

## CHAPTER 9

# ONE CANNOT BE JUST A LITTLE BIT REALIST: PUTNAM AND VAN FRAASSEN\*

Stathis Psillos

(T)he world is not a product. It's just the world.

Hilary Putnam, 1991

### 1. Introduction

Hilary Putnam and Bas C. van Fraassen have been two pivotal figures in the scientific realism debate in the second half of the twentieth century. Their initial perspectives were antithetical – defining an archetypical scientific realist position (Putnam) and a major empiricism-inspired alternative to scientific realism (van Fraassen). But as the years (and the philosophical debates) went on, there have been important lines of convergence in the stances of these two thinkers, mostly motivated by an increasing flirtation with pragmatism and by a growing disdain towards metaphysics.

Putnam's views went through two major turns, in a philosophical journey he aptly described as taking him 'from realism back to realism' (1994, 494). Being an arch-scientific realist in the 1960s and the early 1970s, he moved to a trenchant critique of metaphysical realism and the adoption of a verificationist-'internalist' approach (what he called pragmatic or internal realism), which he upheld roughly until the end of the twentieth century. Then he adopted a direct realist outlook, what

---

\* A paraphrase of Christopher Hitchens's 'One cannot be just a little bit heretical', from *Hitch-22: A Memoir* (419).



he called 'common sense' or 'natural realism', which was based on the denial of at least some of the tenets of his internalist period (e.g., he abandoned a verificationist conception of truth), while at the same time he tried to avoid 'the phantasies of metaphysical realism'.

It is impossible (or almost so) to cover all aspects of Putnam's realist endeavours. I will therefore focus on his changing views about scientific realism.

Van Fraassen occupied a space in the scientific realism debate that was left vacant by Putnam's critique of fictionalism and verificationism – he favoured an agnostic stance towards the ontological commitments of literally understood scientific theories. His positive alternative, to realism, constructive empiricism (CE), was meant to be a position suitable for post-positivist empiricists – that is, philosophers who a) take for granted the empiricist dictum that all (substantive) knowledge stems from experience; b) take science seriously (but not uncritically) as the paradigm of rational inquiry; and c) take to heart all criticism of the positivist approach to science. In more recent work, CE has been placed within a broader framework, known as empiricist structuralism – motivated, at least partly, by Putnam's critique of metaphysical realism.

This chapter will discuss these two philosophers' engagement with scientific realism, hopefully in a way that highlights their overlapping trajectories.

## 2. From realism to *realism* and back again

### 2.1. The 'no miracles' argument

Putnam (1975, 73) is the author of the most famous argument for scientific realism; it has become known as the 'no miracles argument' (NMA).

The positive argument for realism is that it is the only philosophy that does not make the success of science a miracle. That terms in mature scientific theories typically refer (this formulation is due to Richard Boyd), that the theories accepted in a mature science are typically approximately true, that the same terms can refer to the same thing when it occurs in different theories – these statements are viewed not as necessary truths but as part of the only scientific explanation of the success of science, and hence as part of any adequate description of science and its relations to its objects.



I will not go into the heated discussion about this argument here (see Psillos 1999, ch. 4). Instead, I will make some observations about its role in Putnam's philosophy of science.

### 2.1.1. *Against positivist empiricism*

Putnam's NMA is a *positive* argument, which is meant to supplement a *negative* argument – namely, an argument against reductive empiricist or operationalist approaches to scientific theories. Such approaches had once been popular among empiricists, but it was widely accepted in the late 1930s that theories have excess content over whatever can be fully captured in a strict observational language. Yet it was not really until the early 1950s that it became apparent that the 'excess content' that theoretical terms and predicates have is their factual reference: they designate theoretical/unobservable entities. In his writings in the 1960s, Putnam aimed to motivate and defend this view by arguing systematically against verificationist, reductivist and instrumentalist approaches to scientific theories.

Three of his arguments stick out. The first (1962) relates to his attack on the supposed sharp distinction between observational and theoretical terms. The second (1965) relates to what came to be known as Craig's Theorem: for any scientific theory  $T$ ,  $T$  is replaceable by another (axiomatizable) theory  $\text{Craig}(T)$ , which consists of all and only the theorems of  $T$  which are formulated in terms of the observational vocabulary  $VO$ . The new theory  $\text{Craig}(T)$ , which replaces the original theory  $T$ , is 'functionally equivalent' to  $T$ , in that all observational consequences of  $T$  also follow from  $\text{Craig}(T)$ .

Putnam mounted a formidable attack on the philosophical significance of Craig's theorem, arguing a) that theoretical terms are meaningful, taking their meaning from the theories in which they feature, and b) that scientists aim to find out about the *unobservable* world and that theoretical terms provide them with the necessary linguistic tools for talking about things they want to talk about.

Putnam's third argument (1963) relates to the role theories play in the confirmation of observational statements. The idea here is that theories are often necessary for the establishment of inductive connections between seemingly unrelated observational statements.

