

Truth and Correspondence: Putnam's 'Paradox' Revisited

Shamik Dasgupta
Draft of January 2026

1. Introduction

Hilary Putnam's notorious model-theoretic argument (1977, 1982) purported to establish that beyond various epistemic and pragmatic constraints that a theory may or may not satisfy, there is no further constraint of truth considered as some kind of "correspondence with reality". Call this view *alethic anti-realism*. David Lewis (1984) objected that Putnam's argument was based on a simple mistake, and he went on to develop a particular account of truth as "correspondence with reality" in terms of his notion of naturalness. But setting aside the details of Lewis' particular account, the core of his disagreement with Putnam lay in his view that *there is* a further constraint of truth considered as some kind of "correspondence with reality", over and above the various epistemic and pragmatic constraints. Call this *alethic realism*.

Lewis' objection has been sufficiently influential that Putnam's argument is no longer taken seriously in many quarters. I will argue, to the contrary, that it is Lewis who was mistaken. Properly understood, Putnam's argument is much stronger than Lewis realized; strong enough, in particular, to withstand Lewis' critique. Indeed, I believe that Putnam's argument is much stronger than *Putnam* realized! In this paper I will explain why.

Before I start, two notes about the terminology just introduced. First, 'alethic anti-realism' is my term, not Putnam's. Putnam used the result of his model-theoretic argument to develop a position he called 'internal realism', but what internal realism *was*, exactly, was never very clear. Among other things, it seemed to involve (i) an epistemic analysis of truth, specifically in terms of idealized warranted assertibility; and (ii) a theory of warrant on which our standards of warrant are historical products that evolve over time to reflect our interests and values.¹ But note that neither (i) nor (ii) follows from what I called 'alethic anti-realism' above. This is obvious with (ii), since alethic anti-realism says nothing about the nature of warrant. What might be less obvious is that (i) doesn't follow either. To see why it doesn't, notice that even if there is no further constraint of truth as correspondence over and above epistemic-pragmatic constraints, *per* alethic anti-realism, one needn't think that truth reduces to epistemic or pragmatic ingredients, *per* (i). Instead, one could try offering an account of truth in neither epistemic, pragmatic, nor correspondence-like terms; or one could even, in principle, be an eliminativist and reject the notion of truth wholesale.² I'm not suggesting that either option is attractive; the point is just that alethic anti-realism is one thing and an epistemic analysis of truth such as (i) is another. If you like, the former is a negative claim (there is *no such thing* as truth considered as correspondence) while the latter is a positive claim (truth *is* idealized warranted assertibility).

¹ I discuss (i) and (ii) further, with textual evidence from Putnam, in section 5.

² An example of the former might be deflationism about truth. For a view in the vicinity of eliminativism, see Gamester (2023).

Still—and this is the second note—alethic anti-realism arguably implies *something* about idealized warrant, namely that “it makes no sense to suppose that an empirically ideal theory, as verified as can be, might nonetheless be false because the world is not the way the theory says it is”. Indeed, that was how Lewis characterized the conclusion of Putnam’s model-theoretic argument (Lewis 1984, p. 221). Here Lewis was following Putnam, who often characterized his conclusion similarly.³ This is understandable insofar as something along these lines does appear to follow from alethic anti-realism as characterized above. For if *there is no such thing as* “correspondence with reality”, over and above the various epistemic-pragmatic constraints that a theory may or may not satisfy, then an ideal theory can hardly count as false in virtue of *failing* to correspond with reality. In *that* sense, at least, it makes no sense to suppose that an ideal theory is “false because the world is not the way the theory says it is”. But of course the same goes for any theory, ideal or not: given alethic anti-realism, it makes no sense to suppose that *any* theory is false in virtue of failing to correspond with reality!⁴

Nonetheless, one might argue that alethic anti-realism implies something in this ball-park concerning ideal theories specifically. Assume that truth—if there is such a thing—must either be understood as “correspondence with reality” or in epistemic-pragmatic terms. Then if there is no such thing as “correspondence with reality”, the argument would go, then *either* there is no such thing as truth *or* truth must supervene on epistemic and pragmatic features. On the first horn there is presumably no such thing as falsity either, and on the second horn falsity would follow from the *absence* of certain epistemic-pragmatic features. Either way, it would make no sense to suppose that a theory with all the epistemic and pragmatic goodies we can imagine might nonetheless count as *false*. Perhaps that is why Putnam’s conclusion is often put in these terms. Nonetheless, the argument I just gestured at raises many thorny issues, not least how to understand talk of an ‘ideal theory’; better, I think, to drop such talk and focus on alethic anti-realism as characterized above. This is my approach here.⁵

So much for terminology. As I said, I will argue that Putnam’s model-theoretic argument is much stronger than both Lewis and Putnam recognized. *Contra* Lewis, the argument does not rest on a simple mistake; and *contra* Putnam, it succeeds (if it succeeds at all) in refuting not just alethic realism, but internal realism too!

Still, one point on which I agree with Lewis is that, like him, “I am not sure how well I understand Putnam. Sometimes, different things he says seem to point in different directions” (Lewis 1984, p. 222). Indeed, I am not sure how well *Putnam* understood Putnam! For he sometimes presents his argument in a way that *would* be vulnerable to Lewis’ objection, while a close reading of other passages suggests that his argument is in fact much stronger than that. Lewis handles this uncertainty by reconstructing the argument in his own voice. I will do likewise and reconstruct an argument that departs from the letter of both Putnam and Lewis. But

³ For example, in (Putnam 1977) he sets up the argument thus: “let T_1 be an ideal theory, by our lights. ... we can imagine T_1 to have every property *except objective truth*—which is left open—that we like. E.g., T_1 can be imagined complete, consistent, to predict correctly all observation sentences... to be “beautiful”, “simple”, “plausible”, etc. The supposition under consideration [that he goes on to reject] is that T_1 might all this *and still be* (in reality) *false*” (1977, p. 485).

⁴ You might think that if there is no such thing as “correspondence with reality”, then an ideal theory *should* count as false in virtue of not so corresponding. Fine; the point remains that the same goes for any theory, ideal or not.

⁵ Indeed, towards the end I will explain why the argument just sketched doesn’t go through.

it is Putnamian in spirit, and I will offer some textual evidence that suggests it was what Putnam had in mind—in some of his moods, at least.

2. A Putnamian argument

The contention, as I said, is that beyond various epistemic-pragmatic constraints that a theory may or may not satisfy, there is no further constraint of truth considered as some kind of “correspondence with reality”. To show this, let us assume, with Putnam, that it is *sentences* that are in the first instance true or false, and that a theory is a set of sentences (closed under logical implication, if you like, though this will play no role in what follows). This is a simplification, no doubt, but a harmless one.

