

Realism about Structure and Kinds

L. A. Paul

In 1976, Hilary Putnam set forth his model-theoretic argument, claiming that it showed that the semantic realist's program¹ was 'unintelligible', since it implied, *contra* the realist view, that reference is radically indeterminate.² The response, for the most part, was unsympathetic. Semantic realists argued that the only reason Putnam's argument gave us the wildly implausible conclusion that reference is radically indeterminate was because the model-theoretic argument's representation of the realist view was fatally flawed. Putnam failed to represent a crucial part of their view, claimed the realists: that the world itself constrains the reference of terms in theories. Different sorts of semantic realists characterized the constraint differently. 'Pure' causal theorists of reference argued that terms are endowed with (moderately) determinate reference in virtue of baptism ceremonies, while semantic realists of a descriptivist bent argued that interpretations for terms in the language of a theory must respect the natural kinds or properties of the world. The additional constraint, argued both kinds of realists, refuted Putnam's claims by securing the (moderate) determinacy of reference.

Although I find the conclusion that reference is indeterminate unattractive, I shall argue that the descriptivist position needs to be supplemented with a premise about the sorts of kinds, that is, the sort of *structure*, that our world includes. The need for this premise gives a counterintuitive result: the descriptivist account of reference makes the very possibility of determinate reference contingent upon the sorts of kinds that make up the world. The need for this premise can pave the way towards a new kind of skepticism about what we may and may not assume about the nature of the world.

For those who are inclined to accept Putnam's conclusions, my arguments may be sufficient to motivate the adoption of a pragmatic or coherentist theory of truth. But for those who cannot accept the idea that reference is radically indeterminate, the lesson will be that, under the

¹ The realist that Putnam calls a 'metaphysical realist' is who I am calling the 'semantic realist'.

² Putnam claimed that semantic realism was 'unintelligible' because it implied that any ideal theory was true. It followed from his argument that reference was radically indeterminate; the argument and its conclusions are discussed in section three.

descriptivist view, if we wish to preserve realist intuitions about truth we will have to embrace a characterization of reference that violates a strong intuition about how our words refer.

1. Semantic Realism

Define the semantic realist as one who believes that an account of truth relies in part upon there being a reference relation that holds between the words we use and the things in the world. This is intended to be a somewhat loose definition: to be this sort of realist one need not adopt a fancy or detailed correspondence theory of truth; one need only think that in order to have true theories about the world there needs to be a fairly determinate way in which the terms in the theories refer to parts of the external world.³ Under this view, we are able to utter true sentences and think true thoughts if there is a fairly determinate reference relation that holds between objects in the world and the words in the sentences that we use to describe them. (If there is moderate indeterminacy, this could interfere with some, but not most, of the things that we say.) For the semantic realist, since terms of our theories determinately refer, the theorems of our theories about the world can be evaluated as objectively true or false.

Now, this sort of semantic realism seems to be a fundamental assumption, accepted by many philosophers, and implicitly relied upon by many other realist views, as it provides an account of truth and reference that applies to all theories. Anyone who holds that the sentences of her favorite theory are (approximately) true in a way that does not involve a construal of truth that is pragmatic or coherentist, or is not inconsistent with a correspondence theory of truth in some other way, should subscribe to semantic realism. If the scientific, mathematical or moral realist thinks that the reason the sentences of her favorite theory are true is because the terms in the theory actually refer to properties, relations and objects in the world, then she (implicitly) adopts semantic realism by adopting her version of scientific, mathematical or moral realism.

But to make semantic realism a respectable doctrine, especially if we want to make it respectable for those who propound naturalism, we should provide an account of how determinate reference is possible and why we think the terms in our theories determinately refer. If realists are unable to give such an account, it does not prove their theories false, but seriously weakens their case against antirealists. If the semantic realist has naturalist leanings, the concession that one is simply assuming determinate reference is particularly galling.

³ Perfect determinacy may be too much to ask. The reasonable realist may be willing to accept a small amount of indeterminacy, such as the sort of indeterminacy that arises with the vagueness of predicates like 'is bald'. Moreover, as David Lewis (1984, p. 228) argues, we might reasonably accept the idea that 'rabbit' refers indeterminately to rabbits, undetached rabbit parts, rabbit-fusion, etc. Other kinds of moderate indeterminacy, however, are less palatable than these: I discuss this point in section five.

Attempts to give constitutive accounts of determinate reference have been presented by realists. Pure causal theorists of reference argue that reference obtains between words and things in the world in virtue of a causal relation established by a baptism ceremony, which fixes the reference in the original cases, and is then passed along from speaker to speaker. Pure causal theorists reject descriptions as reference-fixing. This theory runs into trouble in accounting for how we can refer to objects that we cannot perceive directly, such as numbers, moral properties, and unobservable objects. Descriptivists argue that reference obtains in virtue of descriptions of a thing in the world that associate a word with an object, property, etc. Causal descriptivism, the view I find most promising, holds that the reference-fixing descriptions are couched in causal terms.

Perhaps the most plausible account of reference would be one that cobbles together some form of casual descriptivism with the pure casual theory of reference. The pure causal theory of reference seems to provide an adequate account of how we endow many proper names with referents, while causal descriptivism seems to provide a good account of how theoretical terms in scientific theories⁴, as well as most other kinds of terms, get their reference. Neither theory is adequate by itself; the pure causal theory in particular faces many counterexamples.⁵ Since my interest is with theories that include a description of the world in terms of its fundamental physical properties, I shall focus on causal descriptivism⁶, which seems to provide the best account of how the terms in fundamental physical theory get their reference, and as such is an essential part of the hybrid view suggested above.⁷ If the terms of fundamental physics fail to refer (moderately) determinately, then our project of coming up with a suitable theory of the world already has more trouble than we can handle.

