

Putnam and Physical Realism

Realism in regard to the physical world seems about as plausible as realism can get. And yet there are many serious and insightful philosophers who reject it. Hilary Putnam is a particularly instructive example, in part because he was for many years 'wedded to' a version of the view in question.¹ Then, in 1976, there was a divorce. Since that time he has offered various objections to physical realism.

In 1974, Putnam was a full-fledged physical realist, and a realist in regard to several other domains as well. In the introduction to *Mathematics, Matter and Method* he wrote:

These papers are all written from what is called a **realist** perspective. The statements of science are in my view either true or false (although it is often the case that we don't know which) and their truth or falsity does not consist in their being highly derived ways of describing regularities in human experience. Reality is not a part of the human mind; rather the human mind is a part - and a small part at that - of reality.²

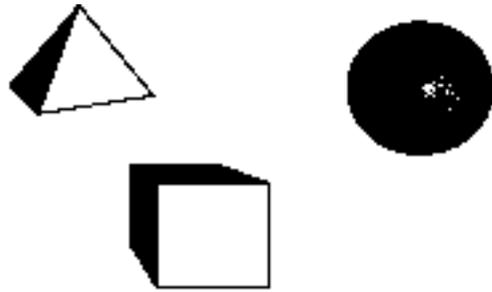
...

To prevent misunderstandings, let me say that by realism I do not mean materialism. Part of the burden of these essays is that a consistent realist has to be realistic not only about the existence of material objects in the customary sense, but also has to be realistic about the objectivity of mathematical necessity and mathematical possibility (or equivalently about the existence of mathematical objects) and about entities which are neither material objects nor mathematical objects - in particular, fields and physical magnitudes.³

This is a view I accept.⁴ It is therefore a bit disconcerting that Putnam came to regard his earlier realism as incoherent. He now attacks 'metaphysical' realism. [This is not, exactly, the view that he himself had held prior to 1976. Perhaps we should think of it as a codification of what he took to be erroneous assumptions underlying the earlier view.] 'Metaphysical realism' is a package-deal which includes most, or perhaps all, of the following claims:

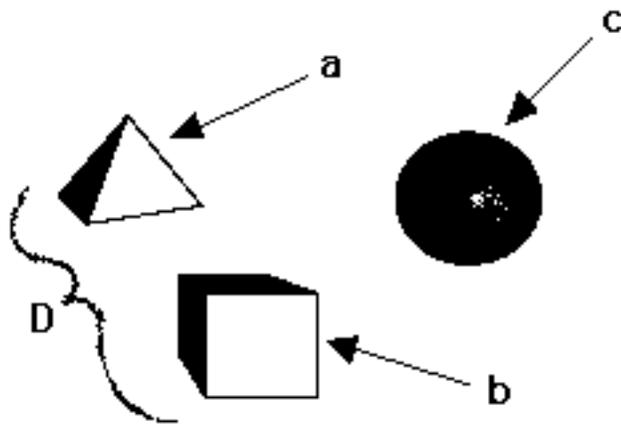
- (1) The world consists of some fixed totality of objects.⁵ That is to say, there is a particular set of objects, O , that constitute the world.⁶
- (2) These objects are mind-independent.
- (3) There is just one true and complete description of the world.
- (4) Truth about the world involves a correspondence between words or thought-signs and 'external' things and sets of things.
- (5) A theory of the world which is 'ideal' from the point of view of operational utility, elegance, plausibility, simplicity, and so on, might, none the less, be false.⁷

These claims are linked to a certain model of the relation that must hold between correct theories, or descriptions, and all, or part of, the world. Consider a world made up of three objects - for instance, Democritean atoms:



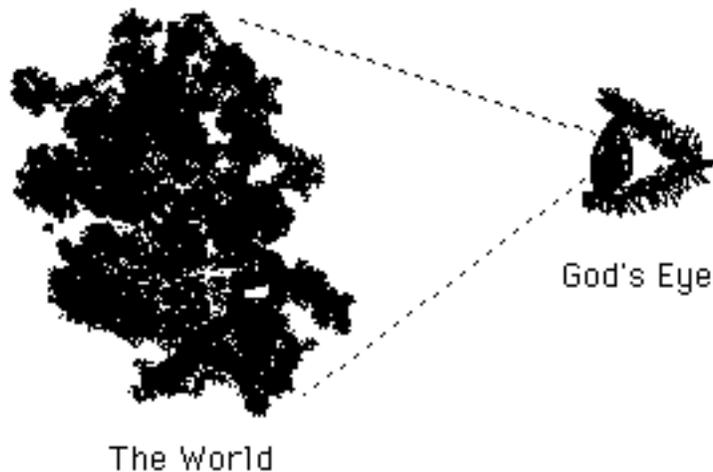
The World

These objects are supposed to be things that are not in any way dependent upon theories about them, or perception of them, or social practices which include them. They are 'mind-independent'. Now imagine a language such that each term in the language denotes a piece, or a kind of piece, of The World. ⁸ For example, there are names 'a', 'b', and 'c', for the three objects, and a general term 'D', denoting a and b, but not c. That's the model in its basic form.



The World

As I have said, Putnam calls the package 'metaphysical realism'. I don't want to endorse this use of the phrase. Instead, I'll call it '**atomistic**' realism. [Perhaps a better name would be 'the One True Theory theory'.] The title is meant to underline similarities between this doctrine and the 'atomism' defended by Moore, Russell, and perhaps Wittgenstein, in the early nineteenth hundreds.⁹ I am going to suggest that Putnam was right in rejecting the package; but, I will add, this leaves **physical realism** (as distinct from atomistic realism) intact, or better than intact - purged of some unnecessary, and implausible, claims. Putnam identifies atomistic realism with an '**externalist**' perspective "... because its favorite point of view is a God's Eye point of view."¹⁰



The picture itself should be acceptable to any physical realist. Don't we think there is a complex, structured, mind-independent, physical world for God to contemplate from 'outside' (assuming, for the moment, that there is a God, and a place outside)?¹¹ Putnam regards the picture as something more controversial. "... the motive of the metaphysical realist is to save the notion of the God's Eye Point of View, i.e. the One True Theory."¹² Why should physical realists accept claim (3)? Why should we hold (a) that a true and complete description of the world is possible,¹³ and (b) that there can, at most, be only one such description?