Given this battery of arguments, the negative argument for scientific realism – namely, that its then extant rivals fail patently to account for the role, scope and aim of scientific theories – was hard to resist.

### 2.1.2. *Realism and the success of science*

Note that Putnam's argument for realism refers to Richard Boyd. In his widely circulated and discussed, but still unpublished, manuscript 'Realism and Scientific Epistemology', Boyd tied the defence of scientific realism with the best explanation of the fact that scientific methodology has succeeded in producing predictively reliable theories.

Boyd viewed scientific realism as a historical thesis about the 'operation of scientific methodology and the relation between scientific theories and the world' (1971, 12). As such, realism is not a thesis only about current science; it is also a thesis about the historical record of science. It claims that there has been convergence to a truer image of the world, even though past theories have been known to have been mistaken in some respects. This historical dimension is necessary if the truth (or partial truth, or significant truth) of scientific theories is to be admitted as the best explanation of the predictive reliability of methodology. For unless continuity-in-theory-change and convergence are established, past failures of scientific theories will act as defeaters of the view that current science is on the right track. If, however, realism aims to explain a historical truth – namely, that scientific theories have been remarkably successful in the prediction and control of natural phenomena – the defence of scientific realism can only be a posteriori and broadly empirical. This kind of defence of realism was very congenial to Putnam's overall approach in the 1960s, which rejected the claim that there are absolutely a priori truths.

## 2.2. What is scientific realism?

In light of all this, it is no accident that Putnam takes scientific realism to incorporate three theses:

- a. Theoretical terms refer to unobservable entities (REFERENCE);
- b. Theories are (approximately) true (TRUTH); and
- c. There is referential continuity in theory change (CONTINUITY).

### 2.2.1. *Literal reading of theories*

(REFERENCE) implies a certain non-verificationist reading of scientific theories – what came to be known as a 'literal or face-value understanding' of theories. Differently put, it implies a non-revisionist semantics for theories: if theories – taken at face value – talk about electrons and the

like, they should be taken to do *exactly* this: to refer to electrons and their ilk. But (REFERENCE) also implies a certain *metaphysical* image of the world: as being populated by unobservable entities. This might not be heavy-weight metaphysics, but, in the context in which it was put forward, it carried considerable weight. It made clear (as Feigl had already recognized) that unobservable entities are no less real than observable entities. It honoured the thought that theoretical entities have independent and irreducible existence. (REFERENCE) implies that the subject matter of science is the unobservable world – at least that it is no less the subject matter of science than are the observable entities.

### 2.2.2. *Truth as correspondence*

(TRUTH) takes realism beyond (REFERENCE) in asserting that t-entities (at least those referred to by t-terms featuring in true theories) are indeed real – they populate the world. But for both Boyd and Putnam, (TRUTH) implies a certain understanding of *truth* – namely, truth as correspondence. The chief motivation for such a conception of truth was explanationist. Putnam (and Boyd) insisted that truth (along with reference) plays a key explanatory role: it explains the success of action (more particularly, the success of scientific theories and methodology, in the case of science). This insistence is quite prominent in Putnam's writings until the middle 1970s and especially in his John Locke Lectures, delivered in Oxford in 1976 (cf. 1978).

When it comes to scientific theories, Putnam makes this point vividly by claiming that theories are maps of the world and, in particular, that it is best to view them as such 'if we are to explain how they help us to guide our conduct as they do' (1978, 100). To be successful maps, theories must correspond to some part of reality – pretty much like the successful map 'corresponds in an appropriate way to a particular part of the earth'. Making truth (and reference) explanatory concepts does not imply that they have to be explicitly mentioned in every single explanation of a successful action. Nor does it imply that they should feature prominently in an explanation of language understanding. I can understand how to turn the light on, by flipping the switch, without understanding (or even having the concept of) electricity. It does *not*, however, follow that electricity does not causally explain why the light comes on. Similarly for truth: language can be understood by mastering the use of words and expressions (what Putnam calls 'a use-theory of

understanding'), and yet a certain theory of truth can be adopted on the basis of offering an explanation of success.

(TRUTH) then has certain metaphysical implications; namely, that scientific theories are answerable to the world and are made true by the world. However, Putnam did not advance anything more than a certain set of general theses about what truth is *not*. One of them is that truth should not be equated with whatever logically follows from accepted scientific theories, even when these theories are empirically adequate and well-confirmed (cf. 1978, 34–5). The point here is not that theories are or tend to be false. Rather, it is that when truth is attributed to the theory, this is a substantive attribution which is meant to imply that the theory is *made* true by the world, which, in its turn, is taken to imply that it is logically possible that an accepted and well-confirmed theory might be false simply because the world might not conform to it. Let's call this view 'the Possibility of Divergence'. It is meant to capture a sense in which the world is independent of theories, beliefs, warrants, epistemic practices, etc.