What then could it be for a sentence to be true in the sense of “corresponding with reality”? One answer takes the general form of a Tarskian theory of truth. Very roughly—and slurring over all sorts of technical and metaphysical details—this consists in two components. The first component specifies a “correspondence” relation between sub-sentential parts of sentences and bits of reality. For example, the relation might associate each name with an object, each predicate with a property or satisfaction-condition, and so on (I’m using “bits of reality” broadly and loosely here to include satisfaction-conditions). And the second component then defines truth recursively in terms of this correspondence relation. A simple base clause might look something like this:

(TC) A sentence of the form ‘a is F’ is *true* if and only if the object that corresponds to ‘a’ has the property that corresponds to ‘F’.

This is just the roughest of sketches, of course, but you get the idea: parts of the sentence “correspond” to certain bits of reality, and the sentence is then true insofar as those bits of reality fit together in the right way. It is in that sense the sentence is true in virtue of “corresponding to reality”. Let us assume that any account of truth as correspondence will take something like this form; this is another simplification, but again, a harmless one.

As I understand it, the core of Putnam’s argument is simply that *there are too many correspondences*. Correspondence relations are cheap, so that given *any* (consistent) set of sentences there will be *some* way in which it corresponds to reality. This is the first premise:

Premise 1: For any (consistent) set of sentences S, there is some correspondence relation between its sub-sentential parts and bits of reality such that every sentence in S counts as true *per* (TC).

And the second premise is that all correspondence relations are on a par; none of them is distinguished over the rest as the “correct” way to correspond with reality.

Premise 2: All correspondence relations are on a par.

The putative constraint that a theory must “correspond to reality” is therefore vacuous: *any* (consistent) theory is *guaranteed* to satisfy the constraint! And a vacuous constraint is no

constraint at all; hence there is no constraint of truth, understood as “correspondence with reality”, over and above the various epistemic-pragmatic constraints on theories. As Putnam put it, “the trouble... is not that correspondences between words or concepts and other entities don’t exist; the trouble is that *too many* correspondences exist” (1981, p. 72-3).

That’s it. That’s the argument. The rest of the paper will unpack it and explain how powerful it is.

But first, where’s the model theory? Putnam called the argument ‘model-theoretic’, but what role does model theory play in the above? As I see it, Putnam just used model theory to establish Premise 1. To this end he assumed that theories can be formalized in a first-order language L . Now, in model theory an *interpretation* of L is defined as a set D , the “domain”, together with a valuation function v from non-logical symbols of L into D . Specifically, v maps every constant in L to an element of D , every 1-place predicate of L to a subset of D , every 2-place predicate of L to a set of tuples of D , and so on (these valuation functions will be the “correspondences”). The notion of *truth in an interpretation* is then defined recursively à la Tarski. For example, a base clause would say something like this:

(TI) A sentence of the form ‘ Fa ’ is true-in-interpretation $\langle D, v \rangle$ iff $v(‘a’)$ is an element of $v(‘F’)$.

This is, of course, just (TC) stated in more formal terms, and with the relativization to different interpretations (i.e., correspondences) made explicit. Now, in model theory it is convenient to think of domains as sets of abstract objects—real numbers, for example, or pure sets. By contrast, let an interpretation *in the actual world* be an interpretation whose domain is a set of “worldly” objects such as tables and chairs, apples and pairs, and so forth. What Putnam shows is that any consistent theory—in the syntactic sense of ‘consistency’—will have an interpretation in the actual world in which it is true. That is his rendition of Premise 1.

So rendered, the proof of Premise 1 is straightforward. Putnam states it explicitly in “Realism and Reason” (1977), p. 485. If a theory T is syntactically consistent, the completeness theorem ensures that there is an interpretation $\langle D, v \rangle$ in which T is true. To find an interpretation “in the actual world”, just take any domain $D_{@}$ of worldly objects with the same cardinality as D , take any one-one function f from D to $D_{@}$, and let $v_{@}$ be the valuation function from L to $D_{@}$ induced by applying f to v . (That is: if $v(‘a’) = \text{element } d \text{ of domain } D$, then $v_{@}(‘a’) = \text{element } f(d) \text{ of domain } D_{@}$; if $v(‘F’) = \text{sub-set } S \text{ of } D$, then $v_{@}(‘F’) = \text{sub-set } f(S) \text{ of } D_{@}$; and so on.) By construction, the resulting interpretation $\langle D_{@}, v_{@} \rangle$ makes T true given that the original interpretation $\langle D, v \rangle$ did. The only question is whether there is a domain $D_{@}$ of worldly objects with the same cardinality as D , but this is guaranteed *if* the actual world is of some infinite cardinality.⁶

⁶ Why? If D is finite, just let $D_{@}$ be any finite subset of worldly objects of the same size. And if D is of some infinite cardinality, the Lowenheim-Skolem theorem guarantees that for any other infinite cardinality, T is also made true by an interpretation whose domain D^* has this other infinite cardinality. So just pick any such interpretation whose domain has the same infinite cardinality as the actual world, and start the argument again.

Admittedly, this last condition means that the constraint of “corresponding with reality” isn’t *entirely* vacuous. For it means that a theory that is only true in interpretations with infinite domains would fail to meet the constraint if the actual world is finite. But as Lewis observed, this result is “incredible enough”, for it means that a theory “can misdescribe the world only by getting its size wrong” (1984, pp. 229-30). So long as alethic realism was understood to maintain that “corresponding with reality” requires more than just getting its size right, the result we have is good enough. In what follows, I will take this caveat to be understood.⁷

Thus, an “interpretation in the actual world” is just a “correspondence” in my earlier sense: an assignment of bits of language to bits of reality (broadly construed). So, if there’s an interpretation in the actual world in which the theory is true, then the theory “corresponds to reality” in my earlier sense—at least, it does relative to *that* correspondence relation. So, Putnam’s model-theoretic rendition of Premise 1 is just a more formal way of saying what I glossed earlier: that for any (consistent) theory there will be *some* way in which it “corresponds to reality”.

That much is uncontroversial. Borderline trivial, in fact. To illustrate, consider the sentence

“Some pigs fly”,

which we can formalize in L as

$(\text{Ex})(\text{Px} \ \& \ \text{Fx})$

What Premise 1 says is that there is *some* interpretation of ‘P’ and ‘F’ in the actual world in which it’s true. And *of course* there is: just map ‘P’ to the set of *birds* and ‘F’ to the set of things that fly! All Premise 1 does is generalize this banal point to any consistent set of sentences.