Hartry Field, a card-carrying realist in regard to the physical world, flatly rejects (3): How could there be only one true and complete description of the way the world is? The concepts we use in describing the world are not inevitable: beings other than ourselves might use predicates whose extensions differ from anything easily definable in our language, or whose extensions differ from anything definable **at all** in our language; indeed, their language might contain no predicates at all, but referential devices of a radically different sort.¹⁴ Putnam replies:

.... suppose metaphysical realism is true. Then there is a definite set I of individuals of which the world consists (in the world-picture of Field's *Science without Numbers*, these would be just the space-time points). And there is a definite set P of all properties and relations of these individuals. Consider a "language" with a name for each member of I and a predicate for each member of P. Such a language is not a **denumerable** language; true. But there is no reason for a **metaphysical realist** to think that the true and complete theory of the world should be expressible in a **denumerable** language. This language is unique (up to isomorphism), and the theory of the world - the set of true sentences - is likewise unique. So there is nothing wrong with the assertion (which Field spends pages in criticizing) that, on a metaphysical realist view, there is "one true theory." Metaphysical realists have always thought in terms of an **ideal** language, not a natural one, with its vagueness, its finite vocabulary, etc. **15**

[Note that Putnam does not treat (3) as part of the definition of 'metaphysical realism'.] Putnam contends that 'metaphysical realists' are committed to claim (1), and that, given (1), it is 'possible' in some sense that there should be a 'language' which includes a distinct name, or term, for each of the ultimate bits, or whatever, that make up the world, and a distinct predicate for each of the properties and relations of these bits. If such a 'language' is possible, then a true and complete description of the physical world is also possible. It is simply the set of all the true sentences of this 'language'. According to Putnam, both the 'language' and the description would be unique (up to isomorphism). So, Putnam thinks, if we accept (1), we must accept (3). And, he might add, if we reject (3), we must reject (1).

Let's grant that (1) implies that a correct and complete description of the world is, in some sense, possible. The next step is more difficult. Does the alleged possibility of this correct and complete description show that there cannot be two or more such descriptions?

Putnam's First Attack

Putnam argues (both before and after his abandonment of physical realism) that there can be correct and complete, but seemingly incompatible, intuitively different, descriptions of the world.

Consider a 'world' that is just a straight line - that's all there is to it. **16**

Theory One: The World is made up of points - infinitely small parts of both the world itself and all its component line segments.

Theory Two: The World is made up of line segments, not points. Talk about so-called 'points' should be understood as talk about convergent sets of line segments.

Putnam regards these theories as seemingly incompatible and intuitively different. After all, one of them says "The world is made up of points", while the other says "The world is not made up of points". It is fairly clear that he thinks there are two theories here, and not just one expressed in two different ways.

In principle, both theories can be built up into complete accounts of The World. The

result is that atomistic realism is in trouble. (3) forces the atomist to hold that at least one of the theories is false. But this is absurd. The theories are correct and equivalent.¹⁷ So (3) is false.

The point is important:

A twentieth-century realist cannot ignore the existence of equivalent descriptions: realism is not committed to there being one true theory (and only one). ...

Assuming there is a 'fact of the matter' as to 'which is true' (if either) whenever we have two intuitively 'different' theories is naive. 'Theories' which differ on which pairs of events are simultaneous are certainly 'intuitively different', but after Einstein we know such 'theories' may, nonetheless, be equivalent. [W]e should be ... sophisticated realists. And a sophisticated realist recognizes the existence of equivalent descriptions, because it follows from his theory of the world that there are these various descriptions, as it follows from a geographer's description of the earth that there are alternative mappings (mercator, polar, etc.).¹⁸

This was Putnam still in his realist phase.¹⁹

[It is thus clear that he did not, as a realist, always adhere to the whole 'metaphysical' package.] I think he was right, and that this 'sophisticated' version of physical realism represents a significant clarification of the theory. What would he have said, in those days, about Theory One and Theory two? He tells us he would have said that Theory One maps the term 'point' onto the very same things that Theory Two calls 'sets of convergent line segments'. That is to say, there are things in The World that bear both labels.²⁰ So where's the problem? Putnam says:

The problem - as Nelson Goodman has been emphasizing for many, many years - is that this story may retain THE WORLD but at the price of giving up any intelligible notion of **how** THE WORLD is. Any sentence that changes truth-value upon passing from one correct theory to another correct theory - e.g. an equivalent description - will express only a **theory-relative** property of THE WORLD. And the more such sentences there are, the more properties of THE WORLD will turn out to be theory-relative.

For example, if we concede that [Theory One and Theory Two] are equivalent descriptions, then the property **being an object** (as opposed to a class or set of things) will be theory-relative.²¹

Putnam's relativistic principal looks very plausible. I think we should accept it. Here it is again:

Any sentence that changes truth-value upon passing from one correct theory to another correct theory expresses only a theory-relative property of THE WORLD.

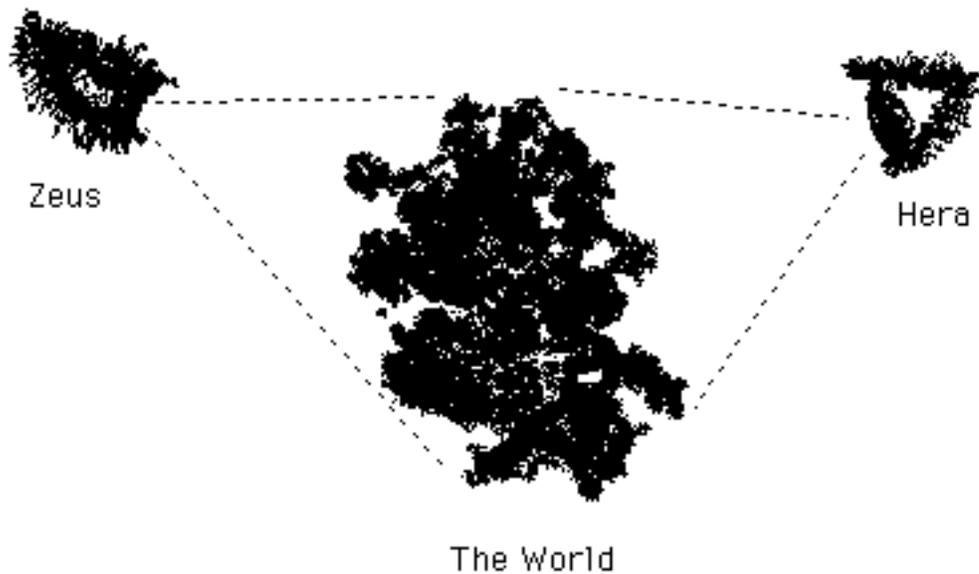
Can we admit this, and remain physical realists? How far towards anti-realism does it drive us? By hypothesis there are things in The World which, according to Theory One, are points and 'objects', but, according to Theory Two, are sets of convergent line segments and not 'objects'. Suppose some god names one of these things 't'. The sentence, "t is an object," is true relative to Theory One and false relative to Theory Two. That is to

say, the property of **being an object** is theory-relative.

