

# Van Fraassen's Dissolution of Putnam's Model-Theoretic Argument\*

Mathias Frisch<sup>†‡</sup>

Department of Philosophy, Northwestern University

---

Bas van Fraassen has recently argued for a “dissolution” of Hilary Putnam’s well-known model-theoretic argument. In this paper I argue that, as it stands, van Fraassen’s reply to Putnam is unsuccessful. Nonetheless, it suggests the form a successful response might take.

---

In a series of recent papers on Hilary Putnam’s well-known model-theoretic argument, Bas van Fraassen has argued that it is a mistake to try to answer the challenge posed by Putnam’s argument directly (van Fraassen 1997a, 1997b, 1997c). It is a mistake, van Fraassen argues, since trying to answer directly Putnam’s argument offers an “entrance into the very kind of metaphysics Putnam intended to undercut” (1997a, 85). Instead, van Fraassen proposes what he calls “a dissolution” of the argument. Van Fraassen believes that if we take our pragmatic situation as speakers of a meaningful language seriously, Putnam’s alleged paradox does not arise. I will argue here that van Fraassen’s reply is not successful, even though it points in the direction of a satisfactory answer to Putnam’s challenge—the answer Putnam

\*Received October 1997; revised January 1998.

<sup>†</sup>Send requests for reprints to the author, Department of Philosophy, Northwestern University, 1818 Hinman Ave, Evanston, IL 60208–1315.

<sup>‡</sup>Most of the research for this paper was done during my stay as a visitor at the Centre for the Philosophy of the Natural and the Social Sciences at the London School of Economics during the 1997 Lent term. I wish to thank the members of the Centre for their hospitality and for many stimulating discussions. For detailed comments and criticisms on various earlier versions of this paper I want to thank Paul Horwich, Lisa Lloyd, Stathis Psillos, Gabriel Siegal, John Worrall, and two anonymous referees for *Philosophy of Science*.

Philosophy of Science, 66 (March 1999) pp. 158–164. 0031-8248/99/6601-0007\$2.00  
Copyright 1999 by the Philosophy of Science Association. All rights reserved.

himself appears to have endorsed when he originally presented the model-theoretic argument (see Putnam 1978, 1983).

In his discussion of Putnam's argument, van Fraassen follows David Lewis's construal of the argument (Lewis 1984). In Lewis's formulation the argument is directed against a certain view of reference which Lewis calls "global descriptivism" and according to which the only constraints on the referents of the non-logical vocabulary of a theory are given by the logical structure of the theory and our intention to come out right. Putnam can be understood as arguing that if the domain of objects, which he calls "THE WORLD," is large enough there will always be multiple mappings from the terms of a theory onto sub-sets of objects of THE WORLD such that the theory comes out true. Any such mapping will satisfy the constraints on reference allowed by global descriptivism. Thus, if global descriptivism is right, then reference is radically indeterminate. Lewis calls the conclusion of Putnam's argument a "paradox," presumably because Lewis takes global descriptivism to be an intuitively plausible theory of reference and yet this theory implies, as Putnam shows, that our words do not refer. Whether this is in fact a paradox or not, Lewis correctly understands Putnam's argument as a *reductio* of global descriptivism and he goes on to argue that this shows that there has to be an additional constraint on reference: Determinate reference is possible only if not all sets of objects are eligible as referents for the terms of our language but only certain elite classes or natural kinds. Thus, Lewis turns Putnam's argument into an argument for "anti-nominalism." And it is this argument that van Fraassen's dissolution of Putnam's challenge is meant to block.

What is it to dissolve Putnam's paradox? Van Fraassen does not explicitly say what he means by a "dissolution," but I take him to mean roughly the following. Putnam's challenge begins with the question: "How is it that terms of our language refer?" Putnam then argues that on a certain view of reference the fact that any of the terms of our language refer becomes unintelligible. A direct answer to Putnam would consist in identifying an additional constraint on reference not allowed by global descriptivism—such as a causal constraint or Lewis's elite classes—that could secure the reference of at least some of the terms of our language. The identification of such a constraint is meant to provide an answer to how it is that the terms of our language refer. By contrast, van Fraassen thinks that Putnam's question should not be answered, rather it should be rejected as in a sense illegitimate. While the question appears to be a "mandatory philosophical question" (1997b, 34), van Fraassen claims that it in fact raises a worry that cannot be raised coherently and, thus, need not be answered. To dissolve the paradox, then, is to show why Putnam's question cannot

sensibly arise. Van Fraassen's response to Putnam's argument is similar to certain responses to varieties of philosophical skepticism. The skeptic asks: What is our justification for a certain class of beliefs, say beliefs about the external world? Instead of offering a justification, one type of response to the skeptic proceeds by arguing that the skeptic's worry is unintelligible. For example, one might argue that the skeptic's question only makes sense if one assumes knowledge of the very things of which the skeptic says we can have no knowledge. Such a response would not provide an answer to the skeptic, but would, if successful, prevent the skeptic's worry from arising.