But what of it? That’s not the *correct* interpretation of ‘P’ and ‘F’, you might say—at least, not if they are to stand for *our* words ‘pigs’ and ‘fly’. The interpretation doesn’t reflect what those words *mean*. ‘Pigs’ *refers* to pigs, the thought is, and the correct interpretation must surely reflect *that*. Putnam puts this point by asking whether it’s the *intended* interpretation of our words. But I think “intended” is misleading in various respects, so I’ll stick with “correct”.

The worry, then, is that while it’s (borderline) trivial that any consistent theory will “correspond to reality” relative to *some* word-world correspondence relation, as Premise 1 states, it is *not at all* trivial that it will do so relative to *the correct* one. And so Putnam has not yet established that the constraint of “corresponding to reality” is vacuous—at least, not when understood as the constraint that a theory must correspond to reality *relative to the correct correspondence relation*. This is where Premise 2 comes in. It says that there is no such thing as the “correct” interpretation: all interpretations—i.e., correspondence relations—are on a par.

⁷ Though I think there is, ultimately, no need for the caveat. The need for it arises only because Putnam treats the meanings of the logical constants—specifically, the quantifiers and the identity sign—as fixed. It’s only because of *that* that it makes sense to think that T might be true only in interpretations with infinite domains. But why treat the logical constants as fixed? I think it’s just a convenient assumption that can ultimately be relaxed. But there is no room to develop this thought, so I will grant the assumption for the sake of argument.

If that's right, the putative constraint that a theory must "correspond to reality" does indeed become vacuous. Any (consistent) theory is guaranteed by Premise 1 to correspond to reality relative to *some* correspondence relation, and according to Premise 2 there is no further question as to whether it's *the correct* correspondence relation. Thus the idea that "corresponding with reality" is a genuine constraint on theories collapses in heap.

3. Just more theory

Clearly, everything hangs on Premise 2. Why accept it? Putnam says frustratingly little at this point. Shortly after giving the proof of Premise 1 sketched above, he offers a brief remark to the effect that distinguishing a 'correct' correspondence relation would require "direct (and mysterious) grasp of Forms" (1977, p. 487). He makes similar remarks again in "Models and Reality (1983b), for example that it would require "non-natural mental powers of directly 'grasping' forms" (p. 1). But it's not obvious (to put it mildly!) how this talk of Platonic *forms* connects with Premise 2 as stated. I will suggest how we might connect these dots later on, but for now we are left to reconstruct an argument for Premise 2 ourselves.

Lewis finds in Putnam's writings the materials to construct the following argument:

- (2.1) Our use of language is not sufficient to fix a correct interpretation (i.e., correspondence relation).
- (2.2) Nothing else could possibly help fix the correct interpretation.

Here, 'our use of language' can be understood to include the sum total of everything we say (and think). So understood, (2.1) follows from the model-theoretic reasoning used to establish Premise 1. For how *could* our use of language fix a correct interpretation? The idea would presumably be that the correct interpretation is that which makes what we say *true*. This is what Lewis called 'Global Descriptivism'. But the model-theoretic reasoning shows that this constrains almost *nothing*. So long as what we say is consistent, there is guaranteed to be an interpretation $I_{@}$ in the actual world that makes everything we say true (Premise 1). And if $I_{@}$ makes everything we say true, then any isomorphic image of $I_{@}$ in the actual world—i.e., any interpretation induced from $I_{@}$ by a one-one function on the domain of $I_{@}$ —will also make everything we say true. Thus, if $I_{@}$ assigns our word 'pigs' to the set of pigs, then for *any* other set S of worldly objects with equal cardinality—a set of birds, for example—there will be *some* interpretation that assigns 'pigs' to S and *also* makes everything we say true! Our use of language therefore constrains almost nothing about what its "correct" interpretation is.⁸

Note, as a consequence, that it's no good trying to pin down the correct interpretation of our speech with more speech. One might for example try saying

"By 'pigs' I mean the animal species that satisfies the following definition:..."

or

⁸ Chapter 2 of *Reason, Truth, and History* (1981) is, in effect, an argument along these lines.

“By ‘pigs’ I mean *that kind of thing*” (frantically gesticulating at a pig).

But the argument above shows that this is futile. There will always be *some* way of interpreting these definitions and gestures that maps ‘pigs’ to pretty much any set of worldly objects you like and *still* makes everything we say true!

So, for truth considered as “correspondence with reality” to be a non-vacuous constraint on theories, something over and above our use of language must help fix the ‘correct’ interpretation of our words. That’s (2.1). And this extra something must help fix a correct interpretation in such a way that some of what we say can turn out to be *false*—otherwise the constraint of “corresponding with reality” would be vacuous.

Lewis proposed a theory of reference that, he claimed, does exactly this. His core idea is that not all interpretations are equal; some are more natural than others. Here he appealed to his metaphysical thesis that some worldly items—properties, sets, relations, functions, objects, whatever—are perfectly natural, and that amongst the rest some are more natural than others. The property *green* is more natural than the property *grue*, for example, and perhaps the property *electron* is perfectly natural. We can then say that one interpretation is more natural than another to the extent that the worldly items it assigns our words are, overall, more natural than the other. Thus, if *green* is more natural than *grue*, then (all else equal) an interpretation that assigns a word to *green* is more natural than one that assigns it to *grue*. Lewis presented this as a theory of *reference*, which we can state as follows:

(R) The referent of a term is given by the interpretation I_R which best balances the two constraints of (i) fit with use, and (ii) naturalness.

Here, ‘fit with use’ just means making what we say (and think) true. So, the interpretation I_R will be one that best balances the two constraints of making what we say (and think) true and assigning our words to worldly items that are natural. Now, Lewis used ‘reference’ interchangeably with ‘intended interpretation’—or as we’re saying, ‘correct interpretation’—so he presumably thought of this as a theory of what makes I_R the ‘intended’ (or ‘correct’) interpretation. To foreshadow, I think this was a sleight of hand that masked Lewis’ mistake. But we’ll come to that later.