If we take (1) to say it's a flat-out fact (not a theory-relative fact) that there is a particular set of objects such that those objects and no others constitute the world, then (1) is false in regard to the 'one line' world.²²

There is also a problem about '**complete**' theories. Given that no objects are, absolutely, **the** objects that constitute the world, we can't say that a 'complete' theory is one which, among other things, provides names, or denoting expressions of some kind, for all those objects.

Theory-relativism may well extend further. Suppose there were two enormously powerful supernatural beings with almost limitless cognitive powers - Hera and Zeus. Each of them has a correct, rich, and accurate, account of the physical world - our world; but they 'see' it in very different ways.



Zeus 'sees' the world as composed, basically, of 'alphatrons'. And, according to his view, there are only denumerable many of these things. Hera, on the other hand, 'sees' the world as made up of non-denumerable many 'omega-bits'.²³ Their theories are, in this regard, 'incompatible'. Is it the god, or the goddess, who 'sees' things as they are? Is the world basically composed of alphatrons, or omega-bits, or neither one? Is the world made up of denumerable or non-denumerable many things? These questions presuppose that at least one of the theories must be wrong. But this presupposition may be false. In fact the

theories may be equivalent, and both correct. Suppose this is the case.

According to Zeus' theory, the physical world is composed of denumerable many basic objects. But does he believe this? Zeus has superlative knowledge. Presumably he knows that there is a theory equivalent to his own which says the world is made up of nondenumerably many basic objects. And, we may suppose, he fully accepts Putnam's relativistic principal. Consequently, what he believes is that **relative to his theory** there are denumerably many basic objects. He does not take this to be a flat-out fact about the world.

Hera runs the same line of thought the other way around. As a result, she doesn't think it's a flat-out fact that the world is composed of non-denumerably many basic objects. The up-shot is that Zeus and Hera are in complete agreement about the number of basic objects. They see this as a theory-relative matter. If they did not, they would both have mistaken conceptions of the world.

In spite of its relativistic aspect, the emerging picture illustrates a form of physical realism. The world is depicted as something that is what it is no matter how Hera and Zeus theorize about it. Each of them wants to 'see' it correctly - to understand it. The adequacy of their divergent views is measured in part by how well they can explain the world's features and predict the events that occur in it. This is physical realism; but there's no hint here of the 'One True Theory' - quite the contrary. Putnam hasn't shown that there are correct theories of our world that differ in regard to the cardinality of the set of 'basic objects' they postulate. But the suggestion is far from absurd. Suppose it's true. Would this refute claim (1) of the atomistic package? I think it would; but there is a hint of unclarity in (1). Let's re-write it in such a way as to guarantee that the supposition would do the job.

(1*) Either (a) there is a particular denumerable set of objects that constitute the world, and this fact is not theory-relative, or (b) there is a particular non-denumerable set of objects that constitute the world, and this fact is not theory-relative.

Physical realists need not accept (1*). Moreover, Putnam has given us grounds for regarding it as implausible. But then, if (1*) is not a crucial component of the physical realist picture, what is? Perhaps we should stick with something modest like this:

(1**) There is a unique, structured, mind-independent, physical world.²⁴

No doubt many claims and beliefs about the world are theory-relative. The question now is: are there any that aren't? If the answer is "no", then the claim that we can describe features of the 'structured, mind-independent world' seems unjustified. This, in fact, is the last move in Putnam's argument.

One can construe space-time points as objects, for example, or as properties. One can construe fields as objects, or do everything with particles acting at a distance (using retarded potentials). The fact is, so many properties of THE WORLD - starting with just the categorical ones, such as cardinality, particulars, or universals, etc. - turn out to be 'theory-relative' that the THE WORLD ends up as a Kantian 'noumenal' world, a mere 'thing in itself'. If one cannot say how THE WORLD is theory-independently, then talk of all these theories as descriptions of 'the world' is

empty.²⁵

Nudged by the relativistic principal, Putnam plunges into anti-realism in regard to the physical world.²⁶ He calls it 'internal realism'. I call it 'anti-realism' because (a) it includes the claim that the physical objects around us are somehow dependent upon our conceptual schemes, and (b) it would have us abandon the idea of a unique, structured, mind-independent, physical world.

Physical realists take current theories in geology, astronomy, cosmology, and so on, to yield putative descriptions of the theory-independent, language-independent, physical world. Putnam claims such talk is 'empty' on the grounds that one cannot say how the world is 'theory-independently'. And, presumably, one cannot say how things are 'theoryindependently' because most, if not all, properties of the physical world are 'theoryrelative'. The idea that many categorical properties are theory-relative is certainly plausible. But this possibility is perfectly compatible with there being a large number of claims about the physical world that are not theory-relative.

Think again of Hera and Zeus. Putnam gives them a way of seeing that some of the claims generated by their respective conceptions of the world are theory-relative. But in doing this it looks as though he also provides them with ways of seeing that some of their claims are flat-out true, or flat-out false. Here are the relevant principals:

(A) Any claim that is true in the context of at least one correct theory and doesn't change to false, or indeterminate, upon passing to any other such theory is flat-out true.

(B) Any claim that is false in the context of at least one correct theory and doesn't change to true, or indeterminate, upon passing to any other such theory is flat-out false.

Are there such claims? I take Putnam to deny it. Nevertheless, they seem plentiful. For instance, consider the claim that our planet is roughly spherical. This is a flat-out fact, and one that must be expressible, and explicable, in the terms provided by any deep, accurate, and reasonably complete, account of the physical world. How could this claim become false or indeterminate when properly transposed into the terms of a correct theory of the world?

On the other hand, suppose someone says our planet is cubical in shape (and means it in the obvious sense). The claim is flat-out false. If the transposition of this claim from the context of our present theories of physics, geology, and astronomy to an alien context changes its truth-value, then that transposition misrepresents the claim.

Zeus and Hera help each other escape both the naïveté of atomistic realism and "the ubiquity of conceptual relativity."²⁷ Each enables the other to distinguish at least some of his or her own theory-relative beliefs from those that are, or at least appear to be, flat out true. In effect, both acquire a number of meta-level, non-theory-relative, beliefs about their own beliefs. I assume they can arrive at complete agreement in regard to the significance of their agreements and 'disagreements' about the physical world. And again, this meta-level assessment of their beliefs sharpens and improves their understanding of the physical world. There is less chance of their mistaking features of their own conceptual schemes for features of the world.

Putnam presumably would deny the possibility of benefiting in this realist way from 'conflicts' between optimal theories. It's as if he thought Hera and Zeus had only two

options in regard to their 'disagreements'. They can be 'metaphysical realists', and go on insisting that, at most, one of them is right (which seems naive). Or they can give up the idea that there is a single, theory-independent, physical world, about which they disagree. 'Sophisticated' physical realism seems preferable to either one of these options.