### 2.2.3. Convergence

(CONTINUITY) takes scientific realism beyond (REFERENCE) and (TRUTH) by capturing the all-important notion of convergence in theory-change. Here again, Putnam states (CONTINUITY) in semantic terms: a t-term that features in different theories can nonetheless refer to the very same unobservable entity. This kind of thesis is necessary for convergence, since it secures that successor theories might well talk about the very same entities that their abandoned predecessors did, even though the now abandoned theories might have mischaracterized these entities. Putnam thought that the failure of (CONTINUITY) would lead to a disastrous 'meta-induction': 'just as no term used in the science of more than fifty (or whatever) years ago referred, so it will turn out that no term used now (except maybe observational terms, if there any such) refers' (1978, 25). Then, (REFERENCE) and (TRUTH) go by the board, too.

In a number of papers in the early 1970s, Putnam argued against reference-variance based on the so-called causal theory of reference. Its thrust is this: the reference of a t-term *t* is fixed by the existential introduction of a referent – an entity causally responsible for certain effects to which the term *t* refers. As can be easily seen, the causal theory disposes of semantic incommensurability and establishes (CONTINUITY). It

also makes available a way to compare theories and to claim that the successor theory is more truthlike than its predecessors. Besides, the causal theory tallies with Putnam's view that the defence of realism is, by and large, an empirical endeavour.

### 2.3. Against verificationism and fictionalism

Putnam's 'no miracles argument' for realism is the culmination of a complex network of arguments and views that aim to render scientific realism – viewed as endorsing a certain combination of positions – the best way to understand science and to explain its empirical successes. To further support this claim, Putnam (1971) pitted fictionalism against verificationism and argued that fictionalism fails to carve a space for genuine doubt over the reality of the entities that are deemed useful fictions.

Verificationism, Putnam (1971, 351ff) argued, superseded fictionalism as the dominant anti-realist position because it promised to close the gap between the claim (a) that *P* is false and the claim (b) that everything in experience is as if *P* were actually true – for example, there are no electrons and yet everything in experience is as if there were actually electrons. Fictionalism takes this combination of claims to be logically consistent precisely because it takes theories at face value. But verificationism closes the gap between (a) and (b) by taking a view of meaning that makes it the case that the meaning of *P* is exhausted (fully captured) by its empirical content (the difference it makes in experience); hence, two statements (or theories) that have exactly the same empirical content are semantically equivalent, no matter how different they appear to be in their theoretical content. Verificationism, then, defies a face-value reading of scientific theories. According to Putnam, verificationism superseded fictionalism in the minds of many philosophers because it was taken to offer an easy and straightforward way out of the *sceptical* challenge.

Putnam claimed that verificationism – and its concomitant anti-scepticism – was the wrong reason to reject fictionalism. He (1971, 352) called a verificationism-based rejection of scepticism 'the worst argument of all'. Verificationism blocks scepticism by denying the Possibility of Divergence. Both realism and fictionalism honour this possibility.

Why then should fictionalism be rejected? Because it does not make sense to have a *merely* fictionalist stance towards a theory that has

been accepted and employed in the explanation and prediction of observable phenomena. The fictionalist would typically read the theory literally, would treat the theoretical concepts as indispensable and would accept a theory 'for scientific purposes' but would refrain from commitment to the reality of the entities implied by the theory since she would take it that theory – though perhaps empirically adequate – is *false*. What possibly could show to a fictionalist that the theory is true? Putnam takes it that the fictionalist would demand a deductive proof of the theory and rightly objects that if this were the golden standard for acceptance as true, no non-trivial observational statements would be accepted as true either. The fictionalist would end up with scepticism. Putnam challenges the fictionalist to draw and motivate a robust distinction between rationally accepting a theory T (but treating its supposed entities as useful fictions) and rationally accepting that that theory is true. As Putnam (1971, 354) put it, if one rationally accepts a theory for scientific purposes, 'what further reasons could one want before one regarded it as rational to *believe* a theory?' His answer was that these reasons are good enough!

There have been two major reactions to this argument. The first is to accept Putnam's challenge and to try to defend fictionalism by showing that a certain theory T which assumes – if literally understood – the reality of certain entities can be replaced by another theory T' which does not imply commitment to the reality of the 'suspicious entities'. This is, in effect, the strategy followed by instrumentalists on the basis of Craig's theorem. Putnam's incisive critique of Craig-theorem-based instrumentalism in the 1960s blocked the revival of this position in the philosophy of science. But a position akin to this was revived in the philosophy of mathematics by Hartry Field (1980).

The second reaction to Putnam's argument is, not to adopt fictionalism, but to be agnostic. This is the position articulated by van Fraassen (1980). On this view, the collapse of verificationism does not make scientific realism the only rational option. We shall discuss this view below, in Section 3.