For now, note that these two constraints can pull in opposite directions. To make *everything* we say true might require an interpretation that assigns our words to very *unnatural* entities; if so, an interpretation that scores higher with respect to naturalness will make some of what we say false. So, it is not guaranteed that everything we say is made true by the interpretation I_R that achieves the best balance. The property of being made true by I_R —or being ‘true-in- I_R ’, as I’ll call it—is therefore a *non-vacuous* constraint on theories. To be true-in- I_R is to correspond to reality relative to the correspondence relation of *reference*, as defined by (R), and it is *not* guaranteed that every consistent theory will correspond to reality in *that* sense. Thus we appear to have exactly the notion of truth as “correspondence with reality” we are looking for: over and above various epistemic-pragmatic constraints on theories, *there is* a further, non-vacuous constraint of “corresponding with reality”, namely the constraint of *being true-in- I_R* .

What does all the work here is the second factor (ii) that goes beyond fit with use, but Lewis's specific proposal involving naturalness is by no means the only way to achieve the same effect. One might instead (or in addition) propose a *causal* factor, to the effect that the referent of a term is given by an interpretation that respects certain causal constraints. Suppose the presence of pigs tends to cause me to say "That's a pig" while the presence of birds doesn't. Then one might propose that it's because of *that* that 'pig' refers to pigs and not birds. One can easily imagine other possible constraints that involve how we're *disposed* to use words, how we'd use them in certain *counterfactual* circumstances, *teleological* facts about the function of our words, and so on. Just survey the literature on reference for a menu of options. Whatever the details, the point is that any proposal along these lines rejects (2.2): in addition to our use of language, *there is* something else—naturalness, causation, whatever—that helps fix reference.

But Putnam thought that any such proposal is a sham. Why? Because it's "just more theory", he complained. When Lewis asserts (R), this is just another thing he says. It's just more *use of language* on Lewis' part. It's now part of the sum total of everything Lewis says, so we can feed it all back into the argument for (2.1): there is guaranteed to be *some* interpretation of everything Lewis said that makes it *all* true, including (R)! This interpretation will assign 'naturalness' to some set of worldly items, 'reference' to some word-world relation, and so on; and we can be assured that this interpretation makes everything Lewis says, *including (R)*, true. The idea that we have a *non-vacuous* constraint on theories collapses once again.

Of course, Putnam can make the same move *whatever* further constraint *C* on reference one offers instead of naturalness. Regardless of whether the constraint involves naturalness, causation, dispositions, counterfactuals, or what have you, Putnam can reply that the constraint is just more theory. Lewis summarized the move thus:

"Constraint *C* is to be imposed by accepting *C*-theory, according to Putnam. But *C*-theory is just more theory, more grist for the mill; and more theory will go the way of all theory" (1984, p. 225).

Thus we have (2.2): over and above our use of language, nothing else can possibly help fix its correct interpretation.

It is at this point that Lewis cried foul:

"*C* is *not* to be imposed just by accepting *C*-theory. That is a misunderstanding of what *C* is. The constraint is *not* that an intended interpretation must somehow make our account of *C* come true. The constraint is that an intended interpretation must conform to *C* itself" (1984, p. 225).

This was Lewis' distinction between *satisfying C-theory* vs *conforming to C*. An interpretation satisfies (R) if it makes (R) true, and if that is all we require of the 'correct' interpretation then Putnam is right that we've made no progress at all. But Lewis' proposal was *not* that the correct interpretation must make (R) true; it was that the correct interpretation is that which best balances the two constraints of use plus naturalness! Putnam's mistake, according to Lewis, is that he simply overlooked this distinction.

As Lewis emphasized, this diagnosis of Putnam’s mistake rests on a certain picture of how language works. It’s not available on what he called a “purely voluntaristic” picture on which language is a tool over which we have *complete* mastery. On this picture, it is up to us what our words refer to *and also how reference is fixed in the first place*. As Lewis put it,

“Take your favourite theory of reference. Let us grant that it is true. But let us ask: what makes it true? And the tempting answer is: *we* make it true, by our referential intentions. We can refer however we like—language is a creature of human convention—and we have seen fit to establish a language in which reference works *thus*. Somehow, implicitly or explicitly, individually or collectively, we have made this theory of reference true by stipulation” (1984, p. 226).

If we want reference to be fixed by use plus naturalness, we can achieve this *by stipulation*—by (implicitly or explicitly, individually or collectively) including (R) in the sum total or what we say and think. If instead we want reference to be fixed by a causal constraint, we can achieve this by stipulation too—all we have to do is scream the constraint loud enough, as it were. On this voluntaristic picture of reference, (R) is indeed just more theory and Putnam was right to call Lewis’ proposal a sham.

But Lewis rejected voluntarism. On his view, reference is *non-voluntaristic* in the sense that how it is fixed is independent of anything we say or do about it:

“Reference isn’t just something we do. What we say and think not only doesn’t settle what we refer to; it doesn’t even settle the prior question of *how* it is to be settled what we refer to” (1984, p. 226).

If reference is fixed by use plus naturalness, *per* (R), that’s not because we *say* it is. It is, rather, a fact that holds independently of anything we say (or think, or do) that reference is so fixed. We could scream ‘causation’ as loud as we like and reference would *still* be fixed by use plus naturalness à la (R). On this non-voluntaristic picture of reference, (R) is not “just more theory” and Lewis was right to cry foul at that point.

And there matters rest. Lewis’ non-voluntaristic picture of reference became widely accepted in many quarters. And so, abstracting from the details of his specific theory of reference (R), it became received folklore in those quarters that Putnam’s model-theoretic argument was based on a simple mistake.

4. Primitive normativity

But I think this is not the end of the story. Lewis’ non-voluntarist picture of reference is not enough to defuse Putnam’s argument. What Lewis needs, in addition, is *primitive normativity*.

To see what I mean, let’s grant Lewis his non-voluntarist picture of reference for the sake of argument. And let’s grant his specific theory of what fixes reference, (R), as a concrete example. To see why this is not enough to defuse Putnam’s argument, recall what (R) says:

(R) The referent of a term is given by the interpretation I_R which best balances the two constraints of (i) fit with use, and (ii) naturalness.

Think of this as describing a relation of correspondence between words and world, the ‘reference’ relation, which is fixed by use plus naturalness as above. This relation then induces the property of being true-in- I_R —i.e., the property of “corresponding to reality” relative to the correspondence relation of reference. As we saw, this property is a non-vacuous constraint on theories in the sense that there is no guarantee that a consistent theory will possess it.

But on its own, this is not enough to diffuse Putnam’s argument. Why? Because it remains that *there are too many correspondences!* To see this, note that if there’s a property of *perfect naturalness* that, let’s suppose, green has and grue lacks, there’s also a property of *perfect graturalness* that grue has and green lacks. After all, properties are cheap—there’s a property for every set. At least, this was Lewis’ own view of properties, and likewise for relations and functions: there’s a relation for every set of n-tuples and a function for every univalent relation. Lewis posited perfect naturalness to distinguish a select few of these entities from the rest, and this works *to an extent*: it distinguishes those that are natural from the rest. But if properties are cheap, there is also a property of perfect graturalness that distinguishes a *different* set of entities from the rest. Green may be perfectly natural, but grue is perfectly gratural!