To summarize, there are two interrelated extensions, or corollaries, of the relativistic principle. First, The principal itself suggests the possibility of finding 'flat-out', nontheory-relative, truths about the physical world. Second, the principal seems to make it possible for rational beings who are theorizing about the physical world to see that some of their own beliefs are theory-relative, while others (perhaps) are not. The principal thus points the way to a 'sophisticated' form of physical realism.

Before we move on to Putnam's second line of attack, I want to say a bit more about 'objects'. Ever since his departure from full-scale realism, Putnam has held that 'objects' are 'theory-relative'. In explaining this idea he has sometimes used the 'cookie cutter' metaphor. As late as 1980 he said things like the following:

'Objects' do not exist independently of conceptual schemes. We cut up the world into objects when we introduce one or another scheme of description.²⁸

It was an interesting mistake. Putnam himself came to hold that the 'cookie-cutter' model is unacceptable.²⁹ Let's look at its defects.

The story goes something like this: Prior to our theorizing, the physical world is like a smooth-textured sheet of dough. Our concepts cut the dough into cookies. Obviously, these cookies didn't exist before we cut them out. And, if our conceptual scheme had been radically different, we would have produced an entirely different batch. The 'objects' in our world are thus, at least partly, 'mind-dependent'.

Putnam's subsequent objection to the metaphor goes deep. Look again at the Democritean world.

How many objects are there? According to what Putnam calls 'Carnapian language' there are three, a, b, and c. But, according to 'the Polish logician' there are seven: a, b, c, a+b, a+c, b+c, and a+b+c. [The Polish logician is Lezniewski, who counts 'mereological' sums of individuals as individuals.] Putnam thinks there are various senses of 'object' at work - as many senses as there are ways of counting 'objects'.³⁰ Which 'language' we choose is up to us - a matter of convention. But it's The World that determines how many 'objects' there are, given a particular sense of 'objects'. That, Putnam says, is a matter of 'external fact'. Carnapian 'language' cuts the world-dough into three cookies, the Lezniewskian into seven. But what is the dough? What are its parts? Any definite answer simply endorses one particular sense of 'object' at the expense of all others.

The Cookie Cutter Metaphor denies (rather than explaining) the phenomenon of conceptual relativity.³¹

Putnam urges us to be more whole hearted in our acceptance of conceptual relativity:

... go one step farther: take the position that one may either treat Carnap's Version

as 'correct' and interpret the Polish Logician's Version as a *façon de parler*, ... or treat the Polish Logician's Version as 'correct' and interpret Carnap's Version as a language in which the range of the individual variables is restricted to atoms (as suggested by the Cookie Cutter Metaphor). That is, take the position that one will be equally 'right' in either case. Then you have arrived at the position I have called 'internal realism'!

What is wrong with the notion of objects existing 'independently' of conceptual schemes is that there are no standards for the use of even the logical notions apart from conceptual choices. 32

A physical realist would see the situation a bit differently. Once one has grasped the point that there are an indefinite number of ways of counting up basic 'objects' in the Democritean world, one should not 'treat' the Carnapian Version as 'correct' and the Polish Logician's Version a mere *façon de parler*. What one should hold is that this is strictly a matter of convention. There is no 'correct' or 'incorrect' Version to be found. The conclusion Putnam draws seems false. Aren't there all sorts of things that do not depend upon conceptual schemes? Let me take a far-fetched example. A 'hoak' is an object made up of an oak tree and the nearest horse (i.e. the horse nearest to the tree). Notice that one and the same horse can be part of several hoaks. Each oak is part of one, and only one, hoak; but some horses wander around as separate and complete entities in themselves.

There are millions of hoaks. Did I create them? Of course not. Hoaks have existed for thousands of years. I created the **name** 'hoak', and the **concept**. Suppose the concept had never been cooked up. Would there have been hoaks? Surely there would. I conclude that hoaks are **not** concept-dependent, language-dependent, things.

Putnam himself is fully aware of this kind of concept/language-independence.

That the sky is blue is causally independent of the way we talk; for, with our language in place, we can certainly say that the sky would still be blue even if we did not use color words ... And the statement that the sky is blue is, in the ordinary sense of "logically independent," logically independent of any description that one might give of our use of color words. For these reasons, I have avoided stating the thesis of conceptual relativity as a thesis of the dependence of the way things are on the way we talk. ...In any sense of 'independent' I can understand, whether the sky **is** blue is independent of the way we talk.33

My suggestion is that the existence or non-existence of hoaks is equally independent of the way we talk (and think). That is to say, hoaks are objects whose existence, or nonexistence, does not depend on the existence or non-existence of conceptual schemes, or on the role they play in someone's conceptual scheme.

Perhaps, in a sense, Putnam would agree. "Yes", he might say, "with your language in place, you can certainly say that hoaks would still exist even if there were no language and no concepts. Nevertheless, hoaks only exist in relation to language and thought!"

Second Attack

Putnam believes that 'metaphysical realism' (and, I presume, physical realism) faces a serious difficulty because of a link between it and the 'brains in a vat' hypothesis. Here is one instance of the linkage claim:

THE WORLD is supposed to be independent of any particular representation we have of it - indeed, it is held that we might be **unable** to represent THE WORLD correctly at all (e.g. we might all be 'brains in a vat', the metaphysical realist tells us).

The most important consequence of metaphysical realism is that truth is supposed to be radically non-epistemic - we might be 'brains in a vat' and so the theory that is 'ideal' from the point of view of operational utility, inner beauty and elegance, 'plausibility', simplicity, 'conservatism', etc., **might be false**. 'Verified' (in any operational sense) does not imply 'true', on the metaphysical realist picture, even in the ideal limit.³⁴

The alleged linkage seems to present a threat because Putnam has previously argued that we **cannot** be 'brains in a vat' - the very idea is incoherent.³⁵

Why should physical realists hold that we may all be brains in a vat? Putnam hints at two distinct lines of thought. First, physical realists are committed to the idea that the physical world is independent of any particular representation we have of it. So, the reasoning goes, they must think that even our 'ideal' representation might, in fact, be false. As it stands, the reasoning is defective. Zeus and Hera might be realists about the world and yet know that their understanding of it couldn't be hopelessly wrong - after all, they made it. The fact that the physical world is as it is independently of how Hera and Zeus represent it to themselves, does not mean that it must be possible that the best representation available to them should be incorrect.