#### **2.4. Isn't 'the worst argument of all' not so bad, after all?**

Somewhat surprisingly, in the dying years of the 1970s, Putnam came to accept a third way to resist his own argument against fictionalism based on Michael Dummett's resuscitation of verificationism. Modern



verificationism, of the form Putnam came to endorse, takes it that truth is not recognition- or evidence-transcendent. Once this view is adopted, it transpires that the Possibility of Divergence noted above is blocked off: there is no logical or conceptual gap between a suitably justified assertion and truth. In his justly famous *Meaning and the Moral Sciences* (1978), Putnam declared *mea culpa*.

#### 2.4.1. *Blocking scepticism without verificationism*

When Putnam (1971) deplored the ‘worst argument’ for verificationism – namely, that it blocks off scepticism – he rightly felt the need to show how scepticism is blocked if verificationism is abandoned. His argument is captivatingly simple. The sceptical hypothesis (e.g., the brains-in-a-vat hypothesis) is yet another hypothesis alongside the realist one (e.g., that there is a world of ordinary material objects and human beings); hence, we should examine whether, and to what degree, the sceptical hypothesis is confirmed by the relevant evidence. Confirmation, however, requires a specification of the prior probabilities of the competing hypotheses. Hence, we should look at the prior probability of the sceptical hypothesis. Ranking alternative hypotheses according to their initial (or a priori) probabilities should be based on their respective plausibilities; that is, on judgements as to how plausible they are. But the sceptical hypothesis is far less plausible than the realist one; hence, it is much less confirmed than the realist hypothesis.

Putnam never said what exactly goes into the plausibility judgements. He took it to be enough to stress that

[t]o accept the plausibility ordering is neither to make a judgement of empirical fact nor to state a theorem of deductive logic; it is to take a methodological stand. (1971, 353)

What exactly is it to ‘take a stand’? Here again, Putnam does not say much. But from the quoted passage, it transpires that taking a stand amounts to making a certain commitment to view the world in a certain way, where this commitment is not idiosyncratic but, in a certain sense, constitutive of rationality. As Putnam (1971, 353) put it, it is the stand taken by all rational human beings – *vis-à-vis* scepticism, at least.

Why isn’t this a good enough answer to scepticism on behalf of a realist? Why, that is, couldn’t a realist (Putnam himself) leave open the



possibility of scepticism (honoured by the Possibility of Divergence) and, at the same time, argue (along the lines Putnam followed) that the sceptical hypothesis is far less credible than the realist one? Why, in yet other words, did Putnam feel the need to go for verificationism instead?

#### 2.4.2. *Blocking scepticism with verificationism*

Part of the answer, I think, is connected to Putnam's critique of metaphysical realism (MR). Putnam associated a number of doctrines with MR. MR is supposed to entertain all of the following:

The WORLD is supposed to be *independent* of any particular representation we have of it – indeed, it is held that we might be *unable* to represent the WORLD correctly (e.g., we might all be 'brains in a vat', the metaphysical realist tells us . . . (1978, 125)

Truth is supposed to be *radically non-epistemic* . . . (1978, 125)

The world consists of some fixed totality of mind-independent objects. There is exactly one true and complete description of 'the way the world is'. Truth involves some sort of correspondence relation between words or thought signs and external things and sets of things. (1981, 49)

(T)here is (. . .) a definite Totality of All Real Objects and a fact of the matter as to which properties of those objects are the intrinsic properties and which are, in some sense, perspectival. (1995, 303)

There isn't enough space here to discuss all these doctrines in any detail, nor to explore their connections. What should be noted is that even if it were rejected that there is such a fixed totality of objects and a fixed set to their intrinsic properties, it would still seem possible that we might be unable to represent the world and that the world might be independent of any particular representation we have of it.

This latter possibility – the Possibility of Divergence – captures more than a kernel of truth in Putnam's characterization of MR: a realist proper (call her metaphysical or not, it does not matter) should honour the Possibility of Divergence, at least in domains where it is extremely plausible to say that there is an external fact of the matter as to what is true or false; something outside our thoughts, language, and minds that is responsible for the correctness of what we come to believe and accept. Putnam's chief point against MR is meant to show that honouring this possibility is incoherent.