2. Putnam's 'Model-Theoretic' Argument

Putnam and the proponents of his model-theoretic argument attack the thesis that the semantic realists' program implies that our words have determinate reference. Putnam does not argue against the individual theories of descriptivism and the causal theory of reference; his is a

⁴ Lewis (1970) provides such an account.

⁵ For some very nice examples, see Peter Unger (1983).

⁶ Although my preference is for causal descriptivism, my conclusions hold for other sorts of descriptivism as well. For this reason, I will not distinguish between causal descriptivism and other forms of descriptivism in what follows, since the differences between the varieties will in no way affect the argument.

⁷ Moreover, I find the response that the causal theorist of reference must give to Putnam unappealing. The causal theorist of reference must claim that the causal theory of reference is true (and gives us enough determinacy), whether or not it can be *proved* that it is true. See footnote nine, below.

general argument that (he claims) shows that the very idea of having determinate reference for the terms of our theories under the realist program is a chimera.

In his argument, Putnam says that we can take a theory T that is 'ideal' from the point of view of operational utility, is consistent, has maximal inner beauty and elegance, 'plausibility,' simplicity, 'conservatism,' etc. , and correctly predicts all observation sentences. We stipulate, after excluding the property a realist would call 'objective truth' for the purposes of the argument, that theory T has all the other 'best properties' a theory should have. Now, a realist, on Putnam's view, argues that such a theory could nevertheless be false (Putnam 1977, p. 484). Putnam maintains that such a theory would have to come out true. This is because he thinks that we can pick a model of T that is the same cardinality as the world, then map the individuals taken as satisfying the predicates 1-to-1 directly into the world. The individuals of the model can be mapped into the world in *any way we please*. So the mapping defines relations of the model directly in the world, i.e., relations between individuals in the model are constructed in the world. We may say that the structure of the model is reflected in the world by the mapping, or that the structure of the model is isomorphic to a structure of the world. Since there are countless many ways to map the model into the world, and countless many ways to interpret the terms of the theory T , there are countless many ways for terms in T to denote. What this means, in effect, is that in each possible mapping, different arbitrary classes of objects in the world are set up as the referents of terms or predicates in the language L , and thus reference is radically indeterminate.

For example, the set of all rabbits might be the extension of the predicate 'is an R' in one interpretation for an ideal theory, but the set of all cherries might be the extension of the same predicate in another interpretation.⁸ For Putnam, both interpretations can serve equally well, as long as the claims of the theory are satisfied. Worse yet, a set made up of rabbits *and* cherries could instantiate the predicate 'is an R' in yet another interpretation. If all these interpretations can assign referents to the terms in the theory such that the claims of the theory are satisfied, then according to Putnam, all of them give us a true theory, and so each interpretation (and thus each way of assigning referents) is equally good.

Putnam argues that all of these interpretations must give us a true theory because the result of mapping the model of the theory into the world is the definition of a satisfaction relation which gives us a correspondence between the terms of the language of the theory (as interpreted by the model) and sets of pieces of the world. His claim is that this correspondence relation is as good as any other; for him, truth just amounts to there being such a correspondence. Since there are many different ways of setting up this correspondence relation, there are many

⁸ Cardinalities of the sets permitting.

different interpretations of the language of a theory such that the theory comes out 'true'. If the theory is ideal, then for Putnam we have all we need to call the theory 'true'. In effect,

'... Putnam's argument retraces the model-theoretic proof that given any consistent theory demanding a universe of cardinality c , and given any set of cardinality c , there will exist a model of the theory whose universe is that set. In virtue of this model which establishes a correspondence between the language of the theory and the set in question, the theory is true of that set (under the given correspondence).' (Merrill, 1980, p. 70)

Now, the realist will reject this notion as a notion of truth, and argue that not just any interpretation could be the correct interpretation; i.e., not just any interpretation I that satisfies the claims of the theory is such that truth under I is truth *simpliciter*. Indeed, the realist will argue, we may be so unlucky that we will have no correct interpretation, and be left with a theory that meets all of our pragmatic desiderata, conforms to all observations, yet nevertheless is false. Putnam's response to the realist is to maintain that it is ridiculous to try to privilege any of these ways of mapping the model into the world. He thinks that the realist has no legitimate way to define particular interpretations as 'intended'. For Putnam, any attempt by the realist to impose a constraint on the eligibility of interpretations to satisfy the theory would require a theory of the constraint, but this theory would be subject to the same indeterminacy as the original theory, since the model-theoretic argument applies equally well to the new addition. Putnam employs this argument against realists who respond to his argument by defending the causal theory of reference, arguing that '[h]ow "causes" can uniquely refer is as much of a puzzle as how "cat" can on the [semantic] realist picture.' (Putnam 1977, p. 486)

3. The Descriptivist Response

Unsurprisingly, realists reject Putnam's view, arguing that there is one more thing aside from theoretical constraints constraining the choice of interpretations for theories—the world itself. David Lewis (1983, 1984) is the foremost proponent of such a view; his response draws on Gary Merrill's (1980) argument that the descriptivist can rebut Putnam by arguing that the world has a unique structure.⁹ We see a contemporary version of the Lewisian view in Sider (2012)'s defense of a world with objective structure.