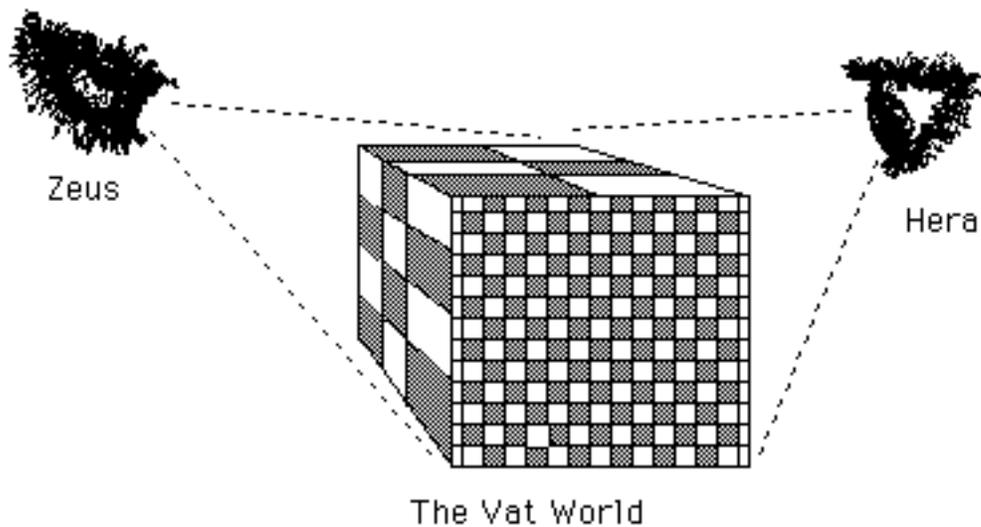
The second line of thought is that physical realists need some way of distinguishing their view from 'Peircean realism' - the notion that a scientific theory is true in so far as it approximates the theory that would be held at the ideal limit of scientific inquiry.³⁶ We need to show that verification of a theory, even in the ideal limit, does not entail that the theory is **true**³⁷

Must we also show that **we ourselves** might, for all we know, be in a set-up that makes Peircean verification miss the mark? It is difficult to see why that should be necessary. And, in any case, why do we need the brains in a vat? Wouldn't it be enough that human science might encounter some limit that would prevent it from becoming altogether true? For instance, the disparity between the way the human mind works and the fundamental nature of wave/particles might be such as to make it impossible for us to get the story absolutely right.

In spite of this fussing about Putnam's argument, the brains in a vat hypothesis deserves investigation. What should physical realists say about those brains? Let me begin by recasting the story.

Once upon a time, Hera and Zeus caused the Big Bang. Of course they did it because

they wanted 'children' like us. They thought the best procedure would be to let us come into existence on a suitable planet by random variation and natural selection. The right sort of Big Bang would eventually produce the right sort of planet. [The waiting period didn't bother them - they could work on mathematics, and have some nice chats.] Back at the planning stage, this wasn't their only option. For instance, they could have imagined a Big Bang just like the one that actually occurred, and imagined the subsequent unfolding of a physical universe like ours, the formation of a planet like Earth, and the evolution of creatures like us. On that basis, they could have made exact replicas of brains like ours. (Obviously, if this had been how things were, our universe wouldn't have existed. Hera and Zeus would have been making replicas of possible brains.) The duplicate brains might have been created in proper temporal order, and sustained in a huge array of vats. Their input experience (generated by an unbelievably sophisticated computer) could have reproduced ordinary human experience. If Hera and Zeus had chosen this course, the whole physical world might have been an enormously complex, well-engineered, self-sustaining, vat-brain-computer cube.



We naturally assume that the basic physical objects, forces, and properties, of Vat World would be those of our world. But this need not be the case. The Vat World might be quite different - something stranger and more improbable than our Quantum Mechanical world, or something more straight-forward. The 'water' in the vat, and in the cells of the brains, might be XYZ, rather than H₂O.

It's an interesting world. The brains 'see' 'mountains' and 'trees'. They 'hear' the 'twittering'

of 'birds'. They 'talk' to each other in languages that have evolved over hundreds, or thousands, of years and are syntactically equivalent to English, Chinese, Arabic, and so on. But what sorts of things do they 'say' in these languages? What, if anything, do they believe about their world?

The 'Cartesian' theory is that the brains think they live and move among mountains, trees, twittering birds, super-markets, and so on. They 'talk' about these alleged things.

According to this theory, they believe they are rational animals, living on, or near, the surface of a beautiful blue planet. In short, they say and believe more or less the same things we say and believe.³⁸ The idea is far from absurd, given the fantastic background provided by the story. It is justified, in part, by the intentions underlying the brains design and the design of the in-put generating master computer. Their structures and modes of operation are explained by those intentions. The brains were intended to be, and are, exact analogs of the brains of certain natural creatures that might have existed (namely us). On this basis, it makes sense to claim that such and such a region of a particular 'brain' is the 'olfactory' region, and that certain input to the brain is, say, the smell of cinnamon. That's how it would have been in the creatures that served as a model.

As I have said, on the Cartesian interpretation the brains would believe roughly the same things we do. Of course they would usually be wrong, while we are by and large approximately right. The crucial difference is 'external' to us and would be to them. It's a matter of how things are in the world that exists independently of our beliefs, linguistic practices, and so on, or of theirs. That is to say, we are employing an externalist-realist notion of what it is that determines truth and falsity in regard to the physical world.

Isn't it clear that this interpretation will fit all the relevant facts? How could it fail? By hypothesis, the input/output relations between the brains and the master computer are, in some abstract sense, the same as the relations between our brains and the world around us. And the brains are, structurally, and in their various internal operations, just like our brains.

Let's pretend that Hera habitually interprets the language and thought of the brains in the Cartesian way. She tends to think they are massively deceived, and it disturbs her. She suspects that she and her consort are not treating their 'children' properly.

Zeus disagrees. He 'hears' the language and thought of the brains in a different way.

"Look", he says to Hera, "Suppose some brain thinks 'I'm a brain in a vat'. How, on your view, did the brain's term 'vat' get hitched to vats?"³⁹

"Well", says Hera, "no doubt the brain got the usual training when it was young. It was, so to speak, shown tubs, cisterns, barrels, and, I suppose, vats."

"So to speak!" Zeus fumes, "The brains have never seen, or tactually explored, anything whatsoever. And, in fact, there are no such things as tubs, cisterns, or barrels - never have been."

"Ah. Now I see what's wrong." Hera replies calmly, "You are assuming that the brains' teachers only show them things that actually exist - individuals, properties, or relations, in the Vat World. But, on my view, the proper context of interpretation isn't the real world at all. It's that other possible world - the one we use as a model."

"We use it as a model. The brains don't know anything about it. You seem to think the brains go to school in that non-existent world." says Zeus.

"But it's our model; our use of it has made it the right context for understanding their experience, their thought, their training and communication. That's why I say they were,

so to speak, shown barrels and vats." Hera responds.

