean IQ of our species had been 60 rather than 100 (and if there had been no other sapient species), no one would ever have thought of the proposition that the existence of atomic nuclei depends on the fact that the bosons that transmit the color-force themselves bear color-charges. But this proposition would nevertheless have existed and it would have been true. (At any rate, it would have been true if it *is* true.) And I find it very hard to believe that there are no propositions that must remain unknown to us because of our intellectual incapacities, but which could be known to finite but more intelligent beings.⁴ If the mean IQ of our species had been 180 rather than 100, then, doubtless, all manner of really interesting propositions would be known to us that will never in fact be thought of.

Truth, I maintain, is a property of propositions, items that stand to assertion as numbers stand to counting and that stand to declarative sentences as numbers stand to numerals. In saying that this property, truth, is radically nonepistemic, I am not embracing any particular theory about what the possession of this property by a proposition consists in. Most especially, I am not saying that the truth of a proposition consists in its sharing a "structure" with a part of the world, or consists in its "mirroring" or "picturing" reality. I am not saying this because I do not know what it means. (To be precise, I do not know what the words in scare-quotes mean.) And, of course, for the same reason, I do not *deny* that the truth of a proposition consists in its sharing a structure with a part of the world or in its mirroring or picturing real-

ity. Some philosophers seem to believe that the thesis that truth is radically nonepistemic commits its holder to the thesis that the truth of a proposition consists in some sort of "Tractarian" relation between that proposition and some segment of the world. Well, perhaps this is right. Doubtless, many of the philosophical theses I hold commit me to propositions that I do not know about because I cannot understand the available descriptions of these propositions. What I am sure of is that I *do not* understand the Tractarian account of truth; and, therefore, I decline to regard myself as committed to it by my belief that truth is radically nonepistemic or by any other thesis I hold.

The sentence 'Truth is radically nonepistemic' is one way of expressing the thesis that is nowadays called realism or metaphysical realism or external realism. Or perhaps it expresses only a part of or a consequence of "realism." I don't know. I have no very reliable sense of the boundaries of what is at issue in the realism/antirealism debate. Perhaps there is more to "realism" than the thesis that truth is radically nonepistemic.⁵ Perhaps some sort of Tractarian "copy" or "picture" theory of truth is an essential part of realism. Nevertheless, my conviction that truth is radically nonepistemic (understood as I have spelled it out in the preceding paragraphs) would not be accepted by any of those philosophers who describe themselves as antirealists. At least I *think* this is true. Consider the proposition that the existence of atomic nuclei depends on the fact that the bosons that transmit the color-force themselves bear color-charges. If there is an antirealist who believes that this proposition would have existed and would have been true even if there were no minds, I hope that antirealist will write to me and tell me so.

In this paper, I wish to examine and reject a certain argument that (if I understand him rightly) Hilary Putnam has employed to show that truth cannot be radically nonepistemic. This argument is not Putnam's only argument for that conclusion, nor, I believe, is it the argument that he places the most weight on. I do not claim, therefore, to have seriously undermined Putnam's defense of antirealism or to have decisively blocked his attack on realism.

Here are the bare bones of the argument:

- (1) If truth is radically nonepistemic, then it is intelligible to suppose that we are all wrong about almost everything.
- (2) It is not intelligible to suppose that we are all wrong about almost everything

hence, Truth is not radically nonepistemic.

We should "all be wrong about almost everything" if we were in the position Descartes imagines himself to be in in *Meditations*, i ("I shall, therefore, suppose . . . that some evil genius, no less powerful and crafty than he is deceitful, has employed all his energies in deceiving me; I shall suppose that the heavens, the earth . . . and all other external things are only illusions. . . ."), or if we were "brains in vats." I say "almost everything" because one might maintain that even if

I am a brain in a vat, I am right in thinking that I am not now in excruciating pain, and that I believe I am now writing a paper for *Midwest Studies in Philosophy*, and that $1 + 1 = 2$.⁶ (And, for that matter, that there are brains and vats.) But I am wrong in thinking that I have limbs or that there is such a country as Austria or that Sirhan Sirhan shot Robert Kennedy: Almost all the statements I (think I) make or (think I) hear others making are false.

I concede that Putnam does not explicitly formulate this argument, or any similar argument. But if he does not mean to be defending this argument, or some argument very much like it, in the first chapter of *Reason, Truth and History*,⁷ then I do not see the point of that chapter.

Most of that part of the first chapter of *Reason, Truth and History* that I take to constitute a presentation of the argument whose “bare bones” are laid out above is (read in that way) a defense of its second premise. I cannot find anything that is clearly supposed to be a defense of the first premise or of any similar proposition. And it is the first premise that will be my concern in this paper. I shall, in fact, argue that there is no reason to accept the first premise. But my argument for this conclusion will consist mainly in an attempt to show that there is no reason why someone who believes that truth is radically nonepistemic should not accept Putnam’s argument for the second premise. I will begin, therefore, by examining Putnam’s attempt to show that it is unintelligible to suppose that we are all wrong about almost everything.