⁹ Realists who are sanguine about the prospects of the causal theory of reference as a complete account for the reference of the terms of our theories could follow Michael Devitt instead, who argues that the causal theory of reference is still the correct theory, whether or not we can establish it in a theory-independent way. So, in response to Putnam's argument that the causal theory of reference just adds more theory to the whole that must be interpreted, Devitt bites the bullet, and argues that the causal

Lewis's response to Putnam is simple and elegant: there are objective samenesses and differences in nature; in other words, there exist objectively natural properties. This stipulation, claims Lewis, is sufficient to give us (fairly) determinate reference. To see the beauty of this response, view Putnam's model-theoretic argument in terms of different ways of carving up the objects in the world into classes. Putnam's claim is that one way of carving up the world is as good as any other, and thus any way of grouping objects into classes is acceptable for the purposes of assigning extensions to predicates. Thus, the class of all grue things is just as legitimate a class to serve as the referent of a predicate as the class of all green things. So, for Putnam, an interpretation that assigns predicates the extensions grue and bleen is just as good as an interpretation that assigns predicates blue and green. (For the uninitiated: define 'grue' as a predicate that applies to all things examined before the year 2011 just in case they are green but to other things just in case they are blue. Define 'bleen' as a predicate that applies to all things examined before the year 2011 just in case they are blue but to other things just in case they are green.)

Lewis disagrees, and argues that there exist objective distinctions among ways in which objects of the world may be grouped into classes, and that the world imposes a requirement upon interpretations: they must respect these objective groupings. Among all the countless things and classes that there are, most are miscellaneous, gerrymandered, ill-demarked. Only an elite minority are carved at the joints, so that their boundaries are established by objective sameness and difference in nature (Lewis 1984, p. 227, and Sider 2012). In other words, out of the countless number of different sets that can be created out of the parts of the world, only some of these classes reflect the objective way the world is broken up. The realist thinks these classes are the only eligible referents for terms in the language of a theory. Thus, some ways of carving the beast of reality are objectively correct ways.

'In the simplest case, suppose that the interpretation of the logical vocabulary somehow takes care of itself, to reveal a standard first-order language whose nonlogical vocabulary consists entirely of predicates. The parts of the world comprise a domain; and sets, sets of pairs,..., from this domain are potential extensions for the predicates. Now suppose we have an all-or-nothing division of properties into natural and unnatural. To have a property is to be a member of a class. Say that a set from the domain is *eligible* to be the extension of a 1-

theory of reference is true (and gives us enough determinacy to eliminate errant interpretations), whether or not it can be *proved* that it is true. Although Devitt's response is sufficient to rebut Putnam, I think it could be counted as a violation of the naturalist methodology of which Devitt, like most who argue for the causal theory of reference, is a proponent.

place predicate iff its members are just those things in the domain that share some natural property; and likewise for many-place predicates and natural relations. An eligible interpretation is one that assigns none but eligible extensions to the predicates. A so-called intended interpretation is an eligible interpretation that satisfies the ideal theory.' (Lewis 1983, pp. 371-2)¹⁰

The extent to which the interpretation respects the elite or natural classes of the world is the extent to which the interpretation can be taken as correct or intended. According to Lewis, if the class of green things is natural and the class of grue things is not, then the class of green things can serve as the extension of a predicate and the class of grue things cannot. So Lewis believes that there is a constraint upon eligible interpretations: the constraint that a property or relation should be assigned by the model to a predicate when its image under the mapping m is *really* one of the properties and relations among objects in the world. This constraint is imposed by the world, not by us, so it is not subject to Putnam's 'just more theory' argument. A realist is one who believes that the world has natural properties, and that any interpretation of a theory of the world must respect these objectively existing classes if it is to make the theory true. There are countless many interpretations which satisfy a consistent theory T : T is 'true' with respect to each of these interpretations, but what realists want is truth under an *eligible* interpretation, not truth under some interpretation or other.

Now, I've been talking as if naturalness and eligibility are all-or-nothing affairs, but this is an oversimplification. Ultimately, Lewis's claim about eligibility of referents makes eligibility a matter of degree, and correct or 'intended' interpretations are defined as those which strike the best balance between eligibility and making the claims of the theory come out true. In this way, we may distinguish interpretations which satisfy the claims of a theory: those that maximize naturalness in their assignments of extensions to predicates are better than those which do not. All else being equal, the more natural the extensions of the predicates assigned by the interpretation, the better the interpretation.

So Lewis has managed to defuse Putnam's bomb quite effectively. If we accept the claim that eligibility of interpretations depends upon naturalness, then the number of eligible interpretations for the language of a theory is greatly reduced. It is again possible for an ideal theory to be false, and Putnam's claim that 'any interpretation which satisfies an ideal theory makes it true' is refuted.

¹⁰ One of Sider's interesting contributions to this discussion is to *not* assume that the first order vocabulary "takes care of itself." Instead, Sider argues that we need a preferred interpretation for the quantifiers as well.

Many semantic realists who accept natural kinds take Lewis to have successfully provided a constitutive account of determinate reference and a decisive refutation of Putnam's model-theoretic argument. The account has been taken as the foundation for several different descriptivist programs, such as analyses of causation, analyses of moral properties, functionalism, objective structure, and various other metaphysical projects.¹¹ Even realists who are less comfortable with kinds accept Lewis's argument as a refutation and grant that his account of determinate reference seems moderately plausible. But as we shall see, the story is not yet complete.