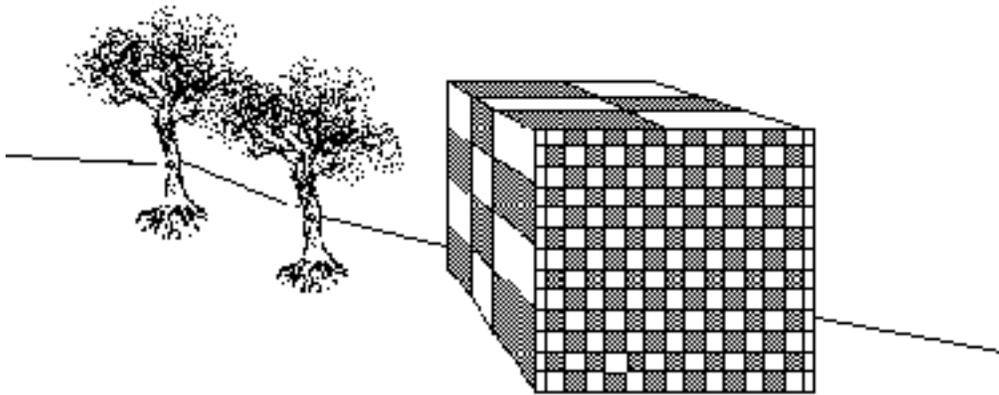
Hera proposes a particular account of what it is that determines the meaning of the terms in the languages of the brains and the contents of their beliefs. It is partly a 'causal' theory and partly a matter of the intentions of the world's creators.

Zeus favors a different account. I don't know exactly what it is; but it's a theory that connects putatively referential terms like 'mountain', 'tree', 'vat', and so on, to things in The Vat World - dispositions in the master computer for example, or features of the abstract model that governs the computers selection of input to the brains. By virtue of these connections, the brains are not massively mistaken.

Let's assume that Zeus' theory can be worked out in detail, and that it fits all the relevant facts about the brains, their behavior, the in-put they receive, and so on.⁴⁰

[Putnam believes, or, more likely, just believes that 'metaphysical realists' ought to believe, that there are several acceptable theories of the sort Zeus favors:

... when the brain in a vat ... thinks 'There is a tree in front of me', his thought does not refer to actual trees. On some theories ... it might refer to trees in the image, or to the electronic impulses that cause tree experiences, or to the features of the program that are responsible for those electronic impulses. These theories are not ruled out ... for there is a close causal connection between the use of the word 'tree' in vat-English and the presence of trees in the image, the presence of electronic impulses of a certain kind, and the presence of certain features in the machine's program. On these theories the brain is right, not wrong in thinking 'There is a tree in front of me.'⁴¹



Is Zeus right, or is Hera? Are they both wrong? Under our present assumptions, the interpretation of Vatish 'speech' and thought is underdetermined by the relevant facts. Should we conclude that what the brains think and mean when they 'talk' or think about 'trees' and so on is 'theory relative'?

I think it would be preferable to say that the brains 'speech' and thought about 'trees' is indeterminate in meaning and reference (and, as a consequence, frequently indeterminate

in truth-value.) There are at least two 'partial interpretations'. According to one, the brains are 'speaking' and thinking about (non-existent) trees. According to the other, they are 'speaking' and thinking about dispositions in the master computer. Consequently, it would be a mistake, sort of, to hold that they are 'speaking', or thinking, of (non-existent) trees, or that they aren't.

Beliefs and claims that are true on both Hera's interpretation and Zeus' are true. Beliefs and claims that are false on both interpretations are false. Beliefs and claims that are true on one interpretation and false on the other are indeterminate in truth-value. [This does not mean that there is some inaccessible fact of the matter.]

My suggestion is that Zeus and Hera should shift to this double-minded way of thinking about what the brains 'say' and believe. Given the context in which the brains operate, and the prima facie 'correctness' of both Zeus' interpretation and Hera's, the 'partial interpretations' approach seems closer to the complex truth than either of the original interpretations. In addition, it brings Hera and Zeus into agreement in regard to the brains language and thought.

[When the brains think about 'trees', what they have in mind is not, exactly, a pattern in the computer, and not exactly trees, and not a disjunctive 'tree-or-pattern' idea. It is, precisely, a thought that can be 'partially interpreted' in each of these ways - a **smudgefication** of the two.]

Of course it doesn't follow that **our** thought and talk about 'trees' is indeterminate in meaning and reference. Our brains operate in a context that makes it possible for us to think and talk about trees, and to be right about them much of the time.

The up-shot is that what may be physically identical brains, receiving physically identical in-put, with long histories of identical activity, can have beliefs, and make claims, that differ in content. Putnam said it long ago: meaning just ain't in the head.⁴²

Suppose a large number of the brains are seriously interested in science. Or, at least, that's how it looks on the 'Hera' style interpretation. On this reading, the brains have made great progress in the study of chemistry, physics, astronomy, geology, biology, and so on. They are to a large extent right, or would be right, in regard to the non-existent world that Hera and Zeus have imagined. Of course the brains tend to be hopelessly wrong about The Vat World (astronomy!).

How does the 'Zeus' style of interpretation handle this activity? From this perspective, the brains are viewed as exploring abstract patterns in the master computer's net-works - and doing it well. They are certainly not trying to understand the physics, chemistry, and so on, of the physical world that surrounds them.

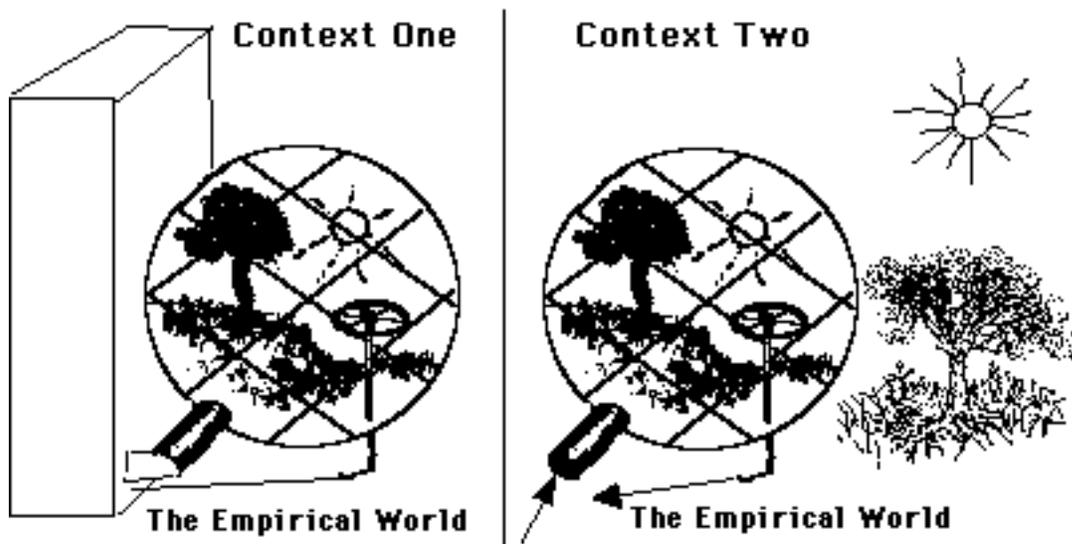
On both readings, the brains are not doing, cannot succeed in doing, what most human scientists hope to do. On one interpretation, the ideal Peircean limit of their investigations would presumably yield a correct theory of a non-existent world, and a hopelessly bad theory of the world they actually inhabit. On the other interpretation, the ideal limit yields a correct representation of part of the master computer's system for generating input.