My first reading of Putnam’s argument for this conclusion was accompanied by a strong feeling of *déjà vu*. The argument is very nearly identical with arguments deployed by two prominent students of Wittgenstein against the intelligibility of two of the great philosophical tales of global deception. The tales are Descartes’s “evil genius” story and Russell’s celebrated “hypothesis” that the world came into existence five minutes ago, “with a population that ‘remembered’ a wholly unreal past.”⁸ The prominent students of Wittgenstein are O. K. Bouwsma (“Descartes’ Evil Genius”)⁹ and Norman Malcolm (“Memory and the Past”).¹⁰ I used to be fascinated by these arguments and to believe that they were both deep and probably correct. I am now less sure about their correctness; but never, during my many years of thinking that they were probably correct, did it occur to me that these arguments ought to make me the least bit uneasy about my allegiance to a radically nonepistemic conception of truth; and I do not now think that I was missing anything.

Putnam’s version of the argument goes like this. (What follows is a free paraphrase of Putnam’s argument. He should not be held responsible for the details of my exposition – all the more so in that I have not hesitated to add my own “improvements” here and there – or for the absence of various subtle qualifications. The argument that follows is intended more as a reminder to the reader of the *sort* of argument that Putnam presents than as a responsible presentation of that argument.)

Let us consider the following tale of global deception: We are all brains in vats of nutrient fluid, continually supplied with artificially produced stimuli that mimic the stimuli we should be receiving from the external world if we were (or had) “nor-

mally embodied" brains. (These stimuli are supplied by a vast computer: a world-processor.) This tale obviously has no implications as regards the intrinsic qualities of our mental states. Therefore, we are tempted to say, there is no way to tell whether or not this tale is true, and it is perfectly conceivable that we are all wrong about almost everything—for, if the tale were true, we should all be wrong about almost everything.

This temptation should, however, be resisted. Here is why. The intrinsic qualities of a mental state cannot determine what extramental events that mental state represents, or what features that state represents them as having. For example, the intrinsic qualities of a mental state can no more, by themselves, make that state a memory (even a false memory) of once having met Winston Churchill than the intrinsic qualities of a pattern of lines on paper can, by themselves, make that pattern of lines a caricature of Winston Churchill. There is more than one philosophical account of what it is besides the intrinsic qualities of a mental state that is needed in order for a mental state to represent something extramental, but all these accounts agree in this: Whatever is needed for a mental state to represent extramental objects, it is something that could not be possessed by a mental state that belonged to a brain in a vat. As a consequence, when a brain in a vat "produces" a sentence like "There's a tree before me," either he [he?] says nothing at all—or he at any rate does not refer to trees or to any others of the things philosophers call "external objects." Either he speaks no language at all, or he speaks a language that is not ours—he was not brought up in the right linguistic community or even in the right kind of environment to speak *our* language—and, therefore, he certainly could not be referring to *trees*. He could not have been referring to trees for the same reason that a Cro-Magnon man could not, when engaged in some piece of paleolithic fireside chatter, have referred to neutrinos. The causal relations that held between paleolithic utterances or mental states and those elusive particles were too remote, too nearly nonexistent, for those utterances or states to refer to or represent neutrinos; similarly, the causal relations that hold between the "utterances" or mental states of a brain in a vat and external objects are too remote for "vattish" utterances or states to represent external objects. (Remember, we are not talking about a case in which a brain in a vat is supplied with an external scanner that "reports" the physical structures of nearby objects; the stimuli that the brain in our story receives are generated by a computer running a world-processing program. If the external world does happen to contain trees as well as vats and computers, that is just an accident.)

A brain in a vat, therefore, either has no language at all or has a language that is not ours. Even if he does have a language, it does not follow that he is wrong when he "says," "There's a tree before me"—for, whatever he may be saying, he is not saying that there's a tree before him. And if he is saying anything, why shouldn't he be saying something *true*? If his words have any meanings, they have *his* meanings. Why not assume that he means by this sentence something that is verified or made true by the situation he is in? (By "the situation he is in," I do not mean simply the mental states he is experiencing at the moment. I mean his total situation, which in-

cludes the program of the world-processor. A brain in a vat can be deceived by the “immediate” aspects of his situation, just as someone who is able to move about in space and manipulate physical objects can. It may well be that his situation is such that his “utterance” of the words ‘There’s a tree before me’ expresses a false proposition, and that he would have asserted a true proposition if he had “said,” “There’s a full-scale papier-mâché model of a tree before me.”) And isn’t that what we should assume about him? Have we any choice?

We should remember that the “hypothesis” we are considering is that we are *all* brains in vats. Presumably, therefore, if any of us speaks a language, we all do. And if that is true, then many of us speak the *same* language—the one that calls itself ‘English’, for example. Let us concentrate on that one. What is the source of the meanings of English words? What could it be besides regularities in the behavior of English-speakers? Don’t such regularities simply establish linguistic conventions? How else could linguistic conventions get established? They might, of course, be established by explicit verbal agreement. The conventions governing, for example, ‘Abelian group’ were so established. But no one would suppose that all linguistic conventions could be established that way; explicit verbal agreement is possible only when language (and hence linguistic conventions) already exist. Isn’t what makes it true that the extension of ‘red’ is a certain class of objects simply the fact that a lot of people agree—the agreement being manifested in their behavior—to apply this term to the same objects?