4. Permutability

Recall that part of the realist's task was to provide an account of moderately determinate reference that could underlie a theory of truth.¹² He has an account of how an ideal theory can still be false. But does this account guarantee us moderately determinate reference? Unfortunately, it does not. There are two ways in which indeterminacy can remain; the lesser of the two is one which the realist might reasonably accept, but the greater of the two is unpalatable to even the most entrenched.

Moderate indeterminacy, of the sort noticed by Quine, is a familiar bugbear of the descriptivist program. Since eligibility is a matter of degree, we may on occasion assign fairly poor referents if such an assignment allows us to maximize the eligibility of referents overall or the truth to falsity ratio of the theory's sentences. Lewis (1984, p. 229) notes that we might have to accept as eligible interpretations that assign rabbit stages, undetached rabbit parts, and a rabbit-fusion as referents. This is not too troubling. But the indeterminacy is not as innocuous as all that, for along with accepting that 'rabbit' might refer to rabbit-stages, we might also have to accept grue and bleen things as referents for, say, 'green' and 'blue'. This sort of moderate indeterminacy, while not as upsetting as radical indeterminacy, is a step beyond accepting rabbit stages as referents, and will make many realists uncomfortable.

¹¹ For example, Menzies (1999), Jackson (1998), Jackson and Pettit (1995a, 1995b), Jackson (1992), Braddon-Mitchell and Jackson (1996). Worrall's (1989) structural realism also seems to depend, implicitly, on the success of the sort of program that Lewis advocates. Further, any philosopher who assumes that we have (fairly) determinate reference and that some form of descriptivism plays a role in establishing reference is relying indirectly on this sort of semantic program, see for example, Sider (2012).

¹² Since my focus here is on descriptions rather than the causal theory of reference, from this point onwards (unless otherwise specified) I will use the terms 'realist', 'semantic realist' and 'descriptivist' interchangeably to refer to the descriptivist who thinks that our terms have (moderately) determinate reference.

For example, suppose we have an interpretation *I* for our best theory that assigns the predicate 'is green' to all grue things, 'is blue' to all bleen things, and does an adequate job of mapping of the rest of the individuals in the domain of the model to moderately eligible parts of the world while satisfying all or most of the sentences of our theory. Now suppose there is no better interpretation, i.e., no interpretation that assigns more eligible referents can satisfy any of the sentences of the theory. Then the realists would have to accept a somewhat unpleasant consequence: 'green' refers to the property grue and 'blue' refers to the property bleen. This might make realists uncomfortable. But this is, of course, not a refutation of the realist view, since if the realist can allow that the difference between referring to green things and referring to grue things is not too obnoxious, such mild cases of indeterminacy can be consistent with a moderate realist lifestyle.¹³ Such a consequence, the realist might argue, is better than the alternative of radically indeterminate reference.

It may seem as though the issue has been settled: we may have a moderate amount of indeterminacy but nothing too outrageous, so the sensible realist may go on as before, undeterred in his philosophical pursuits. But even after he accepts moderate indeterminacy, and accepts naturalness as a condition for eligibility of interpretations, the realist is still in trouble. Without an additional constraint upon interpretations, he is still faced with the prospect of having to accept a large amount of indeterminacy, an indeterminacy so radical that he may not be able to distinguish between referring to rabbits and referring to cherries, or between referring to green things and referring to spherical things. Even though by assuming that natural kinds exist he can refute Putnam's claim that nearly any empirically ideal theory is true of the world, this assumption is not sufficient to refute the thesis that reference is radically, not just moderately, indeterminate. This spells trouble.

The worry surfaces when we realize that although linking naturalness to acceptability of interpretations is enough to avoid Putnam's claim that nearly any interpretation of an ideal theory will make it true, it does not eliminate the possibility that more than one radically different interpretation of the language of an ideal theory could receive the title 'most eligible'. If we had two or more interpretations that maximized naturalness while making the claims of the theory come out true (as our grue and green cases did above), yet the way in which the interpretations assigned referents were very different, we could be faced with indeterminacy that is far more radical than that of rabbit versus rabbit stages.

¹³ Lewis was prepared to accept this much indeterminacy, just as he is prepared, if need be, to accept the sort of indeterminacy we have with respect to rabbits versus rabbit stages, etc. It is not a conclusion he welcomed. I myself find it objectionable.

Recall that Lewis circumvented Putnam by claiming that the world has a structure. But if there exist two or more very different interpretations for a theory that preserve the structure we have attributed to the world, then we are still in trouble. How could such a thing be the case? It could be the case if, for some interpretation that maximizes both naturalness and satisfaction of the claims of the theory, there exists a permutation of the natural properties of the world into itself which also maximizes both desiderata. In other words, there could be an alternative, equally good, mapping of terms in the language of the theory to natural properties. This would give us more than one (potentially) radically different interpretation for our best theory.