One of the interesting features of The Vat World is that it affects the brains' beliefs and claims in two radically different ways. First, it is one of the factors that determines the content of the brains' beliefs and the meaning of their claims. Second, in many cases it is the factor that determines the truth, falsity, or indeterminacy, of particular beliefs and claims.

These 'externalist' ideas are far from trivial in their consequences. Let me sketch a bit of historical background. To put it very crudely, Descartes, Kant, and many others, have held that the content of belief, and immediate experience, are, in some sense, independent of their 'transcendental' or 'metaphysical' context. A creature's experience, and beliefs, might well be the exactly the same whether it is a disembodied spirit in a non-physical world, a mind/brain in The Vat World, or a mind/brain in the skull of a bipedal primate living in the suburbs.

Descartes thought, at least for a while, that he might be an incorporeal thinking-thing in the clutches of a deceiving demon. He would have rejected the suggestion that the content of his beliefs is, to some extent, dependent upon whether he is, or is not in that predicament. As he saw it, the content of his beliefs is determined entirely by factors within himself - his innate concepts (for instance, the concept of God).

Kant's view is in some important ways similar. The 'world' epistemically accessible to us is the 'empirical world', generated, so to speak, by the interaction of our raw input, our perceptual procedures, and our basic concepts. These concepts and procedures spring from hard-wired human predispositions and are not context-relative.⁴³ Where does the 'input' come from? And who, or what, has the 'concepts'? These are questions about the **transcendental context** of the 'empirical world' - questions about things we can never know.



Suppose some dogmatic naturalist asserts that we are rational animals who come into existence by chance, live out our brief lives, and are obliterated.

... to meet [such an assertion] we can propound a transcendental hypothesis, namely, that all life is, strictly speaking, intelligible only, is not subject to changes of time, and neither begins in birth nor ends in death; that this life is an appearance only, that is, a sensible representation of the purely spiritual life, and that the whole sensible world is a mere picture which in our present mode of knowledge hovers before us, and like a dream has in itself no objective reality; that if we could intuit ourselves and things **as they are**, we should see ourselves in a world of spiritual beings, our sole and true community with which has not begun through birth and will not cease through bodily death - both birth and death being mere appearances.⁴⁴

Kant's claim is that this may be our real situation, our transcendental context. On the other hand, he thinks, the context might be the sort of thing described by the naturalist. And there are many other such possibilities. Speculative reason leaves us completely in the dark in regard to these 'metaphysical' matters. But note that the 'sensible world' (i.e. the 'empirical world') would be exactly the same in all such contexts - no difference discernable from 'inside', and none discernable from 'outside' either - in fact, no difference, period.

Phenomenology is, or was, pure description of the human 'empirical world' - no 'metaphysics' allowed. But the whole idea of an 'empirical world', in anything like the Kantian sense, is loaded with questionable metaphysical presuppositions. In particular, it presupposes that the nature of our collective experience, the meaning of our language, and the content of our beliefs, would, so to speak, be unaffected by a shift of our 'empirical world' from one transcendental context to another. We now have externalist reasons to reject, or at least be very suspicious of, that presupposition. If meaning just ain't in the head, it just ain't in the 'empirical world' either.

The Kantian 'empirical world' may well be a myth much like the 'Cartesian Theater' of the mind.⁴⁵ Of course, one must be careful not to throw out babies with the bath-water. For example, there needn't be anything bogus about trying to describe the way the physical world is perceived, or experienced, by various kinds of creatures including ourselves.

Before we abandon the brains in a vat, I should point out that Putnam's brains inhabit a different world than mine. He insists that all the sentient beings in his world are in the vat - no Hera and no Zeus.

The differences between his world and mine aren't trivial. The intentions of Hera and Zeus connect The Vat World to a possible world like our own. Putnam makes this impossible. Putnam's world springs into existence by pure chance, or, perhaps, has

always existed - it wasn't designed. These features add to the difficulty of interpreting what is going on in it. What facts support the claim that such and such input to a 'brain' is associated in some way with a certain odor?

The problem can be made worse by replacing the 'brains' with electronic analogs of brains. On this plan, the whole world is electronic - no vats, no protoplasm. Might there be, here, a sentient electronic entity savoring the odor of cinnamon?

As we have seen, Putnam insists (or insists that 'metaphysical realists' should insist) that the brains in his world 'talk' and think about things that actually exist in that world - states of the master computer and so on. I think this insistence is more plausible in regard to his world than mine.⁴⁶ Suppose he is right. In that case he has provided yet another possible 'metaphysical context' for the brains to occupy, and, presumably, another interpretation of their beliefs and languages.

FootNotes to Putnam on physical realism

1 See Putnam's, *Meaning and the Moral Sciences*, Routledge & Kegan Paul, Boston, 1978, p. 129.

2 Hilary Putnam, *Mathematics, Matter and Method*, Philosophical Papers Vol. 1, second edition, Cambridge University Press, Cambridge, 1979, p.vii.

3 Ibid. pp. vii-viii.

4 However, I reject the Dummettean idea that realism in regard to a domain is roughly equivalent to the view that typical statements of the domain are either true or false (not indeterminate).

5 See Hilary Putnam's, *Reason, Truth and History*, Cambridge University Press, Cambridge, 1981, p. 49. Claims (2) - (4) are taken, more or less verbatim, from this same page.

6 See Putnam's "Reply to Two Realists", *The Journal of Philosophy*, Oct. '82, p. 575. I quote the relevant passage later in this chapter.

7 This feature of 'metaphysical realism' is stressed in "Realism and Reason", *Meaning and the Moral Sciences*, pp. 123 - 138. See, e.g. p. 125.

8 See *Meaning and the Moral Sciences*, p. 124.

9 There are differences that deserve emphasis. Some of the 'atoms' Russell and Moore had in mind were plainly mental (e.g. pleasures and pains). And all of them were 'simple'.

10 *Reason, Truth and History*, p. 49. On p. 50, Putnam suggests that the 'externalist' presupposes "... a No Eye view of truth - truth as independent of observers altogether." This seems nearer the mark than the earlier claim.