Van Fraassen's response begins by noting the anthropomorphic way in which Putnam presents his argument. Putnam says "pick a model *M*" and "map the individuals of *M* one-to-one into the pieces of THE WORLD" (Putnam 1978, 126). But, van Fraassen says, we cannot map objects into the world, because without being able to describe the world independently we cannot define any mapping function from objects of the model to objects of the world. Thus, he thinks Putnam is faced with the following dilemma:

- (A) If we cannot describe the elements of THE WORLD, neither can we describe/define/identify any function that assigns extensions to our predicates in THE WORLD;
- (B) If we can describe those elements then we can also distinguish between right and wrong assignments of extensions to our predicates in THE WORLD. (van Fraassen 1997b, 21)

How does this dilemma affect Putnam's argument? Does it dissolve the paradox? Unfortunately van Fraassen does not offer much help in answering these questions. After posing the dilemma, he simply says that this dissolves the paradox. Van Fraassen must think that Putnam's challenge can arise on neither of the two horns. Not on the first horn, because then we cannot specify the multiple mapping functions that are responsible for the failure of reference; and not on the second horn, because if we can describe the world independently, we can determine a unique intended interpretation for our terms so that the threatened indeterminacy cannot arise.

Now Putnam would surely agree to the second horn of the dilemma. If we could specify the domain of objects independently, in some interpreted meta-language, say, then we could single out one mapping from the terms of our theory onto objects in the domain as providing us with the intended interpretation of those terms, and Putnam's challenge could not arise. But, Putnam would argue that this meta-language would go the way of all theory: The referents of its terms

would have to be fixed in some way. Thus, appealing to a meta-language is of no help. But, what is more, we do not need a theory in a meta-language to distinguish between right and wrong assignments of extensions to our predicates. Putnam would say to van Fraassen that we can specify the elements of THE WORLD and we can do so by using the terms of the very theory at issue. We can specify those elements, because the terms of the theory refer. What Putnam is asking is: How is it that they refer?

Moreover, Putnam would not take issue with the first horn, either. Rather, he would reply that it is the first horn that gives rise to his challenge. It is precisely because, *according to the view of reference at issue*, we cannot identify any function that assigns extensions to our predicates that his challenge arises. Only because we cannot identify any function can we not choose among the many functions that exist and pick one as providing our terms with their intended interpretation. What van Fraassen's reply does show is that Putnam's formulation of the argument is misleading. Putnam's challenge does not arise because we can pick out too many different models of our theory; rather, it arises because there are too many models and, on the view of reference against which his argument is directed, none of them are in any way privileged.

Van Fraassen's dilemma offers a useful way of putting Putnam's argument: We can describe the elements of THE WORLD, because many of the terms of our best theories refer. But it follows from global descriptivism that we cannot describe those elements. Therefore, global descriptivism is false. But if the argument is put this way, one can also see that the dilemma does nothing to dissolve Putnam's question "in virtue of what do the terms of our language refer?"

Van Fraassen also offers another line of response, which does not rely on Putnam's anthropomorphic formulation of the argument. Van Fraassen argues that Putnam's challenge can only arise if we put ourselves into the position of a Quinean field-linguist with respect to our own language. Van Fraassen takes Putnam's general philosophical question "How is it that the terms of our language refer?" to reflect more specific worries, such as "Does 'cat' really refer to cats?" It is these specific worries which we have to address when "we get down to brass tacks" (1997b, 36) with the general philosophical question. But, van Fraassen argues, the latter worry makes no sense. As field-linguists investigating an alien language we can wonder whether we have interpreted that language correctly. We can, for example, ask whether "Katze" really refers to cats. Putnam's paradox arises, according to van Fraassen, because Putnam asks us to consider the equivalent question for our own language: Does "cat" refer to cats? But for our own

language that question cannot sensibly arise. Sentences like “‘cat’ refers to cat,” van Fraassen says, are pragmatic tautologies.<sup>1</sup> We might aEnglish speaking field-linguists not be able to determine the extension of “Katze.” But we are not, according to van Fraassen, in the position envisaged by Putnam where we cannot fix the extension of the term “cat.” We know that “cat” refers to cats and any mapping that does not assign all and only cats to “cat” has to be rejected by us. From this van Fraassen concludes that the question “how is reference fixed?” is nothing “but a pseudo problem” (1997b, 37).