One familiar way that radical indeterminacy could occur arises with the possibility of isomorphic models for the same (syntactic) theory.¹⁴ We have isomorphic models for a theory when there is a way of pairing the objects of the domain of one model with the objects of the domain of the second model in such a way that relations which hold between objects in the first model correspond to relations that hold between objects in the second model. If it is possible to state a rule by which the elements of the system $\Sigma 1$ are paired in a mutually unique manner with the elements of the system $\Sigma 2$, so that elements in $\Sigma 1$ between which R (or R' , ...) holds correspond to elements in $\Sigma 2$ between which the relation Q (or Q' , respectively) holds, so that R is correlated with Q , and R' is correlated with Q' , etc., even though the R s and Q s may mean entirely different things, then the two domains are isomorphic (i.e., they share the same structure). A simple example is the progression of the odd natural numbers, which is isomorphic to the progression of the even natural numbers when we set up a correspondence whereby 1 corresponds to 2, 3 to 4, etc.

Now, we might have an interpretation I^* that assigned the predicate 'is a G ' to green things, and 'is an R ' to rabbits. In addition, we might have an interpretation J^* that assigns 'is a G ' to cherries, and 'is an R ' to meteors. (Of course, many referents of other predicates in the two interpretations would also differ. But assume that we are able to assign the rest of the referents in some satisfactory way.) Both interpretations map predicates to natural kinds. Since both interpretations assign eligible referents to their predicates, if both turn out to give us the same balance between eligible kinds and satisfying the claims of the theory, then both are equally eligible interpretations. If they are the best available interpretations, and the theory is ideal in other respects, we should, if we follow realist doctrine, accept both of them. However, this means that we must also accept a significant amount of indeterminacy with respect to the referents of the predicates 'is a G ' and 'is an R ', indeterminacy far more radical than most realists would feel comfortable with.¹⁵ And note that the interpretations of predicates for

¹⁴ Of course, the unsurprising fact that there exist identity transformations is not relevant here.

¹⁵ Michael Smith (1994, pp. 48-54) has a good example of local permutability, based on the (supposed) symmetry of the color wheel, that shows how permutations can confound uniqueness claims.

relations such as “counterfactually depends”, “causes” and “is earlier (later) than” are also up for grabs.¹⁶

When evaluating interpretations for our best theory of the world, we could end up with two interpretations such that the objects referred to by predicates in both interpretations were natural kinds, and the relations which obtained between the properties in one model were as natural as those which obtained in the second model. This means that we would end up with two ‘best’ or maximally natural, interpretations.¹⁷ If these models were isomorphs of one another, so that one interpretation mapped into the other, then they would be elementarily equivalent models, i.e., the same sentences would be made true by both interpretations. But note that isomorphism isn’t a requirement, merely a sufficient condition: even if we can’t construct isomorphs that maximize naturalness, we might still be able to permute the natural properties in the domain of a model for the theory so as to get ‘approximate’ isomorphisms: alternative interpretations that do an approximately equal job of maximizing naturalness and satisfying the claims of the theory.¹⁸ If we assume that the same natural properties in the world are subsets of each of the domains (or subsets of sets of pairs, triples, etc. of the domains) of competing models for the theory, then the models are automorphs, not just isomorphs, of one another. It is possible that an interpretation for our best theory of the world that maximized naturalness might be one that would admit of such permutations, automorphic, isomorphic or otherwise.¹⁹

For a physicalist or materialist, the most elite properties and relations are most likely things like counterfactual dependence, or having mass, charge, and quark flavor. Less elite properties such as ‘is a cherry’ could be supposed to be connected to these fundamental ones by long chains of definability.²⁰ The problem with interpretations would arise most vividly at the level of the assignation of fundamental properties and relations. Thus quark flavor, quark color and protons might be assigned to the predicates *A*, *B* and *C* in one interpretation that maximizes naturalness, and mass, charge and electrons to *A*, *B* and *C* in another, equally eligible interpretation. Or, since

¹⁶ And if the interpretation for the fundamental quantifier is up for grabs, then things get even crazier. This is relevant to the arguments in Sider (2012).

¹⁷ If the structure of the world exhibits enough symmetry, we will certainly have alternative competing interpretations.

¹⁸ Since naturalness and satisfaction of claims of the theory are both matters of degree, and the realist holds that these two desiderata may need to be balanced against one another.

¹⁹ Moreover, I’ve been assuming so far that the language of first order logic has a fixed interpretation. Putnam seems to assume this, although I see no reason for it. But if we do not assume, as Sider (2012) does not (and as I do not), that the language of first order logic has a fixed interpretation any more than the rest of our language has a fixed interpretation, then the possibility of competing radical interpretations is significantly increased.

²⁰ Lewis (1984), p. 228.

maximizing naturalness and the satisfaction of claims of the theory involves balancing matters of degree, we might have the interpretation that assigns quark flavor, quark color and protons to the predicates *A*, *B* and *C* (respectively) competing with an equally eligible interpretation that assigns quark flavor and some protons to the predicate *A*, quark color and some protons to the predicate *B*, and the rest of the protons to the predicate *C*. Such variance at the fundamental level would most likely give us extreme variance at the macro-level of cherries and rabbits, as in the example given earlier, given that the fundamental properties are used to define these less elite properties.