“But doesn’t that mean that if everyone agreed that the earth was flat, it would be flat?” It means that if everyone agreed to apply the term ‘flat’ to the earth, then ‘flat’ would be a correct term to apply to the earth.¹¹ It does not follow that ‘flat’ would mean what it actually means—after all, we actually agree *not* to apply the term ‘flat’ to the Earth. The “agreement” I am talking about is, in any case, agreement in all possible situations, including, say, observing the Earth from ten thousand miles above its surface. If we imagine a situation in which the earth is round but in which English-speakers, observing the earth from ten thousand miles up, say, “It’s flat, all right,” we are imagining a situation in which ‘flat’ means something other than what it actually means—perhaps ‘round’, perhaps something more complicated.¹²

Speaking of Russell’s “hypothesis” of an unreal past, Malcolm says that this hypothesis

requires that the newly created people should make judgments about the past, all of which are false. What criterion would there be, then, for saying that they have a past tense in their language? An omniscient observer ought to conclude that they do not have one. The case would be similar to that of a people who apparently speak English but whose “color judgments” are always or usually false. An observer ought to conclude that, contrary to what had first been supposed, they are not making color judgments.¹³ What originally looked like that has to be interpreted in some other way. The same holds for those sentences that at first appeared to express statements about the past.¹⁴

And the same holds for those statements (made by the brains in the vats) that at first appeared to express statements about trees and buildings and brains and vats and the rest of the furniture of the world. In particular, this holds for statements expressed by the words 'I am not a brain in a vat'. If someone is (and has always been) a brain in a vat, then—assuming he has a language at all—he means by 'brain' and 'vat' and 'in' something different from what these words mean in the mouths of people who are, or have, normally embodied brains. What *he* means by 'I am not a brain in a vat' is something that could be true even if it were asserted by a brain in a vat. But if that is right, then the thesis that *I* am a brain in a vat collapses. If I try to suppose that I am a brain in a vat, then I must suppose that I do not speak any language (a supposition I cannot coherently make, since one can suppose that one cannot speak only if one can speak), or else I must suppose that what I mean by 'I am not a brain in a vat' is something that could be true if I were a brain in a vat. To attempt to make either of these suppositions would involve me in a pragmatic contradiction: I can no more coherently say 'I am a brain in a vat' than I can say 'It's raining but I don't believe that it's raining'. *Someone* may be a brain in a vat, just as *someone* may not believe that it is raining when it is raining. But no one can coherently say that *he* is in either of these situations. The two cases are not perfectly parallel, however. One cannot coherently *assert* that one satisfies the description 'does not believe it is raining although it is', but one can without falling into incoherency *wonder* whether one satisfies that description. But one cannot, without falling into incoherency, wonder whether one is a brain in a vat. One can coherently raise the question whether a proposition one does not assent to is in fact true, but (if the argument I am presenting is correct) one cannot even coherently raise the question whether one is a brain in a vat. If this is correct, then the question whether one is a brain in a vat is like the question whether one in fact speaks any language at all, or the question whether one is in fact capable of raising any questions. One cannot coherently raise the question whether one speaks any language, because it is a conceptual truth that a person who spoke no language could not raise the question whether he spoke any language. One cannot coherently raise the question whether one is a brain in a vat, because it is a conceptual truth that a brain in a vat could not raise the question whether he was a brain in a vat.

Most philosophers find this argument unconvincing. The easiest way to find it unconvincing is to imagine that while one is picking one's way through its intricacies, one *is* a brain in a vat and that the Vat Keeper is monitoring one's thoughts (*via* his cerebroscope) and laughing at one. ("Hey, everybody! Look! This is priceless! He's just thought of a philosophical argument that proves he isn't a brain in a vat!") Well, *if* the argument is correct, this is an impossible supposition—since it involves the impossible supposition that one is a brain in a vat.

A more profound unease, or at least one that is not open to this quick reply, can be generated if we "turn the whole thing around" and imagine ourselves in the place of the Vat Keepers. *We* examine *our* cerebroscope in *our* laboratory and discover that the brain in *our* vat thinks he is someone called 'Hilary Putnam' and that

he is in a study in a (wholly fictional) place called 'Widener Library', constructing an argument proving that one cannot coherently suppose that one is a brain in a vat. We are, of course, immensely amused.

But, if Putnam's argument—the real Putnam's argument—is correct, we are misusing our cerebroscope. It is not telling us the thoughts of our captive brain; it is telling us the thoughts that *would* be "tokened" (as they say) in the normally embodied brain of an English-speaker if that normally embodied brain were in the same physical state as the brain in the vat. And that is *our* fault. *We* are the ones who calibrated (or whatever the word is) the cerebroscope. We have, evidently, in calibrating the cerebroscope, used the brains of English-speakers as our standards. But the brain in the vat is not the brain of an English-speaker, and the immensely amusing joke is on us. The brain in the vat either speaks no language at all—in which case we are certainly wrong in thinking that he is constructing a subtle philosophical argument for any conclusion whatever—or he speaks a language in which 'brain' and 'vat' and 'in' (and perhaps even 'prove' and 'coherently') have meanings different from the meanings they have in English. In the second case, the conclusion of the argument the brain is constructing, a conclusion that is perhaps not even *expressible* in English, may well be true.