So, if the relation of reference is fixed by use plus naturalness, *per* (R), there is also a relation of *greference* fixed by use plus graturalness in a precisely analogous manner:

- (G) The *greferent* of a term is given by the interpretation I_G which best balances the two constraints of (i) fit with use, and (ii) graturalness.

Reference is *one* correspondence relation between words and world, and greference is another. Thus a word will “correspond” to more than one worldly entity at once: it will *refer* to one thing and *grefer* to something else at the very same time! This might sound odd, but it is no stranger than the fact that I am both the *son of* Carol and the *father of* Elowyn at the same time. *I* can stand in different relations to different things, and our words are no different.

And just as there was the property of being true-in- I_R , defined in terms of reference, there is also an analogous property of being *true-in- I_G* defined analogously in terms of greference—i.e., the property of “corresponding to reality” relative to the correspondence relation of greference. Moreover, this property is a non-vacuous constraint on theories in just the same sense that truth-in- I_R is, namely there is no guarantee that a consistent theory will possess it. There’re *both* non-vacuous constraints; they’re just *different* constraints.

So *for all Lewis says* there remain too many correspondences! To be true-in- I_R is to “correspond to reality” relative to the correspondence relation of *reference*; to be true-in- I_G is to “correspond to reality” relative to the correspondence relation of *greference*, and so on. More generally, given *any* consistent set of sentences there will be *some* way in which it corresponds to reality—there will be *some* correspondence relation C , that is, between sub-sentential parts and bits of reality such that every sentence in the theory counts as true-in- I_C ! This was Premise 1, remember, and Lewis’ theory of reference (R) does *nothing* to change that. And for all we’ve been told so far, all these correspondence relations are on a par—that was Premise 2, and Lewis’ theory of reference (R) does *nothing* to change that either. All (R) does is *describe* (or *analyze*)

one of those correspondence relations, but it doesn't follow from *that* that it's distinguished over other correspondence relations as the 'correct' way to correspond to reality. After all, (G) describes a different correspondence relation, and that obviously doesn't imply that it's the 'correct' way to correspond to reality either!

To illustrate, consider our word 'pigs'. It refers to pigs, of course, but it grefers to something else—birds, let's suppose. Thus, the sentence

“Some pigs fly”

is not true-in- I_R , but it *is* true-in- I_G !⁹ Here I don't mean that we could have used the word 'pigs' differently, in such a way that it would have referred to birds—that is just the banal point that reference depends (to some extent) on use. No, I mean that the word 'pigs' *in our mouths, given how we English speakers actually use it*, grefers to birds. Hence the sentence “Some pigs fly”—again, *in our mouths, given how we actually use the expressions*—is true-in- I_G ! The sentence does not correspond to reality relative to the relation of reference, of course, but it *does* correspond to reality relative to the relation of greference! Thus the constraint that what we say should “correspond to reality” collapses into vacuity once again: we could say “Some pigs fly” and *still* satisfy the constraint!

Note that all this is so even while we assume Lewis' non-voluntaristic picture of reference. We are assuming that reference is fixed by use plus naturalness, *per* (R), not because we *say so* but because that's how the reference relation works. Given this assumption, Putnam's “just more theory” move is indeed illegitimate, just as Lewis said. But what goes for reference goes for greference too: just as what we say and think doesn't settle how reference is fixed, it doesn't settle how *greference* is fixed either. Greference is a relation fixed by use plus *graturalness* not because we say so but because that's how the greference relation works.

Thus, even if we grant Lewis everything he wanted—his non-voluntaristic picture of reference, and his specific theory of reference (R)—it remains that there are too many correspondence relations (Premise 1) all of which are on a par (Premise 2). Which was, remember, Putnam's core insight all along:

“the trouble... is not that correspondences between words or concepts and other entities don't exist; the trouble is that *too many* correspondences exist” (1981, p. 72-3).

My point is that this core argument survives Lewis' critique entirely unscathed—indeed, Lewis doesn't even so much as *engage* with it!

What Lewis needs, clearly, is a further claim that breaks the symmetry between all these correspondence relations and distinguishes one of them—reference, presumably—over the rest as the 'correct' way to correspond with reality. And the key point here is that this further claim is a *normative* claim. If this isn't obvious, recall that there are *gazillions* of non-vacuous correspondence constraints: there's the constraint of corresponding with reality relative to the correspondence relation of reference, the constraint of corresponding with reality relative to the

⁹ More accurately: there is some word-world relation C such that 'Some pigs fly' is true-in- I_C .

correspondence relation of preference, and so on. *Any* (consistent) theory is guaranteed to satisfy *some* of those constraints and not others. What Premise 2 says is that all those constraints are on a par: i.e., that satisfying any one of them is as “good” as satisfying any other. So, what Lewis needs is the further claim that they are *not* all on a par: that the constraint one *ought* to satisfy is the constraint of corresponding with reality relative to the relation of reference; i.e., that one *should* aim at truth-in- I_R . That’s the further claim, and the normativity inherent in it should now be clear. There are of course arguments to be had about how to *put* this normative claim, e.g. whether it’s most perspicuously expressed in terms of ‘should’, or ‘ought’, or ‘right’, or ‘correct’, or ‘good’, or what have you. Thus instead of saying that one *should* aim at truth-in- I_R , we could instead say that being true-in- I_R is the *right* way to correspond with reality, or that reference is the *correct* correspondence relation, or that reference gives the *correct* interpretation of our words. But setting those details aside, the claim is that reference, and therefore truth-in- I_R , has a distinguished normative status. As I’ll put it,

(NS) Reference is *normatively significant* over other correspondence relations.

This is what Lewis needs to reject Premise 2 and neutralize Putnam’s argument.