John Winnie (1967, pp. 226-7) gives an example that shows how the existence of two (or more) competing interpretations is not implausible. Imagine a theory with one interpretation, *K*, that assigns what we now call neutrons to the predicate *N* and electrons to the predicate *E*. Then imagine another interpretation, *K**, that assigns all neutrons but one to the predicate *N* and all the electrons plus one neutron to the predicate *E*. We may then assign other referents so as to satisfy the claims of the theory. We are left with a measure of indeterminacy at the micro-level and perhaps much more radical indeterminacy at the macro-level, depending upon the chains of definability in which the neutron is involved.²¹ Of course, we can come up with many more such models in which only a small switch is made. Now, in this simple case, we know that *K* is a better interpretation than *K**, but that is because we are assuming they do equally well with respect to satisfaction. But since naturalness and satisfaction are matters of degree, the existence of competing interpretations like *K* and *K**, but where *K** does a bit better with satisfaction while *K* does a bit better with respect to eligibility may not be too far-fetched. For a (more far-fetched) case where only perfectly natural kinds are referred to, imagine an interpretation *I*** that assigned the property of having positive charge to the predicate *POS* and the property of having negative charge to the predicate *NEG*. Then imagine an alternative interpretation *J*** that assigned the property of having positive charge to the predicate *NEG* and the property of having negative charge to the predicate *POS*. If the world exhibited enough global symmetry, then both these interpretations could do an equal job of maximizing naturalness and satisfaction of claims—so by the realist's own lights, the predicates *POS* and *NEG* would refer indeterminately.

If our world is one in which the natural kinds actually are permutable in any of the ways I have specified, that is, if the world exhibits the right sorts of deep-level symmetries, then the existence of competing interpretations is physically possible. If the world admits more than one interpretation as eligible, then *contra* the realist program, we would still have radical

²¹ Winnie's proof is relevant in the event that so-called observational predicates do not have a fixed interpretation in sentences that contain both observational and theoretical predicates, as is the case here, with discussions about what Lewis (1984) calls 'global descriptivism'. His argument does not refute Lewis (1970), since Lewis takes as a premise that the interpretations of observational (or 'old') predicates remain fixed in 'mixed' sentences containing the old predicates and new (theoretical) predicates.

indeterminacy of reference, where the indeterminacy involves reference to objects or kinds that are very dissimilar to one another.

5. Realist Responses

There might be a way out for the semantic realist who hopes for a constitutive account of determinate reference: deny that our actual world is one in which the natural kinds are permutable. The realist could agree that an alternative interpretation that permutes the kinds that the predicates of our best theory are to apply to while preserving empirical virtues is possible, but not in our world. This amounts to a claim about the sort of structure that our world has.

Such a claim makes sense, the semantic realist might argue, given what we know about the world. Science tells us about the world, and its main business is to try to discover natural kinds and talk about them in its theories. The referents of so-called theoretical and observational terms are the natural kinds of the world. Science tells us what the natural kinds are, such as the natural kind of being an electron and the natural kind of being a quark, and these, at least *prima facie*, don't seem to allow us to construct competing 'most natural' interpretations. However, the realist might argue, surely part of the language of our theory of fundamental physics will be indexical, and our use of 'here' and 'now', combined with assertions about our use of predicates about change, location and the direction of time²², will eliminate alternative interpretations. Although competing interpretations are logically possible, we don't have to worry about them because it doesn't seem like the kinds we have *are* permutable in a way that would maximize naturalness.²³ The realist could then try to use this argument to refute the examples in the previous section which suggested that the existence of competing interpretations was not implausible.

Now, this argument deserves close examination. The first objection to make to it is to point out that the indexical properties and fundamental relations described by our ideal theory are as open to interpretation as any of the other terms—we're talking *global* descriptivism here—and so the possibility of permutation still exists.²⁴ Recall that all we are working with is a set of existentially generalized sentences. Hereness and nowness, along with "depends upon",

²² Assuming, of course, that the realist accepts the thesis that temporal or other relevant fundamental relations are asymmetric.

²³ Lewis (1997) takes this line in response to worries about uniqueness with respect to the referents of color terms.

²⁴ And does the realist really want to assume that, e.g., the property of hereness is a natural property?

“causes” and “is earlier than” are just more natural kinds that enter into the equation when we balance naturalness and satisfaction of claims.

But perhaps a very strong version of scientific realism could imply that the kinds are not permutable. The realist could argue that science tells us the natural kinds, and that it does not seem that an interpretation that respected these kinds would be permutable. Thus, the semantic realist assumes that the kinds science tells us about are really all and only the kinds that there are. Since (via science) we know what all and only the kinds of the world are, claims the semantic realist, we can judge that they are not permutable and therefore that reference is (moderately) determinate.

This seems, *prima facie*, to be a plausible line of defense for the realist. After all, most scientific realists think that we have fairly strong reasons to think that our theoretical desiderata lead us to the truth. In other words, the semantic realist seems to be assuming something fairly reasonable for a scientific realist to assume—that our scientific theories are (approximately) true claims about the world (including the world’s unobservable realms.) The realist who accepts a version of scientific realism that tells her what the kinds are and implies that they are not permutable would seem to have a reasonable response to permutation worries.²⁵

If we grant the semantic realist the assumption that we know, via science, that the kinds of our world are not permutable and that this ensures (fairly) determinate reference, three conclusions follow immediately.²⁶ First, the semantic realist must assume that science will not discover any new natural kinds, kinds which would allow for permutations involving the predicates of current theory. Second, in order for one to be a semantic realist, one must first embrace strong scientific realist views such as the view that our current theory is reasonably close to what our ideal theory will be. These two results show us that a rather optimistic realist view of science is called for, one which places great faith in today’s science (and rejects the idea that the world has deep or fundamental unbreakable symmetries).²⁷

If the realist is right, we must now base semantic realism upon scientific realism, reversing the more usual view that realism about scientific theories is bolstered by semantic realism. But there is more: the realist must now claim that determinate reference is contingent (in a surprising new way). Determinate reference, according to the realist view, is dependent upon the nature of the natural kinds that make up a world. If a world has the wrong sort of kinds, then

²⁵ I’m assuming here that most semantic realists would accept abductive inferences as reasonable and as fairly likely to be (objectively) true.