Bouwsma has made a similar point. Speaking of Descartes's evil genius, he says,

I intend to show that the evil genius is himself befuddled, and that if we exhaust some of our energies in sleuthing after the peculiarities of his diction, we need not be deceived either.¹⁵

Bouwsma means that a being who attempts global deceptions of the Cartesian variety (whether he stands outside Nature or merely outside a vat) succeeds only in producing a creature who, if he speaks any language at all, speaks a language that is semantically radically unlike the Global Deceiver's,¹⁶ although it may be homophonic¹⁷ with the Deceiver's language. This language—if it exists—is one that the Evil Genius or the Vat Keeper or the Creator of a World-That-Began-Five-Minutes-Ago cannot, for all his airs, understand. He cannot understand it because he was not brought up in a community in which it was spoken. (Nor does he belong to a community in which it could be taught to him: He and those he thinks he is deceiving have no common repertory of ostensive gestures—they, in fact, do not make what he calls 'gestures';¹⁸ he cannot even understand what *they* mean by 'gesture' or 'over there'. If we imagine the Vat Keeper to be human and a member of our linguistic community—we have no way of imagining his two colleagues in attempted global deception, and they certainly are not members of *our* linguistic community—we can suppose that when the brain in the vat says, 'Now this is what we call a nasturtium,' a computer screen displays a picture of a man pointing at a flower. We can imagine that this picture is constructed from cerebroscopic data, eked out with information about the Master Computer's world-processing program. But a picture must be interpreted, and a picture can have more than one interpretation. If the Vat Keeper

deduces from the picture he sees that the brain in the vat currently believes himself to be pointing at a flower, he is misinterpreting that picture.) If the Evil Genius or the Vat Keeper thinks he is deceiving someone about almost everything, he is deceiving only himself. Like Screwtape's unfortunate nephew Wormwood, he is an incompetent deceiver.

To complete our account of how attempted global deception looks "from the other end," we should note that the brain in *our* vat (1) cannot understand what is expressed by the *English* predicate 'is a brain in a vat', and (2) cannot coherently wonder whether he satisfies the predicate of *his* language—always assuming he has one—'is a brain in a vat', although he knows perfectly well (and we cannot know) what is expressed by that predicate. In making this second statement, of course, I am assuming that the brain in the vat is such that, if he were caused to be normally embodied, the resulting man-shaped organism would pass the behavioral tests by which we pick out competent speakers of English—as opposed to speakers of a language in which 'is a brain in a vat' means something different from what it means in English.

This essentially completes my exposition of Putnam's argument for his second premise ('It is not intelligible to suppose that we are all wrong about almost everything'). We need only note that the "hypothesis" that we are all brains in vats is not importantly different from any other tale of global deception: If Putnam's argument shows that it is unintelligible to suppose that we are brains in vats, it could easily be modified to show, with respect to any given tale of global deception, that we cannot intelligibly suppose that tale to be true. But if we are all wrong about almost everything, then some tale of global deception must be true. Whether Putnam's argument is cogent depends on many disputed questions in the philosophy of language and the philosophy of mind. Theories of mind and language according to which this argument is certainly *not* cogent have recently been developed by several philosophers, most notably John Searle in *Intentionality*¹⁹ and Roderick M. Chisholm in *The First Person*.²⁰ (Searle and Chisholm do not seem to say anything in these books that would count as a statement about the nature of truth.) It is not my purpose to try to decide who is right about the second premise of Putnam's argument. My business, as I have said, is with its first premise ('If truth is radically nonepistemic, then it is intelligible to suppose that we are all wrong about almost everything').

Why does Putnam believe this? Why does he think that a (metaphysical or external) realist should believe that it is intelligible to suppose that we are all wrong about almost everything? What premises in Putnam's argument for the incoherency of supposing that we are brains in vats must the realist reject?

In *Reason, Truth and History*, Putnam says,

Before I give the argument [the argument presented above for the conclusion that it is unintelligible to suppose that we are all brains in vats], let us consider why it seems so strange that such an argument can be given (at least to philosophers who subscribe to a 'copy' conception of truth). . . . it is compatible

with physical law that there should be a world in which all sentient beings are brains in a vat. As philosophers say, there is a 'possible world' in which all sentient beings are brains in a vat. . . . The humans in that possible world have exactly the same experiences that *we* do. They think the same thoughts we do (at least, the same words, images, thought-forms, etc., go through their minds). Yet I am claiming that there is an argument we can give that shows we are not brains in a vat. How can there be? And why couldn't the people in the possible worlds who really are brains in a vat give it too? (p. 8)

What Putnam asks us to consider in this passage is the question, Why does it seem so strange to philosophers who subscribe to a copy theory of truth that there is an argument that shows that we are not brains in vats? He does not ask us to consider the question we are at present interested in: Why must philosophers who believe that truth is radically nonepistemic *reject* the thesis that there is an argument that shows that we are not brains in vats? But perhaps an answer to the former question would provide the materials out of which an answer to the latter could be constructed. The main trouble with the passage I have quoted is that it does not seem to provide us with much of an answer to the question that it asks us to consider. Insofar as it does contain an answer to this question, this answer would seem to be contained in the two rhetorical questions with which it ends. We are, presumably, to suppose that those questions are being asked by subscribers to a 'copy' conception of truth. Here is why the thesis that a certain kind of argument is possible seems strange to them: It raises (for them, if not for more enlightened philosophers) those questions.