### 2.4.3. *The model-theoretic argument against metaphysical realism*

To show this, Putnam calls us to envisage an ideal theory *T* of the world; a theory that satisfies all operational and theoretical constraints; that possesses any property that we can imagine or please except objective truth – which is left open. He takes it that for a metaphysical realist *T* might still (in reality) be false. His argument then is a *reductio*: if we assume that ‘ideal *T*’ might still be false, we end up with absurdity. Let me sketch the argument. Suppose that *T* says that there are infinitely many things in the world. *T* is consistent (by hypothesis) and it has only infinite models. By the Löwenheim-Skolem theorem, *T* has a model of every infinite cardinality (greater than or equal to the cardinality of extra-logical symbols of the language of *T*). Now, pick a model *M* of *T*, having the same cardinality as the *WORLD*. Devise a one-to-one mapping *m* of the individuals of *M* onto the pieces of the *WORLD* and map the relations between the individuals of *M* directly into the *WORLD*. These mappings generate a satisfaction relation (in Tarski’s sense) – call it SAT\* – between (sets of) pieces of the *WORLD* and terms and predicates of the language of ‘ideal *T*’ such that *T* comes out true of the *WORLD*. (That is, the *WORLD* is isomorphic to the model *M* in which *T* is true.) The ideal theory has been shown to be true of the *WORLD*. Then how can we claim that the ideal theory might really be false? This, we are told, would be absurd.

Putnam’s challenge is that the very notion of a unique interpretation fixed by the world – implicit in a non-epistemic theory of truth – makes no good sense. Actually, Putnam anticipated an objection that a realist (himself a few years back!) would make: a causal theory of reference would show that, and *explain why*, a particular referential scheme for a language *L* – call it the intended interpretation – is picked out. Putnam’s retort was that a causal theory of reference would be of no help to the realist, since the model-theoretic argument can be extended to a word like ‘cause’: ‘cause’ can be reinterpreted no less than other words; in each model *M*, reference *M* will be defined in terms of cause *M*. Then, Putnam said,

unless the word ‘cause’ is already glued to one definite relation with metaphysical glue, this does not fix a determinate extension for ‘refers’ at all. (1980, 477)

#### 2.4.4. *Stopping the endless dialogue*

As Clark Glymour (1982, 177) has nicely put it, Putnam's argument seems like 'a kind of endless dialogue': whenever one says something about what singles out a referential scheme, Putnam says it is insufficient for the job, for what one says 'adds more theory' which may be reinterpreted in countless ways, and hence it is itself referentially indeterminate. Is the dialogue really endless, though? Note that if it is, Putnam's aim to show that MR is incoherent is far from proven. At most, we have a draw or a stand-off. But the dialogue is not really endless, for Putnam's 'more theory' move misinterprets what the realist claims. As David Lewis (1984) has observed, when a realist suggests an extra constraint – call it *C*-constraint – that fixes the intended referential scheme, what one suggests is that *in order for an interpretation to be intended, it must conform to C*. Then, the real issue is not whether the *theory of C* will come out true under unintended interpretations. Rather, it is what exactly *C* is and how it operates. In light of this, Lewis's suggestion is that the appeal to causal considerations in fixing the intended referential scheme is not just adding more theory but offering *constraints* to which an interpretation must conform in order to be intended.

Perhaps the best way to understand what these constraints might be has been suggested by G. H. Merrill (1980), who argued that it is questionable that realism conceives of the world merely as a *set of individuals*; that is, as a model-theoretic *universe of discourse*. The world, a realist would say, is a *structured entity*. Its individuals stand in specific relations to one another or to subsets of individuals. In particular, whereas Putnam's assumption is that the language precedes the world and 'structures' it, the realist position is that the world is *already structured, independently* of the language. It is then easily seen why the model-theoretic argument fails. For, an interpretation of the language – that is, a referential scheme – either matches the language to the existing structured world or it does not. If it does not, there is a clear-cut case in which even an ideal theory *might* be false. In particular, if the WORLD is a structured domain, in order for Putnam to have his model-theoretic argument, he would have to show that the mappings from a model *M* of *T* onto the WORLD are structure-preserving. Yet, it simply is *not* always possible to produce structure-preserving isomorphisms.

### 2.5. Scientific realism without the metaphysics?

This way of neutralizing Putnam's model-theoretic argument rests on an inflation of the metaphysics of realism. Where does all this leave *scientific* realism? In the midst of his conversion to internal or pragmatic realism, as Putnam tended to call his new verificationist position, he published a piece in which he did endorse scientific realism, suitably dissociated from both materialism *and* metaphysical realism (cf. 1982). What then is scientific realism? Prima facie, it still is what it was taken to be by the early Putnam: theoretical entities have irreducible existence; theoretical terms featuring in distinct theories can and do refer to the same entities; there is convergence in the scientific image of the world; and scientific statements can be (and are) *true*. But – there is always a 'but' – truth is now 'correct assertibility in the language we use' (1982, 197). So scientific realism is retained but dressed up in a verificationist garment.

Is this scientific realism enough? Note that even if truth is tied to justification, one should be careful how exactly this tie is effected. Putnam is indeed extremely careful. As he (1983, 162) put it, he looks for 'a realism which recognizes a difference between 'p' and 'I think that p', between being *right*, and *merely thinking one is right* without locating the objectivity in either transcendental correspondence or mere consensus'. Truth is not a property that can be lost – nor does it have a sell-by date. Hence, the verificationist notion that replaces (or captures) truth should be such that it retains this property of truth. The 'correctness' of an assertion is a property that can be lost, especially if it is judged by reference to current standards or consensus (which come and go). Because of this Putnam ties correctness to 'the verdict on which inquiry would ultimately settle' (1982, 200).