²⁶ Among other things, this implies the view that science tells us that the world is asymmetric (overall).

²⁷ For a discussion of how a local symmetry of color properties can be broken by taking a more global perspective, see Lewis (1997).

determinate reference might not be possible. We've granted to the semantic realist who holds a strong version of scientific realism that our world has the right sort of kinds so as to avoid permutability, but surely there are many other worlds where the kinds *are* permutable. So even if we grant the realist the claim that science tells us what the kinds are and that these kinds are not permutable, the realist conception of reference makes reference contingent—not only in the trivial sense that it is a contingent matter how languages and theories are constructed—but contingent upon the world having the right sort of structure. If we had lived in a world with different fundamental structure, we might not be able to refer determinately.

Now, here's the rub: the realist can't argue that he knows that his strong (or optimistic?) version of scientific realism is true simply because current theory seems to do a fairly good job of meeting our empirical and theoretical desiderata, since he has already argued against Putnam that such a theory can be false.²⁸ So he can't *argue* that scientific theory must be true. Instead, within the context of his rejoinder to Putnam, the realist must assume, based on his *faith* that scientific theory is true, that the kinds of our world are not permutable.

But although the strong scientific realist can make this assumption about the kinds, the antirealist about science will not be so sanguine. She will reject the assumption as unjustified, and argue that therefore no satisfactory account of determinate reference is available. Skepticism about the existence of natural kinds is familiar; we may add to this skepticism about contingent properties of kinds.

More importantly, semantic realists who prefer moderate scientific realism will be caught in a bind, for *they* cannot use their version of scientific realism to claim that the kinds are not permutable.²⁹ Unless she is exceedingly optimistic about the status of current scientific theory, the moderate scientific realist who thinks descriptivism can give us moderately determinate reference will also have to make an additional assumption (without scientific justification) about the contingent structure of the physical world.

²⁸ We can see how this debate drops into a debate in the philosophy of science: the realist maintains that science is about discovering the real, i.e., natural, properties of the world, whereas the antirealist argues that the properties that science discovers and defines are based on pragmatic (human) desiderata. See van Fraassen (1989, 1995, 1997) and Elgin (1995).

²⁹ By 'moderate scientific realist' I mean a scientific realist who thinks that we have good reason to believe that our theories of fundamental physics are approximately true, but doesn't want to rule out the possibility that there might be significant revisions made to these theories, or a realist who thinks that part but not all of fundamental physics should be held to be approximately true, or a realist who doesn't want to rule out the possibility that we might discover new natural kinds that would allow for permutations involving the predicates of current fundamental physical theory.

We come then to a sort of skepticism. In order to provide an adequate account of how reference is determinate, the realist must make the assumption that the world has structure. But in order to secure the sufficiency of this account, the realist who accepts only moderate scientific realism must make an additional substantial ontological commitment: she must assume that the structure of the world is such that it does not permit the possibility of multiple, radically different, ‘most natural’ interpretations for our best theory. Realists already tend to take the thesis that kinds exist as a primitive assumption about the world; now, unless they are exceedingly optimistic about current science, they must assume, without independent evidence, that the structure of the world is non-permutable.

Providing ways to motivate acceptance of such an assumption will not be easy. The semantic realist who is not comfortable with assuming strong metaphysical theses to ensure determinate reference may well find the burden of accepting yet another controversial metaphysical claim too much to bear. Perhaps she thought, when she accepted the semantic realist view, that accepting natural kinds was sufficient to refute Putnam’s skeptical arguments. Perhaps she thought, when she accepted the semantic realist view, that even though accepting the idea that there exist natural kinds was somewhat repugnant, that such an assumption—just barely—outweighed the consequences. After all, accepting the existence of natural kinds, even a much weaker version than Lewis’s, helps to solve other philosophical problems. Moreover, there is a clear reason why the assumption that natural kinds exist is acceptable to realists: it conforms to our intuitions about truth and reference that interpretations for theories whose terms refer and whose theorems are true do so in virtue of the fact that they respect the actual, real differences and samenesses between objects in the world. Using the assumption that natural kinds exist to refute Putnam reflects traditional realist values; as Lewis puts it, ‘the realism that recognizes a nontrivial enterprise of discovering the truth about the world needs the traditional realism that recognizes objective sameness and difference, joints in the world, discriminatory classifications not of our own making.’ (Lewis 1984, p. 228)

But while there is support for the idea that the world has some structure, there is no such support for the stronger idea of nonpermutability—the realist does not have a story about the intuitions captured by denying deep structural symmetry. Yet, she must attribute non-permutability to the structure of the world. This is a problem for the realist. The attribution seems *ad hoc*, and the fact that it must be accepted on faith sounds more like a layman’s argument for the existence of God than an argument for a respectable naturalist philosophy of reference and truth.³⁰ The moderate scientific realist may well feel that the need to accept the

³⁰ There may be another option (for which I am indebted to David Lewis): reject the assumption that the kinds are permutable and instead accept the thesis that a term has determinate reference if it refers to the same entity under all eligible interpretations. However, this means accepting the possibility that truth

additional assumptions about structure, coupled with the counterintuitive result that determinate reference is contingent upon the properties of the kinds, tips the balance towards a view of semantics that construes reference differently from its construal under the descriptivist view.