But why? Why should it raise those questions *for them* – for them in particular? Any philosopher *might* be puzzled by the announcement of an argument that proves that we are not brains in vats; and that puzzlement might lead him to ask those questions. But why should those questions be particularly closely associated with a 'copy' theory of truth – *or* with the thesis that truth is radically nonepistemic? Why shouldn't the "copyist," when he has seen Putnam's argument for the incoherency of supposing that we are brains in vats simply say, "Oh, now I see," and continue to be a copyist? Why shouldn't the proponent of the thesis that truth is radically nonepistemic react in that way, too? This is precisely the question we have been asking. I do not know of any answer to it. I think it has no answer. (On p. 22 of *Reason, Truth and History*, Putnam says that the reason it is surprising that the brain-in-a-vat hypothesis turns out to be incoherent is that "we are inclined to think that *what goes on in our heads* must determine what we mean and what our words refer to." Well, perhaps we are inclined to think this, and perhaps it is false. But why should the inclination to think this be associated with realism? On this point, see notes 5 and 25 to the present paper.)

If Putnam is right about brains in vats, then, as we have seen, to suppose that we might all be brains in vats is to fall into pragmatic contradiction. Supposing that we might all be brains in vats could be compared with supposing that no one speaks a language, or that no one is able to suppose that some thesis is true. I expect that

most philosophers, whatever theories they might hold about the nature of truth, would concede that a person who supposes that no one can speak any language has fallen into incoherency (and this despite the fact that the proposition that no one speaks any language is not self-contradictory). And this thing that all philosophers would concede does not seem to favor any given theory about truth over any other. In particular, the (metaphysical) realist can accept it with perfect equanimity. But then what prevents the realist from accepting the corresponding thesis about supposing that everyone might be a brain in a vat? To suppose that everyone might be a brain in a vat is either to fall into a pragmatic contradiction or it is not. In the latter case, Putnam's argument is unsound. In the former case—well, the realist admits that there are *some* self-consistent propositions such that a person cannot, without contradiction, suppose with respect to one of those propositions that it is true—why should a proof that any particular proposition falls into this category disturb him?

It may be instructive to substitute some other one of these propositions in the argument I have attributed to Putnam. For example:

- (1a) If truth is radically nonepistemic, then it is intelligible to suppose that no one speaks any language.
- (2a) It is not intelligible to suppose that no one speaks any language.

hence, Truth is not radically nonepistemic.

In this case, I dare say, almost everyone would grant the second premise. But why should anyone grant the first? And why is the first premise of Putnam's argument—that is, of the argument I have attributed to Putnam—any more plausible than (1a)? (It will not do to argue that 'We are all wrong about almost everything' has epistemic implications, and is thus particularly apt to have some sort of conceptual connection with 'truth is radically nonepistemic'. To speak a language entails being right about many, many things. For example, to speak English entails being right in believing that 'cat' does not mean the same as 'dog', that 'colored' applies to anything 'red' applies to, and so on. Therefore, the thesis that no one speaks a language has epistemic implications only slightly less profound than those of the thesis that we are all wrong about almost everything.)

In my view, there is no answer to this question, and Putnam's argument for the incoherency of the "hypothesis" that we are brains in vats has, even if it is correct, no implications whatever for the thesis that truth is radically nonepistemic. What Putnam's argument shows (assuming it to be correct) is that the *capacity to state and believe propositions* is "radically epistemic": If we were not right about most things, we could not even be *wrong* about very much. Thus, the supposition that we are all wrong about almost everything is (if Putnam's argument is correct) incoherent. But this is a far cry from saying that *truth* is radically epistemic. The only truths that depend for their status as truths on our knowing lots of things or on our almost always being right are the truths that everyone would admit had that feature, and the types of dependence involved are obvious and boring: It is because (formal cause)

our beliefs are almost always right that it is true that our beliefs are almost always right; it is because (efficient cause) we are often wrong about our own motives that our judgments about other people's motives are often wrong. And so on.

The realist will distinguish in the following way between the conditions necessary for the acceptance and statement of propositions and the conditions necessary for the existence of truths. (At least many realists will. At least I will.) Suppose someone believes that Mount Everest is 8,847.7 meters high. Or suppose he says, "Mount Everest is 8,847.7 meters high." (Let us concentrate on the problems raised by the former supposition, for what we shall say about belief can be adapted easily enough to the case of assertion.) What he believes is true. The fact that what he believes is true is (as the realist sees matters) in a fairly obvious way a conjunction or combination of two facts:

- (F1) He believes (or, as I prefer to say, accepts) the proposition that Mount Everest is 8,847.7 meters high.
- (F2) The proposition that Mount Everest is 8,847.7 meters high is true.