Why did Lewis miss the need for (NS)? I think it’s because he used ‘reference’ interchangeably with ‘intended interpretation’—or as we’re saying, ‘*correct* interpretation’. By doing so, he gave the impression that (NS) follows from his theory of reference (R) *by definition*. Put otherwise, if ‘reference’ is used synonymously with ‘correct interpretation’, then his theory of reference (R) *just is* a theory of what the *correct* correspondence relation is: it’s the one fixed by use plus naturalness. Lewis is free to use these terms as he likes, of course; the substantive point is that his theory then has *two* components, a descriptive component and a normative component. The descriptive component just describes (or analyzes) one relation of correspondence between words and world, which we can label ‘reference’—that is how I’ve been reading (R). And the normative component then adds that this correspondence relation gives the *correct* interpretation of our words and is therefore the *correct* way to correspond with reality—this is (NS). If you want, you can combine these two components into one package by using ‘reference’ synonymously with ‘correct interpretation’. It’s still a free country and I’m not here to police language. But I *am* here to insist that both components are needed to neutralize Putnam’s argument.

What of it? Why not just add (NS) explicitly to Lewis’ package? Well, you could! But here’s the kicker: normative significance must be a *sui generis* property, not identical or reducible to anything else. If reference is normatively significant, that is, this cannot hold purely in virtue of non-normative facts about the material cosmos (or the divine realm, or what have you). If that’s right, then (NS) implies *normative primitivism*, the view that normativity is an extra, *sui generis* component of reality over and above the material (and divine) world order.¹⁰

This may not be obvious. Why couldn’t reference be normatively significant in virtue of some non-normative property it possesses? Well, it could. But if so, we’ll find irreducible

¹⁰ This is sometimes called “normative non-naturalism” or “robust realism” about the normative, but I prefer the label “normative primitivism”. Enoch (2011) is perhaps the most explicit contemporary representative of this position.

normativity elsewhere. To see this, suppose one tries saying that reference is normatively significant over *grreference* in virtue of being the correspondence relation fixed by *naturalness*. OK, but of course *graturalness* stands to *grreference* in just the same way that *naturalness* stands to reference. So, for reference to be normatively significant over *grreference*, *naturalness* must already be normatively significant over *graturalness*—if not, there is nothing to distinguish *naturalness* and the relation of reference it fixes over *graturalness* and the relation of *grreference* it fixes. Now, if the fact that *naturalness* is normatively significant is fundamental, we’re done: we have a fundamental fact about normative significance. If instead you say that some property of *naturalness*, N_1 , confers this significance on *naturalness*, well, there’ll be another property G_1 that stands to *graturalness* just as N_1 stands to *naturalness*... and we’re off on a regress. Thus, if reference is to have normative significance over *grreference*, *either* the regress must end in some property N_i that has normative significance fundamentally, not in virtue of anything else; *or* the chain of properties each of which is normatively significant in virtue of the next goes on forever. Either way, the property of being normatively significant is not reducible or explicable in terms of something else: it is an extra, *sui generis* component of reality over and above the material (and divine) world order.

Notice, as a consequence, that nothing about *us* can confer normative significance on anything either. Suppose for example that it *serves our interests* to aim at truth-in- I_R and not truth-in- I_G . Then one might try suggesting that this is what makes truth-in- I_R normatively significant over truth-in- I_G . But this gets us no further. To see this, think of ‘interests’ as some gradable property of ours, such as health, which is promoted by truth-in- I_R —i.e., we tend to have more of it when we accept theories that are true-in- I_R , not true-in- I_G . Well, in addition to interests we also have *grinterests*, a gradable property that is promoted by truth-in- I_G . This is a strange property to which we pay little attention, but since properties are cheap we must have *some* such property.¹¹ Thus, if our interests single out the property of truth-in- I_R , then, equally, our *grinterests* will single out the property of truth-in- I_G . For truth-in- I_R to be normatively significant over truth-in- I_G , therefore, *interests* must already have normative significance over *grinterests*. If not, there’s nothing to distinguish our interests and the property of truth-in- I_R they pick out as significant over our *grinterests* and the property of truth-in- I_G they pick out; neither is distinguished over the other as the “correct” way to correspond with reality.

This is why Putnam’s argument is so much stronger than Lewis recognized. To neutralize it, it’s not enough to offer a non-voluntaristic theory of reference such as (R). What’s needed is primitive normativity.

Indeed, we can now reconstruct Putnam’s argument in its strongest form. Recall that it is based on two premises.

¹¹ To see this, note that the property of having certain interests can be represented by a mapping from persons to real numbers that increases as a function of, say, the person’s health. Or more precisely: a function from possible worlds to such mappings that increases as a function of the person’s health in that world. So, to find the gradable property of *interests** just take a function from possible worlds to such mappings that typically increases as a function of how much they maximize total pleasure. If properties are cheap, this function represents *some* gradable property of persons (this is just the idea that every set of possibilia corresponds to a property, adjusted to the case of gradable properties).

Premise 1: For any (consistent) set of sentences S, there is some correspondence relation between its sub-sentential parts and bits of reality such that every sentence in S counts as true *per* (TC).

Premise 2: All correspondence relations are on a par.

Premise 1 is uncontroversial, we saw earlier. And it's now clear that the rationale for Premise 2, properly understood, is *not* the argument (2.1) and (2.2) that Lewis reconstructed—that took us down the rabbit-hole of looking for a theory of *reference*, which we now know is beside the point. Rather, Premise 2 is properly understood as the claim that no correspondence relation is *normatively significant* over the rest, which I suggest is best supported as follows:

Premise 2(a): Normative significance is a primitive normative property (i.e., not reducible in non-normative terms).

Premise 2(b): There is no primitive normativity.

I just gave the rationale for Premise 2(a) above. I haven't offered any support for Premise 2(b), but that is beyond the scope of this paper. Rather, what I've tried to show is that the only way out of this Putnamian argument is to embrace primitive normativity. If that's right, anyone who *rejects* primitive normativity—which includes Lewis!—must accept the conclusion: beyond various epistemic-pragmatic constraints on theories, there is no further constraint of truth considered as some kind of “correspondence with reality”.

That's the Putnamian argument as I reconstruct it. But is it *Putnam's* argument? Personally, I don't much care: I'm more interested in what works than what Putnam said. But it might bother you that premises 2(a) and 2(b) stray from Putnam's preferred vernacular, for he appears to say little about normativity in the relevant texts. If so, have another look at *Reason, Truth, and History* (1981), p. 46, where he asks

“Given that there are many ‘correspondences’ between words and things... what *singles out* one particular correspondence *R*?”

Nothing does, says Putnam, who concludes thus:

“It seems as if the fact that *R* is reference must be a *metaphysically unexplainable* fact, a kind of primitive, surd, metaphysical truth” (p. 46).