6. Conclusion

At the start of this paper, I argued that a descriptivist theory of reference seemed to do the best job of capturing our views about how reference for theoretical and other terms in our language is established. A pure causal theory of reference might be suitable as an account of how proper names get their reference, but it isn't going to give us all the reference we want or need. For these reasons, a descriptivist theory seems to be the best theory of reference that we have (for many of the terms of our language). In the face of Putnam's model-theoretic argument, Lewis and other realists have marshaled elegant and interesting theses that embellish descriptivism so as to refute Putnam's claim that their views allow for radical indeterminacy of reference.

But the problem of permutability shows us that, according to the realist's constitutive account of reference, the possibility of determinate reference is contingent upon the properties of the kinds that make up our world. The semantic realist who is willing to accept strong scientific realism on faith must accept this conclusion; the semantic realist who is a moderate scientific realist must accept this conclusion *and* make the additional assumption that the kinds of the world are not permutable (in a way that jeopardizes determinate reference). I find this new contingency of determinate reference quite counterintuitive, even apart from the need to hold a strong version of scientific realism or to make assumptions about the properties of worldly structure in order to bolster the view. We knew, as per the realist view, that reference depends on the external world in the sense that if we were living on twin-earth, our tokens of 'water' should refer to XYZ instead of H₂O. But the permutation problem shows us something new: reference also depends on contingent properties of the structure of the world. In other words, using tokens of 'water' on our world in the way that we do is not enough to secure moderately determinate reference to H₂O; our world must also be one where the kinds are not permutable. Similarly on twin-earth: tokens of 'water' on twin-earth refer to XYZ, but might also refer to vastly different things if XYZ is permutable with the other kinds of twin-earth. According to the realist view of reference, if we are unlucky and the structure of the world admits the possibility

may be much gappier than is normally thought to be the case. Moreover, if we do not assume within the context of the model-theoretic argument that the language of first order logic has determinate reference (and I don't think we ought to assume this—see footnote nineteen), some cherished truths (such as logical or mathematical truths) may go by the wayside.

of permutability, then *nothing we can do or say* will secure us even moderately determinate reference. Such a consequence is grounds for rejecting the descriptivist view as an adequate characterization of the reference of the terms in our theories.

References

- David Braddon-Mitchell and Frank Jackson, 1996. *Philosophy of Mind and Cognition*, Blackwell.
- Michael Devitt, 1983. 'Realism and the Renegade Putnam: A Critical Study of *Meaning and the Moral Sciences*', *Nous*, 291-301.
- Catherine Elgin, 1995. 'Unnatural Science', *Journal of Philosophy*, 289-30.
- Frank Jackson 1992, 'Critical Notice of Susan Hurley's *Natural Reason*', *The Australasian Journal of Philosophy* 70, 475-87.
- Frank Jackson and Philip Pettit, 1995. 'Moral Functionalism and Moral Motivation', *The Philosophical Quarterly*, 45, 20-39.
- Frank Jackson and Philip Pettit, 1996. 'Moral Functionalism, Supervenience and Reductionism', *The Philosophical Quarterly* 46, 82-86.
- Frank Jackson, 1998. *From Metaphysics to Ethics*, Oxford: Oxford University Press.
- David Lewis, 1970. 'How to Define Theoretical Terms', reprinted in Lewis 1986, *Philosophical Papers*, vol. 1, Oxford Univ. Press, 78-95.
- David Lewis, 1983. 'New Work For a Theory of Universals', *The Australasian Journal of Philosophy* 61, 343-77.
- David Lewis, 1984. 'Putnam's Paradox', *The Australasian Journal of Philosophy* 62, 221-236.
- David Lewis, 1997. 'Naming the Colours', *The Australasian Journal of Philosophy* 75, 325-42.
- Peter Menzies, 1999. 'Intrinsic Versus Extrinsic Conceptions of Causation', in *Laws and Causation: Australasian Studies in the History and Philosophy of Science*, Howard Sankey (ed.), Kluwer, 313-29.
- Gary Merrill, 1980. 'The Model-Theoretic Argument Against Realism' *Philosophy of Science* 47, 69-81.
- Sider, Theodore, (2012). *Writing the Book of the World*. Oxford University Press.
- Hilary Putnam, 1977. 'Realism and Reason', *American Philosophical Association Proceedings and Addresses*, vol. 50, no. 6, 483-498.
- Hilary Putnam, 1980. 'Models and Reality', *Journal of Symbolic Logic* 45, 464-482.
- Peter Unger, 1983. 'The Causal Theory of Reference', *Philosophical Studies* 43, 1-45.
- Bas van Fraassen, 1989. *Laws and Symmetry*, Oxford Univ. Press.
- Bas van Fraassen, 1995. 'Elgin on Lewis's "Putnam's Paradox"', *Journal of Philosophy*, 85-93.
- Bas van Fraassen, 1997. 'Putnam's Paradox: Metaphysical Realism Revamped and Evaded', *Philosophical Perspectives*.
- John Winnie, 1967. 'The Implicit Definition of Theoretical Terms', *British Journal of Philosophy*.

John Worrall, 1989. 'Structural Realism: The Best of Both Worlds?', *Dialectica* **43**, 99-124.