Understanding the nature of the former fact is a deep problem in the philosophy of mind. (I should say that it was one of the two central problems in the philosophy of mind, the other being the problem of the nature of sensuous experience.) Some philosophers (Chisholm and Searle, for example) say that the existence or holding or obtaining of a fact of this sort is entirely a matter of what goes on in the subject's head (or in his mind or soul—the mind-body problem is not what is at issue here). Other philosophers (Putnam and Kripke, for example) say that it involves other things as well—perhaps the subject's being a member of a certain linguistic community, or his being causally related to a certain large mass of rock. If Putnam's argument for the incoherency of supposing that everyone is a brain in a vat is correct, then Chisholm and Searle must be wrong; it certainly does follow from their position that a person's "vat-Doppelgänger" accepts exactly the propositions that that person accepts. (In saying that this is a consequence of their position, I pass over the issue of propositions expressed by sentences containing indexicals.) And, if Putnam's argument is correct, then either one's vat-Doppelgänger accepts hardly any propositions at all, or else he accepts a set of propositions most of which one could not even grasp. From the realist's point of view, Putnam's argument has nothing to do with the nature of truth. It has nothing to do with the nature of the relation that must hold between a proposition and the world in order for the latter to *confer truth on* the former. It has, rather, to do with the nature of the relation that must hold between a proposition and a subject in order for the latter to *accept* the former. The realist who accepts Putnam's argument will describe its import like this: Putnam has shown that certain of the necessary conditions for accepting the proposition that everyone is a brain in a vat—or for accepting *any* proposition about brains and vats—could not be satisfied by anyone if everyone were a brain in a vat. Of course (the realist might add), if everyone were a brain in a vat, this proposition would be *true*; it is just that no one would be able to accept it or grasp it or consider it.

The nature of the second fact, (F2)—that the proposition that Mount Everest is 8,847.7 meters high is true—represents a problem of an entirely different sort. (This problem is not the problem solved by Tarski:²¹ the problem of giving a mathematical characterization of the set of sentences in a formal language *L*—a language satisfying certain constraints—that satisfy the predicate ‘expresses in *L* a true proposition’.²² It has next to nothing to do with *that* problem. Nor does it have much to do with the important extensions of Tarski’s work by Saul Kripke²³ and by Anil Gupta,²⁴ which treat formal languages that are in a certain philosophically important respect more like natural languages than is the sort of language for which Tarski solved the problem. I do not mean to depreciate the work of Tarski, Kripke, and Gupta, but merely to classify it. The classification, moreover, represents a philosophical position: When I say that Tarski’s results have next to nothing to do with the problem of what it is for a proposition to be true, I do not claim to be stating an obvious fact; I am only saying where I stand.) I have little to say about this problem. I note only that, as a realist, I have the following picture: Propositions, like numbers, are necessarily existent, abstract objects; the same propositions exist in every possible world; in different possible worlds, there are different concrete objects, and even in two worlds that have the same population of concrete objects, these objects may have different intrinsic qualities and be related differently to one another; what change from world to world, therefore, are the population, qualities, and arrangement of concrete objects, and it is these features of concrete objects that are the formal causes of the truth and falsity of propositions. The problem of the nature of the fact (F2) is, in its general form, the problem of the nature of the relation that holds between a world composed of certain objects having certain intrinsic qualities, and related to one another in a certain way, on the one hand, and a proposition, on the other, in virtue of which the world is the formal cause of the truth-value of the proposition. The realist position is: This relation has nothing to do with the existence of minds, or with their contemplation of the proposition or of the arrangement of objects that causes it to be true. *Some* worlds, indeed, contain minds, and some propositions are about minds, but the relations “makes true” and “makes false” hold not only in those worlds, but also in worlds (if such there be) wholly devoid of mind.

This is the realist position on the nature of truth. Perhaps it is an obscure position. It is, after all, a position on a fundamental philosophical question, and these positions tend to be obscure. It is certainly no more obscure than the antirealist’s position on the nature of truth (rather less so, if you ask me). But its obscurity or lack of it is neither here nor there. The essential point for our purposes is this: Putnam’s brain-in-a-vat argument has nothing to do with this position. The brain-in-a-vat argument is relevant only to this question: What is it for a proposition to be accepted, grasped, or entertained? It is not relevant to the question: What is it for a proposition to be *true*. I do not, as I have said, understand a “copy” or Tractarian account of truth. But I see no reason why an adherent of a copy account of truth should not accept Putnam’s argument. I can see no reason why he should find it “strange” that such an argument can be given.²⁵ Why shouldn’t the “copyist” say this: “I accept the proposi-

tion that Mount Everest is 8,847.7 meters high. And this proposition is true because it and a part of the world have the same structure. My vat-Doppelgänger accepts a proposition that *he* expresses by the words 'Mount Everest is 8,847.7 meters high'. This proposition is true because it and a part of the world have the same structure"? As far as I can see, the only thing that could prevent his saying this would be certain theories in the philosophy of mind, theories that, although consistent with realism, are by no means entailed by realism, theories according to which the propositions accepted by a thinker are entirely determined by "what is in his head."