Replace ‘metaphysical’ with ‘normative’, and you have Premise 2(a). And insofar as he clearly rejects this ‘primitive’ truth, you have Premise 2(b).¹²

¹² This reconstruction of Putnam's argument also resolves an interpretative puzzle in Lewis' reading of Putnam. In Lewis' reading, Putnam's argument starts by showing that Global Descriptivism doesn't suffice to fix reference—that's (2.1)—and then goes on to show that any other theory of reference is “just more theory”—that's premise (2.2). But Global Descriptivism *is* a theory of reference, so why wouldn't Putnam reject *that* as “just more theory” too? Why should Global Descriptivism get special treatment? The answer, I suggest, is that it doesn't, for the argument I just reconstructed doesn't go by way of (2.1) or (2.2) in the first place.

5. Epistemic-pragmatic constraints?

That is why Putnam's argument is much stronger than Lewis recognized. I'll now show that it's also stronger than *Putnam* recognized!

To see this, recall that Putnam used the result of his argument to develop a position he called 'internal realism'. I characterized this earlier as having two core components:

- (i) An epistemic analysis of truth, specifically as idealized warranted assertibility.
- (ii) A theory of warrant on which our standards of warrant change over time to reflect our interests and values.

Putnam was never very precise about either component, but what he did say is clear enough. Regarding (i), he said in *Reason, Truth, and History* that truth is "some sort of (idealized) rational acceptability" (1981, p. 49); and later, in *Realism With a Human Face*, that "a true statement is one that could be justified were epistemic conditions ideal" (1990, p. vii).¹³ And he articulates (ii) in *Realism With a Human Face* (1990, p. 21), where he lists five principles about warrant that include

"Our norms and standards of warranted assertibility are historical products; they evolve in time."

and

"Our norms and standards always reflect our interests and values. Our picture of intellectual flourishing is part of, and only makes sense as part of, our picture of human flourishing in general."

Thus, having argued that there is no constraint of truth considered as "correspondence with reality", Putnam appears to maintain that *there are* epistemic-pragmatic constraints on theories. We have *epistemic* standards of warrant (rational acceptability, justification) that a theory may or may not satisfy; and the relevant standards are those that play a *pragmatic* role, i.e. of reflecting our interests and values. As our interests and values change over time, so too do our standards of warrant.

Now, since our standards of warrant can change over time, it will help to understand warrant as a relative notion: a theory is warranted relative to a given epistemic standard. Examples of epistemic standards include induction, modus ponens, trusting one's senses, and other familiar standards; but also counter-induction, affirming the consequent, and other standards we reject. Thus, if all Fs observed so far have been Gs, then

"All Fs are Gs"

¹³ The specific turn of phrase I use in (i) is from his introduction to *Realism and Reason: Philosophical Papers Volume 3* (1983b, p. xviii), where he says "if assertibility (in the sense of *warranted* assertibility) is not formalizable, idealized warranted assertibility (truth) is even less so".

is warranted relative to induction, but not relative to counter-induction. Putnam's idea, then, is that our interests and values at a time fix a set epistemic standards—which, let's suppose, currently includes induction—and these standards then operate as epistemic constraints on theories: we are to accept (assert, believe) a theory only if it is warranted relative those standards.

But the problem should now be obvious. In addition to having *interests* which distinguish one set of standards, we also have *grinterests* which distinguish others. More fully, think of a standard as a function from a body of evidence to a theory, i.e. the set of sentences that meet the standard. But a function is just a set of tuples, so there are *gazillions* of such functions. Induction is one and counter-induction is another, but that's just the beginning. If induction maps our current body of evidence E to a theory that includes "Pigs don't fly", then there is also a standard of *grinduction* that agrees with induction on all our *past* bodies of evidence but which maps E to a theory that includes "Some pigs fly"! And if our *interests* distinguish the standard of induction, then we will inevitably have other properties, call them *grinterests*, that distinguish the standard of grinduction instead.¹⁴ Thus, given our current body of evidence E,

"Pigs don't fly"

is warranted relative to epistemic standards that reflect our *interests*; but equally,

"Some pigs fly"

is warranted relative to epistemic standards that reflect our *grinterests*! More generally, given *any* body of evidence, *any* theory is guaranteed to be warranted relative to *some* epistemic standard that reflects *some* property of ours! Putnam's idea that what say and think is constrained by epistemic standards that reflect something about us therefore collapses into vacuity: *anything* we say or think is *guaranteed* to satisfy this constraint!

Or at least, this is so *if* interests and grinterests are on a par. Putnam can of course avoid this result by maintaining that our *interests* and the epistemic standards they pick out are distinguished over our *grinterests* and the epistemic standards *they* pick out. He can avoid this result, that is, if he says that interests are *normatively significant* over grinterests. But to say that, he must embrace primitive normativity—in which case his argument that there is no further constraint of "corresponding with reality" collapses!

This is why Putnam's argument is so much stronger than he recognized. *If* it succeeds in showing that there is no constraint of "corresponding with reality", then it *also* succeeds in showing that there are no epistemic-pragmatic constraints on theories either!

Moreover, there is clearly nothing special about epistemic-pragmatic constraints specifically. If Putnam's argument succeeds in showing anything at all, it shows that *there are no constraints on theories at all*. Anything we say or think is "just as valid" as anything else!

6. Internal and external constraints

¹⁴ Grinterests are whatever property of ours stands to grinduction just as our interests stand to induction; since properties are cheap, we are guaranteed to have some such property.

Let's pause to reflect on what I just said. "There are no constraints at all"—isn't this *obviously* absurd? Lewis said of Putnam's argument against truth as correspondence that

"We are in the presence of paradox... It is out of the question to follow the argument where it leads. We know in advance that there is something wrong, and the challenge is to find out where" (1984, p. 221).

Isn't this *doubly* so if the argument purports to demolish not just truth as correspondence, but *any* constraint on theories whatsoever?

I don't think so. Let me explain why. The argument shows that anything we say or think is guaranteed to satisfy *some* epistemic standard on a par with all others, to "correspond with reality" relative to *some* correspondence relation on a par with all others, and so on. Nonetheless, we do of course have a *practice* of evaluating what we say and think—of policing each other (and ourselves) in speech and thought, so to speak. Precisely what this practice *is* is an empirical question I won't try to answer, but it could very well involve the kinds of constraints that Putnam and Lewis talked about all along! For example, it may be the case, as a matter of empirical fact, that we tend to accept a theory only if it is warranted relative to induction and other epistemic standards that reflect our interests. It may be that we tend to criticize others (and ourselves) when they accept theories that do not meet these epistemic standards, praise them when they accept theories that do, and so forth. It may even be the case—again, as a matter of empirical fact—that this practice tends to result in us accepting theories that are true-in- I_R and rejecting theories that are not. Our practice might turn out, that is, be one that tends to "track" truth-in- I_R . In virtue of this practice, we might even be said to be 'aiming at truth-in- I_R '!