Putnam took it that truth should be constitutively linked with justification, for otherwise 'we cannot say what role it [truth] could play in our lives' (1982, 197). So, there is a set of epistemic constraints that a verificationist notion of truth should satisfy. As Nick Jardine (1986, 35) has aptly put it, the needed concept of truth should be neither too 'secular' nor too 'theological'. It should avoid an awkward dependence of truth on the vagaries of our evolving epistemic values, but it should link truth to *some* notion of ideal epistemic justification. But in its attempt to break away from 'secular' notions of truth and to make truth a standing and stable property, verificationism moves towards a 'theological'

notion: the justification procedures become so ideal that they lose any intended connection with humanly realizable conditions. In the end, it becomes either 'secular', resulting in an implausible relativism, or else 'theological' and hence not so radically different from a (realist) non-epistemic understanding of truth, according to which truth *outruns* the possibility of (even ideal-limit) justification.

Why should scientific realism incorporate the claim of mind independence (as elaborated by the Possibility of Divergence? Why, that is, couldn't someone who accepted the reality of unobservable entities without also rendering them mind-independent (in the above sense) be a scientific realist?

A moral that can be drawn from Putnam's early defence of scientific realism is that the success of science – success that realism is meant to explain – is hard won. It is neither trivial nor in any way guaranteed. The heated debate over the pessimistic induction (see Psillos 1999, ch. 5) has driven the point home that if there is continuity in theory change, this has been a considerable achievement, emerging from among a mixture of successes and failures of past scientific theories. A realist non-epistemic conception of truth and in particular the Possibility of Divergence do justice to this hard-won fact of empirical success and convergence. Given that there is no guarantee that science converges to the truth or that whatever scientists come to accept in the ideal limit of inquiry or under suitably ideal epistemic conditions will (have to) be true, the claim that science does get to the truth (based mostly on explanatory considerations of the sort we have already seen) is quite substantive and highly non-trivial. If, on the other hand, the Possibility of Divergence is denied, the explanation of the success of science becomes almost trivial: success is *guaranteed* by a suitably chosen epistemic notion of truth, since – ultimately – science will reach a point in which it will make no sense to worry whether there is a possible gap between the way the world is described by scientific theories and the way the world is.

It is wrong to pose to the realist the following dilemma: either the concept of truth should be such that cognitive success is guaranteed or else any cognitive success is a matter of pure luck. What Putnam has taught us, to be sure, is that the success of the realist project requires some epistemic luck: if the world were not mappable, science would not succeed in mapping it. But the realist has a story to tell us as to why and

how cognitive success, though fortunate and *not* a priori guaranteed, is not *merely* lucky or a matter of chance. The realist story (cf. Psillos 1999) will have to be phrased in terms of the reliability of scientific method and its defence. But there is good reason to think that this story is both sensible and credible.

### 3. Constructive empiricism

CE, as characterized by van Fraassen, is a mixture of two theses: an *axiological* and a *doxastic*.

- (i) Science aims at empirically adequate theories; and
- (ii) acceptance of scientific theories involves belief only in their empirical adequacy (though acceptance involves *more* than belief: namely, commitment to a theory).

As such, it is contrasted to an analogous doublet of realist theses:

- (i') the aim of science is true theories; and
- (ii') acceptance of theories implies belief in their truth.

Given the discussion we have had so far, this is a non-standard way to view scientific realism. But this is not accidental. Van Fraassen approached both scientific realism and constructive empiricism as ways to view science; that is, as ways to view a particular activity or game. This is most naturally understood in terms of its aim (what is the aim of science conceived of as an activity or game?) and of what it is to be counted as success in it (what is involved in meeting the aim of science?). CE, van Fraassen insists, is not an epistemology but a philosophy of science.

If science is seen as an activity, and if SR and CE are seen as rival accounts of this activity, what exactly is the issue between them? For realists, the key issue was the explanation of the (empirically certified and hard-won) success of science and the implication for the epistemology of science that this explanation should have. Not quite so for van Fraassen. Seen as rival accounts of an activity, SR and CE are compared vis-à-vis their ability to explain or accommodate the main actual

features of this activity (success being, if anything, just one of them). The explanandum is, we may say, the phenomenology of scientific practice, which, to be sure, should not include the intentions and doxastic attitudes of individual scientists but, instead, the salient features of the activity they are engaged in. The question then is, are there salient features of science (qua an activity or game) that force upon us philosophers of science SR? Or is CE a viable option, too?