Again, I do not see how considerations like these can block Putnam’s challenge. Putnam, like van Fraassen, starts from the assumption that our language is interpreted and that most of its terms refer. He then asks, given that many of the terms of our best theories refer, how is this possible? In virtue of what do our words refer? Nothing of what van Fraassen says shows how that question cannot arise for our own language. The field-linguist’s worry is not Putnam’s worry. The field-linguist might doubt whether “Katze” really refers to cats; Putnam knows that “cat” refers to cats and goes on to argue that a certain view of language cannot account for this fact. Putnam asks: “*How* is it that ‘cat’ refers?” The confused philosophical field-linguist who is the target of van Fraassen’s argument asks: “*What* is the referent of ‘cat’?” Showing that the latter question is incoherent does not impugn the intelligibility of the former.

Putnam would agree with van Fraassen that “‘cat’ refers to cats” is a pragmatic tautology. But what makes this sentence a pragmatic tautology is that “cat” is an interpreted word of our language. Contrast this sentence with the sentence “‘P’ refers to P” where P is an uninterpreted symbol. The latter sentence is either false or meaningless. The second occurrence of “P” is used to pick out the putative referent of “P.” But if “P” is uninterpreted it has no referent. So the question Putnam is asking may also be put in the following way: Given that the sentence “‘cat’ refers to cats” is a pragmatic tautology, what makes it the case that that is so? This is a question that can surely arise for our own language. Of course it is always open to van Fraassen simply to reject Putnam’s question, but that is not the same as a dissolution. If van Fraassen were to offer a dissolution of Putnam’s paradox he would have to show that Putnam’s problem is somehow ill-founded or cannot arise, but as far as I can tell, none of his considerations show that.

1. In what follows I am, like van Fraassen, ignoring the fact that we would not assent to all instances of a schema of the form “‘s’ refers to s”, where “s” is a meaningful word in our language. If reference implies existence, there are instances of the schema, such as “‘phlogiston’ refers to phlogiston”, that we would deny.

Even though van Fraassen's proposed dissolution of the paradox is unsuccessful, his reply suggests the direction a successful response to Putnam's challenge might take. On many views of reference many instances of the schema " 's' refers to s" constitute what van Fraassen calls "pragmatic tautologies." But on some views of reference, like causal theories or Lewis's amended descriptivism, noting that instances of the schema cannot sensibly be denied by us if the word in question is a meaningful word in our language with a non-empty extension still leaves open the question how the word refers. Van Fraassen's response to Putnam is unsuccessful if it is not supplemented with a view of reference that prevents this question from arising. Such a view is suggested by Putnam's own response to the argument. Putnam's own reply begins in a way that is strikingly similar to the one proposed by van Fraassen. Like van Fraassen after him, Putnam points out that we are never in the position of having to interpret our own language. Like van Fraassen, Putnam would say that the sentence " 'cat' refers to cats" cannot sensibly be denied. But while van Fraassen's reply ends here, Putnam goes on to give a reason for why this prevents his paradox from arising. Putnam advocates what Paul Horwich calls "a deflationary view of reference." According to Putnam, " 'cat' refers to cats" is true, because it follows from the definition of the word "refers" (see Putnam 1978, 136). That is, according to Putnam, "refers" is defined implicitly through instances of the schema " 's' refers to s," where s is any meaningful word in our language. But now there is no further question to be asked in virtue of what "s" refers to s. It simply follows from the meaning of the word "refers" that instances of the schema are true.

Taking this line requires that we can give an account of what it is to understand the meaning of words that does not appeal to the notion of reference. Putnam believes that such an account can be given. He maintains that our use of our words determines their meaning and once a language "has a full program of use" (1983, 443) there can be no further question of its interpretation: "Either the use *already* fixes the 'interpretation' or *nothing* can" (1983, 443; italics in original). But then what Putnam offers in reply to his own puzzle is more than a dissolution. He offers an answer to the puzzle: The assumption of global descriptivism that only our intention to come out right constrains the interpretation of our words needs to be rejected, not because causal conditions or natural kinds supply additional constraints on reference, but because our use of our language determines what we mean by our words.

## REFERENCES

- Lewis, David (1984), "Putnam's Paradox", *Australasian Journal of Philosophy* 62: 221–236.  
 Putnam, Hilary (1978), *Meaning and the Moral Sciences*. London: Routledge.

- . (1983), “Models and Reality”, in Paul Benacerraf and Hilary Putnam (eds.), *Philosophy of Mathematics*. Cambridge: Cambridge University Press, 421–444.
- van Fraassen, Bas (1997a), “Elgin on Lewis’s Putnam’s Paradox”, *Journal of Philosophy* 94: 85–93.
- . (1997b), “Putnam’s Paradox: Metaphysical Realism Revamped and Evaded”, in James E. Tomberlin (ed.), *Mind, Causation, and World*. Malden, MA: Blackwell, 17–42.
- . (1997c), “Structure and Perspective: Philosophical Perplexity and Paradox”, in Maria L. Dalla Chiara (ed.), *Logic and Scientific Methods*. Dordrecht: Kluwer, 511–530.