The same point can be made about the philosopher who believes that truth is radically nonepistemic. Why shouldn't he say: "I accept the proposition that Mount Everest is 8,847.7 meters high. The way the world is makes this proposition true, and this 'making true' relation is one that would hold between the way the world was and certain propositions even if there were no minds. My vat-Doppelgänger accepts a proposition that *he* expresses by the words Mount Everest is 8,847.7 meters high'. The way the world is makes this proposition true, and this 'making true' relation is one that would hold between the way the world was and certain propositions, even if there were no minds"? There is, of course, one important difference between you and me and the propositions about Mount Everest that *we* accept, on the one hand, and the propositions expressed by "Vat-English" sentences containing the words 'Mount Everest', on the other. The propositions about Mount Everest that you and I accept (the geological ones, anyway, as opposed to the historical and sociological ones) do not depend for their truth-values on the thoughts or actions of human beings. But the proposition expressed by the Vat-English sentence 'Mount Everest is 8,847.7 meters high' would have been false if the world-processing program controlling the stimuli received by the brain in the vat had been different in certain minor ways, or, *a fortiori*, if it had never been written. That proposition would have *existed* even if there had never been a physical universe (like all propositions, it is necessarily existent), but its truth-value is intimately connected with the specific actions of specific human beings. But that does not affect my point. That is only to say that the thoughts and actions of human beings are among the efficient causes of the truth of this proposition; it is not to say that the thoughts and actions of human beings have anything to do with *what it is* for that proposition to be true. (We should note, by the way, that the sentence 'The fact that Mount Everest is 8,847.7 meters high in no way depends on the thoughts or actions of human beings' expresses, in Vat-English, a true proposition.)

I conclude that defenses of antirealism, and attacks on metaphysical realism and the copy theory of truth and the thesis that truth is radically nonepistemic, must be based on something other than arguments showing that we cannot coherently suppose that we are brains in vats.

Notes

1. Compare Hilary Putnam, "Realism and Reason," in *Meaning and the Moral Sciences* (London, 1978), 125.

2. I do not, however, wish to deny that if someone were to say, "Mt. Everest is 8847.7 meters high" (that is, if someone were to pronounce those words in a context in which his pronouncing them would constitute his saying that Mt. Everest was 8847.7 meters high), then his saying this would have certain "conversational implicatures" or presuppositions that would be false if various of the states of affairs enumerated in the text failed to obtain. For example: that 'Mt. Everest' is a name for a certain mountain; that he has some way of knowing how high that mountain is; that 'meter' is a name for a unit of linear measure. But none of these propositions whose truth is presupposed by his asserting the proposition that Mt. Everest is 8847.7 meters high (in those words) is entailed by that proposition.

3. Or no one but God. I do not quite know how to deal with God—that is, with a necessarily existent and essentially omniscient being—in a discussion of the present topic. I should *like* to say that even though (because of God's necessary existence and essential omniscience) there is no proposition that no one ever thinks of, and no true proposition that is not believed by anyone, this fact is somehow irrelevant to what is at issue in the realism/antirealism debate. But I am not sure how best to articulate this inclination. Tentatively, I would say this: Although God's "mental states" are among the efficient causes of the truth-values of all contingent propositions, they are not among the efficient causes of the truth-values of necessarily true or necessarily false propositions—for the truth-values of noncontingent propositions have no efficient causes. And both of the following propositions are necessary truths: (1) the proposition that snow is white exists; (2) if snow is white, and if the proposition that snow is white exists, then that proposition is true. These two necessary truths, together with the contingent truth that snow *is* white, entail that the proposition that snow is white is true. God's "mental states" are doubtless among the efficient causes of snow's being white. They are, therefore, among the efficient causes of the truth of the proposition that snow is white. But they are not among its formal causes; they are not constituents of the fact that the proposition that snow is white is true; they need not be mentioned in a correct account of what it *is* for that proposition to be true. An account of what it is for that proposition to be true—I do not have a nontrivial one—need mention only the intrinsic or nonrelational properties of the proposition that snow is white, and the fact that snow is white (how the whiteness of snow came about would not be relevant).

The philosopher who believes that truth is radically nonepistemic, and who is also an atheist, has available to him an effective "intuition pump" (in Dennett's sense) that he can use in good conscience. He can imagine a world in which there are no minds, but (of course) lots of truths and falsehoods. I have made some use of this intuition pump in the text, but not in very good conscience, since I do not believe that there are any such worlds. My only defense is that my use of this pump is not supposed to constitute an argument for the mind-independence of truth, but only to provide the idea that truth is mind-independent with some intuitive content—and, anyway, the present paper is not a defense of the mind-independence of truth.

4. 'Known to' here refers to "knowledge by acquaintance." The proposition that the truths of arithmetic can be specified recursively is a *false* proposition that is "known" to us in this sense, and which could not have been known to us if we had been very much less intelligent than we are.