For van Fraassen, CE offers an alternative-to-SR philosophical image of science: it views science as an activity or practice which is intelligible and successful, without also accepting that science aims at, and succeeds in, delivering truth. He suggests that it is precisely *this* image – CE – that modern empiricism should juxtapose to scientific realism. The crucial issue, of course, is the grounds on which CE is to be preferred to SR. As we shall see later, van Fraassen has shifted his position somewhat here. The tone of his *Scientific Image* was that CE is rationally superior to SR. Later on, and after a shift in his conception of rationality, it transpired that CE is a rationally permitted view of science.

CE, it should be stressed, is markedly different from old empiricist-instrumentalist positions. Unlike traditional instrumentalism, CE agrees with realism that theories and their concomitant theoretical commitments in science are ineliminable. Van Fraassen accepts that scientific theories should be taken at face value and be understood literally as purporting to describe the reality behind the phenomena:

[I]f a theory says that something exists, then a literal construal may elaborate on what this something is, but will not remove the implication of existence. (1980, 11)

So theoretical terms should not be understood as useful shorthands for complicated connections among observables. Rather, they should be taken to refer to unobservable entities in the world. Scientific theories are taken as constructions with truth values: true, when the world is the way the theory says it is, and false otherwise (cf. 1980, 10, 38). Unlike Putnam, van Fraassen downplayed the usefulness of the philosophy of language to the philosophy of science. He thought that nothing much could be gained by analysing the language of theories or by looking into the issue of the meaning of theoretical terms (cf. van Fraassen 1980, 56).

Part of the reason for this is that van Fraassen adopted (in fact, co-introduced) the so-called semantic view of theories. According to



van Fraassen, theories represent the world *correctly* by having it (the world) as one of their models. But CE does not require that the theories get the world right. Rather, it requires that theories be empirically adequate. This idea of empirical adequacy was meant to capture the old instrumentalist conception that theories should aim to save the phenomena. But whereas the traditional conception, bound as it was to the syntactic view of theories, took it that a theory is empirically adequate (i.e., it saves the phenomena) if and only if all of its observational consequences are true, van Fraassen cast this requirement in model-theoretic terms: for a theory to be empirically adequate it should be the case that the structure of the phenomena is embedded in one of the models of the theory (i.e., that the structure of the phenomena is isomorphic to an empirical substructure of a model of a theory). This way of casting the requirement of empirical adequacy frees it from the commitment to a distinction between observational and theoretical *vocabulary*. At the same time, it honours the instrumentalist (and, as noted above, fictionalist) claim that a theory may be empirically adequate and false: a theory may save all observable phenomena and yet fail to correctly describe their unobservable causes. However, if a theory is solely about observable entities, empirical adequacy and truth coincide.

In recent work, van Fraassen has defended CE as a species of structural empiricism (or empiricist structuralism, as he prefers to put it). This is a view about science and not a view about nature, as van Fraassen (2008, 239) is fond of saying. It incorporates the following two theses:

- I. Science represents the empirical phenomena as embeddable in certain *abstract structures* (theoretical models).
- II. Those abstract structures are describable only up to structural isomorphism.

Being empiricist, this position is focused on observable phenomena. These are taken to be the sole target of scientific representation. The means or the vehicles of representation are theoretical models – qua abstract mathematical structures – but precisely because mathematical structures can represent only up to isomorphism, the phenomena are described – through science – only up to isomorphism. So all we can know – through science anyway – is structure. This, to be sure, is

fully consistent with the thought that we know a lot about observable entities – which knowledge is not structural, since it is supposed to be theory-independent. We shall discuss later some of the problems that empiricist structuralism faces – problems which are intimately connected with those Putnam thought metaphysical realism faces.

### 3.1. Against metaphysics

Van Fraassen has taken it to be a merit of his own empiricism that it delivers us from metaphysics. Realism, van Fraassen says, buys into inflationary metaphysics. What is certainly right is that realist views aim to explain certain phenomena (broadly understood) by positing entities which are said to be causally-nomologically responsible for the explananda. Let's call this realist stance 'explanation by postulation'. To some empiricists (notably Duhem and van Fraassen), the critique of metaphysics is tied to the critique of explanation by postulation. This is supposed to be the pinnacle of inflationary metaphysics.

There are at least two distinct ways in which scientific realism can buy into metaphysics. The first is to adopt the view that the world has a deep and, by and large, unobservable structure – it is made up of entities and causal-nomological relations among them – which is constitutive of and causally responsible for the behaviour of the observable entities. Why, one may wonder, is this kind of explanation-by-postulation inflationary? In a sense, it obviously is: it proceeds by positing further entities that are meant to explain the life-world and its (typically non-strict) laws. But in another sense, it isn't. For if you think of it, it proceeds by positing micro-constituents of macro-objects, whose main difference from them is that they are, typically, unobservable. That a putative entity is unobservable is, if anything, a relational property of this entity and has to do with the presence of observers with certain sensory modalities (of the kind people have) and not others. No interesting metaphysical conclusions follow from this fact, nor any seriously controversial ontological inflation.