11 In his post-realist phase, Putnam seems to dismiss the picture. "There is no God's Eye point of view that we can know or usefully imagine; there are only the various points of view of actual persons reflecting various interests and purposes that their descriptions and theories subserve." Ibid. p. 50. The language here is a bit guarded. Putnam does not deny the existence of God. Nor does he say God has no 'point of view' - no conceptual scheme in terms of which he (God) comprehends the physical world.

12 *Reason, Truth and History*, p. 73.

13 Michael Scriven once wrote: "There is no such thing as a 'complete description' of the contents of a laboratory. Not because there are an infinite number of atoms in the room, but because there is an indefinite number of things in the room and an indefinite number of relations between them and always an indefinite number of answers to the question, 'What is the best description?' ("Definitions, Explanations, and Theories", in *Minnesota Studies in the Philosophy of Science*, Vol. II, ed. by H. Feigl, M. Scriven, and G. Maxwell, University of Minnesota Press, Minneapolis, 1958, p. 127.) Scriven attributes the point to Stuart

Hampshire. See also Putnam's *Representation and Reality*, The MIT Press, Cambridge, Mass., 1988, pp. 110-113.

14 Hartry Field, "Realism and Relativism", *The Journal of Philosophy*, Oct. '82, pp. 553-554.

15 Hilary Putnam, "Reply to Two Realists", *Ibid.*, p. 575.

16 See *Meaning and the Moral Sciences*, pp. 130-133.

17 *Ibid.* pp. 131-132.

18 *Ibid.* pp. 50-51.

19 One of the interesting things about *Meaning and the Moral Sciences* is that in it we can watch Putnam shift from physical realism to anti-realism. In the John Locke lectures, (i.e. up to p. 80) he is a realist. In Part Four, "Realism and Reason", he has made the transition.

20 *Ibid.* p. 132.

21 *Ibid.*

22 Both Theory One and Theory Two say there exists a set of objects, O, such that the members of O constitute The World. They agree about that. The disagreement is about which set it is that does the job. (In this note, I treat 'O' as a variable. In the text, it's a name.)

23 Putnam sketches a theory of the 'one line' world according which it is composed of only denumerably many objects. See *Meaning and the Moral Sciences*, pp. 132-133.

24 Field writes: "In my view ... realists - that is, metaphysical realists¹ - should drop all pretext of being separated from their opponents by some issue about the existence of a uniquely correct theory of the world. A unique and mind-independent world is enough." ("Realism and Relativism", p. 554). Metaphysical realists¹ are people who hold that the world consists of some fixed totality of mind-independent objects. Field apparently takes this later belief to be more or less equivalent to belief in a unique, mind-independent, world.

25 *Meaning and the Moral Sciences*, p. 133.

26 Of course the relativistic principal wasn't the only factor pushing Putnam towards anti-realism. There are at least two other important lines of argument. Then too, as I say at the end of the paper, Putnam thinks that 'metaphysical realism' leads naturally to physicalism, which, in turn, leads naturally to acceptance of the alleged 'fact/value' distinction. He regards this as a disastrous consequence.

27 "Truth, Activation Vectors and Possession Conditions for Concepts", *Philosophy and Phenomenological Research*, Vol. LII, No. 2, June 1992, p. ??

28 Putnam, *Reason, Truth and History*, p. 52.

29 Putnam attacks the cookie-cutter metaphor in *The Many Faces of Realism* (Open Court, LaSalle, Illinois, 1987, pp. 33-36) and repeats the attack in *Representation and Reality* (The MIT Press, Cambridge, Mass. 1988, pp. 113-114).

30 *The Many Faces of Realism*, p. 33.

31 *Ibid.* p. 34.

32 *Ibid.* pp. 35-36.

33 Putnam, "Truth, Activation Vectors and Possession Conditions for Concepts", *Philosophy and Phenomenological Research*, Vol. LII, No. 2, June 1992, p. ?????

34 *Meaning and the Moral Sciences*, p. 125.

35 See *Reason, Truth and History*, Chapter 1.

36 This line of thought can, perhaps, be found in *Meaning and the Moral Sciences*, p. 125.

37 See my "Cartesian Semantics", *Canadian Journal of Philosophy*, Vol. 17, N.1, March 1987, pp. 63-70.

38 Proper names for entities other than the brains themselves are a problem. Clearly they cannot have beliefs about the North Atlantic or Disney Land. It's a matter of reference failure. Presumably they think (mistakenly) that there are places and things that satisfy such and such descriptions.

39 See *Reason, Truth and History*, p. 51.

40 I'm not absolutely certain that this can be done. One possible test for the adequacy of Zeus' theory might be that it yields a meaningful and reasonable alternative to every claim or belief that is meaningful and reasonable on Hera's reading. But then consider, for example, an 'utterance' Hera interprets as meaning that the brains might all be brains in a vat. What could it mean on Zeus' interpretation? Presumably it means something like "We may all be G23 type patterns in the master computer." Don't the brains (on this interpretation) know that G23 patterns are incapable of thought? But, if they know this, how can they find the suggestion that they themselves may be such patterns tempting and worthy of exploration?

41 Reason, Truth and History, p. 14. The talk about 'trees in the image' suggests that Putnam believes in something like the Kantian 'empirical world'.

42 The slogan can be found in Putnam's "The Meaning of Meaning", Mind Language and Reality, Cambridge University Press, Cambridge, 1975, p. 227.

43 The 'basic concepts' are those purely positive ideas - the genuine predicates - to be found in our innate concept of God. Kant apparently accepted the standard Leibnizian view on this matter. See, for example, Kant's Lectures on Philosophical Theology, Translated by Allen W. Wood and Gertrude M. Clark, Cornell University Press, Ithaca, 1978, pp. 33-34.

44 Immanuel Kant, Critique of Pure Reason, A 780, B 808, translated by Norman Kemp Smith.

45 The Kantian 'empirical world' and its 'transcendental context' seem inextricably linked together. If the 'empirical world' is bogus, its allegedly inaccessible, unknowable, context is suspect too.

46 Nevertheless, I am not at all certain that Hera's initial style of interpretation is out of the running. Even in Putnam's world, the master computer organizes the brains input in a way that 'makes sense', and is predictable, if it is read as mimicking a world like ours. And the hypothesis of such mimicry provides a way of understanding the brains 'design'. [Note added March 7, 2010: I am now even more inclined to defend Hera's interpretation. Surely, for instance, the brain's talk about time refers to real time, and some of their up-to-date 'physics' links space (or 'space') and time? If their physics isn't to be regarded as nonsense don't we have to go with Hera?]

47 Hartry Field, "Realism and Relativism", Op.Cit. p. 562.