5. In "Realism and Reason" (p. 125) and elsewhere, Putnam says that metaphysical realism is the thesis that there is a "determinate relation of *reference*" between the terms of a language and pieces of "the world." (In "Realism and Reason," Putnam writes 'THE WORLD' rather than 'the world'. I think we may take the world—lowercase or capitalized—to be constituted by those objects, if such there be, whose existence and attributes are independent of the mental states of the speakers of the language in question, or whose existence and attributes "depend" on those states only in a crude, causal sense. The function of the capitals appears to be to suggest that realists treat the attribute of mind-independence as conferring a sort of quasi-religious awfulness on its possessors.) If this is what "realism" is, then I am a realist in this sense, too. Putnam tells us in "Realism and Reason" (p. 125) that the thesis that truth is radically nonepistemic is a "consequence" of the thesis that there is a determinate notion of reference. I have a certain amount of trouble with this contention in that, in my view, truth is a property of items that do not belong to any language—propositions—and, thus, I see the connection between theses about reference and theses about truth as rather problematical. Perhaps Putnam's contention is this: The thesis that there is a determinate relation of reference entails the thesis that *expressing some true proposition or other* (a

relational property of some sentences) is radically nonepistemic. I shall not discuss this contention. I should say that in accepting the position that there is a determinate relation of reference, I do not mean to commit myself to any theory about the “determinants” of this relation; in particular, I do not mean to commit myself to the thesis that the referent of, for example, ‘Mt. Everest’, as used by Sir Edmund Hilary, is wholly determined what goes on “in the head” of Sir Edmund. (Compare note 25.) I should also say that, because I do not understand the “copy” theory of truth, I decline to regard my belief in a determinate relation of reference as committing me to that theory.

Putnam also suggests in “Realism and Reason” (p. 125) that “realism” is the thesis that an ideal scientific theory might be false— “ ‘ideal’ from the point of view of operational utility, inner beauty and elegance, ‘plausibility’, simplicity, ‘conservatism’, etc.” I will not discuss this form of “realism,” except to remark that the conceivability of a false ideal theory seems extremely plausible on scientific grounds. It seems plausible that there might be two competing and logically incompatible theories, each of which was “ideal” in this sense—and one of them, presumably, would have to be false. Two incompatible competing theories might both be “ideal” because it might be physically impossible (even demonstrably physically impossible) to perform any experiment of the sort that would be needed to choose between them. (And this physical impossibility might be a consequence of *both* theories.) Current physical theory suggests that this is not only a logical but an epistemic possibility. Some physicists have speculated that an experiment that could choose among competing versions of the “Grand Unified Theory” would require energies on the order of those that would be produced by a linear accelerator light-years in length and powered by the total output of several stars. It is reasonable to suppose that, given the actual distribution of available energy in the universe, the production of such energies would be not only technologically impossible, but thermodynamically impossible. And what if it could be shown that the production of such energies would be thermodynamically impossible in any comic epoch in which the average energy-density of the universe was compatible with the existence of living organisms?

6. One might maintain these things—or one might not. It is consistent with everything I shall say later to suppose that a brain in a vat is not right even about simple a priori truths or about its own pains and beliefs.

7. Cambridge, 1981.

8. See *The Analysis of Mind* (New York, 1921), 159–60.

9. *The Philosophical Review* 58 (1949): 141–51.

10. In *Knowledge and Certainty: Essays and Lectures* (Englewood Cliffs, N.J., 1963), 187–202.

11. Compare Wittgenstein, *Philosophical Investigations*, 241.

12. It might not be a shape-predicate at all—if, for example, they refused to apply it to billiard balls.

13. Here, I believe, Malcolm overstates his case. All that follows is that they are not making the color judgments we should make using the same words: Perhaps they use ‘red’ to mean what we mean by ‘yellow’, and so on. Malcolm’s conclusion would require a stronger set of assumptions; it would require at least the assumption that their apparent color-judgments are not related systematically to the colors of the objects in their environment.

14. “Memory and the Past,” 196.

15. “Descartes’ Evil Genius,” 142.

16. I imagine that Bouwsma would have smiled to read this highfalutin paraphrase of his words: “ ‘Semantically radically unlike—so *that’s* what I meant! My, my!”

17. Whatever, exactly, *homophonic* may mean in this context.

18. To my mind, this is a good reason for saying that the brain in the vat does not speak any language at all.

19. Cambridge, 1983.

20. Minneapolis, 1981.

21. “The Concept of Truth in Formalized Languages,” in *Logic, Semantics, Metamathematics* (Oxford, 1956), 182–278.

22. Tarski, of course, would say ‘is true in L,’ where I say ‘expresses in L a true proposition’ or (even better) ‘expresses in L some true proposition or other’.

23. "Outline of a Theory of Truth," *Journal of Philosophy* 72 (1975): 690-716.

24. "Truth and Paradox," *Journal of Philosophical Logic* 11 (1982): 1-60.

25. Nor do I see any reason to suppose that an adherent of the thesis that there exists a "determinate relation of reference" should find it strange that such an argument can be given. Such a philosopher will contend that this determinate relation of reference (or, more generally, extension) holds between the English word 'vat' and certain large containers. And, if he accepts Putnam's argument, he will contend that the "Vat-English" word 'vat' bears this same determinate relation to something-we-know-not-what—provided that there *is* such a language as Vat-English. He must, of course, deny that what determines this relation to hold between the words of a given language and various objects or sets of objects is entirely "inside the heads" of the speakers of that language. Compare note 5.