I am of course slurring over the details of our practice. This is partly because the details are an empirical matter that are not our concern here. But it's also, and more importantly, because *any* description of our practice will be contentious for precisely the reasons we've been discussing. Let induction* be a standard that overlaps with induction on all bodies of evidence encountered so far and diverges in future cases. Then if we have tended so far to accept a theory only if it is warranted relative to induction, we have *also* tended so far to accept a theory only if it is warranted relative to induction*. We have engaged in *both* practices at once; or, if you like, these are two *descriptions* of our practice. We don't ordinarily describe our practice in the latter terms—but *that is just another part our practice*. So, the empirical hypothesis I asked you to entertain in the previous paragraph is better put like this: that we happen to engage in a practice of evaluating speech and thought that *we happen to describe as* 'using induction' and 'aiming at truth-in- I_R '.

To reflect this, say that C is an *internal* constraint iff we engage in a practice of evaluating speech and thought that *we describe as* being constrained by C. By contrast, say that C is an *external* constraint iff C is a constraint on speech and thought independently of our practices. An external constraint is imposed "from the outside", as it were: its status *as a constraint* has nothing to do with us or our practices. By contrast, an internal constraint is just a rather abstract description of *what we actually do*.

With this distinction in hand, we can be more exact about what the Putnamian argument shows: it shows that there are no *external* constraints on speech and thought. For truth-in-I_R to be an external constraint, for example, it would have to be normatively significant over truth-in-I_G and other correspondence-like standards; so if all such standards are on a par (Premise 2), it isn't an external constraint. Likewise for epistemic-pragmatic constraints: to say that induction is an external constraint is to say that induction (or our interests that induction reflects) is normatively significant over *grinduction* (or our *grinterests* that *grinduction* reflects). But Premises 2(a) and 2(b) imply that nothing is normatively significant; hence induction isn't an external constraint either. More generally, if nothing is normatively significant then there are no external constraints at all. *That's* what the Putnamian argument shows.

Still, this leaves untouched the obvious platitude that there are *internal* constraints on speech and thought—induction, presumably, and perhaps even truth-in-I_R! But all this *amounts to* is that we happen, as a matter of empirical fact, to engage in a practice of acceptance and rejection, praise and criticism, etc., that *we happen to describe* as 'using induction' and 'aiming at truth-in-I_R'. This may well be the case, but it's just an empirical theory about *what we do*. And as an *empirical* theory, it's not what Putnam's argument had in its sights in the first place.¹⁵

To illustrate, suppose I say

“Some pigs fly.”

Have I made a mistake? Did I say something incorrect? Internally, yes: you will of course criticize me, tell me that there is inductive evidence to the contrary, insist that it's not true-in-I_R, and so on. If I really persist, you might even have me institutionalized. But externally, I made no mistake at all! What I said satisfies *some* epistemic standard on a par with all others; is true-in-I_C relative to *some* correspondence relation C on a par with all others, and so on. Externally, an alternative practice of *praising* what I said and treating it as gospel is on a par with our actual practice of *criticizing* it! So, you can criticize me and shun me all you like, but as the Big Lebowski said, “that's just, like, your opinion, man”.

In light of this, what are we to make of Putnam's claims (i) and (ii)? Well, we can read them externally or internally. Externally, the idea is that our interests and the epistemic standards they pick out are normatively significant over our *grinterests* and the epistemic standards *they* pick out. The former epistemic standards therefore function as *external* constraints on what we say and think; constraints imposed 'from the outside'. That's component (ii). Component (i) then says that, in addition to the (external) constraint of being *in fact* warranted relative to those epistemic standards, given our current body of evidence, there is a further external constraint of *truth*—not truth considered as correspondence, but truth considered as what *would be* warranted relative to those epistemic standards were epistemic conditions ideal. That's the external reading, and we saw in the last section that the Putnamian argument *demolishes* it. If that's the reading Putnam had in mind, he didn't Putnam hard enough.

¹⁵ Putnam was clear about this at the very beginning of “Realism and Reason” (1977), which starts thus: “In one way of conceiving it, realism is an empirical theory” (p. 483). He then says in the second paragraph that this empirical theory may include the idea that our worlds “correspond” to things in the world; and in the third paragraph that this is *not* his target.

Understood internally, by contrast, (i) and (ii) are just empirical descriptions of our actual practices. That reading is left standing by the Putnamian argument. But the trouble is that when read internally, component (i)—the epistemic analysis of truth—seems *obviously wrong!* As Lewis rightly emphasized, it just seems absurd to say that there can be no unverifiable truths: as a matter of empirical fact, that is not the kind of thing we say or think! We simply *don't* ordinarily think that truth is somehow limited by our capacity to *know* it. This is no objection if (i) is read *externally*—i.e., as a suggestion of what the external constraint of truth could *be* if not correspondence with reality. For we are then in the realm of Metaphysics (with a capital M), and whether we *ordinarily* think of truth as limited by our capacity to know it is, arguably, neither here nor there. But read internally, as an empirical description of our practices, Putnam's epistemic analysis of truth seems clearly incorrect.

But no matter: if Putnam had appreciated the full power of his argument, he'd have seen that there was no need for an epistemic analysis of truth in the first place! For what his argument shows is that *external* constraints are a will-o'-the-wisp. All we can do, when thinking about constraints on what we say and think, is offer empirical hypotheses about *what we actually do*. And as an empirical hypothesis, it may indeed be that we aim at truth-in-I_R—i.e., that “corresponding with reality” relative to the relation of reference is a constraint we hold each other to in speech and thought. So, you're only left needing an epistemic analysis of truth if you only follow the Putnamian argument half-way. But if you Putnam hard enough, you come out the other side.

References

- Putnam, Hilary (1977). “Realism and Reason”. *Proceedings and Addresses of the American Philosophical Association* 50(6): 483-498.
- Putnam, Hilary (1981). *Reason, Truth, and History*. Cambridge: Cambridge University Press.
- Putnam, Hilary (1983a). *Realism and Reason: Philosophical Papers Volume 3*.
- Putnam, Hilary (1983b). “Models and Reality”. In *Realism and Reason: Philosophical Papers Volume 3*.
- Putnam, Hilary (1990). *Realism With a Human Face*. Harvard University Press.
- Lewis, David (1984). “Putnam's Paradox”. *Australasian Journal of Philosophy* 62(3): 221-236.