The other way in which scientific realism can go into metaphysics is to adopt a certain neo-Aristotelian conception of the deep structure of the world; in particular, one that posits 'regularity enforcers' (e.g., powers). Indeed, an increasing number of realists wed realism with neo-Aristotelianism (cf. Sankey 2008). But, I think (cf. Psillos 2011a), this isn't mandatory for realism.

Van Fraassen's critique of realism has been based, at least occasionally, on running together these two ways to link scientific realism to metaphysics. Consider the following arguments:

From the medieval debates, we recall the nominalist response that the basic regularities are merely brute regularities, and have no explanation. So here the antirealist must similarly say: that the observable phenomena exhibit these regularities, because of which they fit the theory, is merely a brute fact, and may or may not have an explanation in terms of unobservable facts 'behind the phenomena' – it really does not matter to the goodness of the theory, nor to our understanding of the world. (1980, 24)

The realist asks us to choose between different hypotheses that explain the regularities in certain ways; but his opponent always wishes to choose among hypotheses of the form 'theory  $T_i$  is empirically adequate'. So the realist will need his special extra premiss that every universal regularity in nature needs an explanation, before the rule will make realists of us all. And that is just the premiss that distinguishes the realist from his opponents. (1980, 21)

A number of points can be made against them. First, from the (suspicious anyway) claim that every regularity needs an explanation, it does not follow that it needs a non-regularity-based explanation. A realist can be happy with the thought that it is regularities all the way down; that is, that less fundamental (observable) regularities are explained by more fundamental (framed in terms of unobservables) regularities. Scientific realism does not have to explain the regularity there is in the world by positing regularity enforcers – that is, entities of distinct metaphysical kind that ground, hence explain, the regularities. Second, the claim that a theory is empirically adequate is already 'inflated' vis-à-vis the available data, which show at most that a theory is unrefuted. This could be taken as a brute fact. The very idea that this fact could be explained by the claim that the theory is empirically adequate shows that even the constructive empiricist does not stomach brute facts all too easily.

Concerning this last point, van Fraassen fully grants that going for empirical adequacy involves making a claim that goes well beyond the available data. He nonetheless takes it that

there is a difference: the assertion of empirical adequacy is a great deal weaker than the assertion of truth, and the restraint to acceptance delivers us from metaphysics. (1980, 69)



But this is an illusion. It does not deliver us from metaphysics. At the very best, it gets away with less metaphysics!

What then is van Fraassen's case *against* scientific realism? Ultimately, it is that the advocates of SR make a 'leap of faith', which is 'not *dictated* by reason and evidence'. Well, this is fine; especially if we read this 'dictation' as we should – namely, that scientific realism (or the truth of a theory) is *proved* by reason and/or evidence. No such proof is, or ever was, forthcoming. We have already seen Putnam claiming that some judgements are neither the product of empirical evidence nor of logic, but amount to 'taking a stand': to viewing the world in a certain way. Putnam, you may recall, took this stand-taking to be constitutive of rationality, since the latter isn't exhausted by the dipole: empirical evidence and logic.

In his writings from the late 1980s on, van Fraassen has developed a 'new conception of rationality' – foreshadowed in the passage above – in view of which, even by his own lights, empiricism and realism are compatible. According to it,

what it is rational to believe includes anything that one is not rationally compelled to disbelieve. And (. . .) the rational ways to change your opinion include any that remain within the bounds of rationality (1989, 172–3).

It follows that though it is rational to form beliefs that go beyond the evidence, these beliefs are not rationally *compelling* by virtue of substantive principles and ampliative rules. It can be easily seen that one can be a scientific realist and adopt van Fraassen's conception of rationality: belief in electrons and the like may well come out rational under van Fraassen's conception of rationality, since it is not rationally forbidden. But so may disbelief in them (or agnosticism about them). Hence, van Fraassen's conception of rationality is suitable for constructive empiricists in that it shows that belief solely in the empirical adequacy of theories is rational.

The point is not that CE is irrational – of course it isn't! Rather, it is that a) van Fraassen's conception of rationality is too liberal and b) there is still room for comparative judgement of rationality: some positions are more rational to occupy than others (cf. Darwinism and creationism). I have criticized van Fraassen's new conception of rationality elsewhere (see Psillos 2007).



### 3.2. Philosophical therapy?

As noted already, scientific realism does avoid Putnam's model-theoretic argument by buying into a certain metaphysics: a way of viewing the world as having a certain determinate natural structure.

This is clearly something van Fraassen is not happy with. But his own empiricist structuralism falls prey to Putnam's model-theoretic argument as well. For if an ideal theory cannot fail to be false – if, you may recall, there is always some satisfaction relation, SAT\*, between (sets of) pieces of the WORLD and terms and predicates of the language of 'ideal T' such that T comes out true of the WORLD – the very gap between empirical adequacy and truth collapses. So here is a dilemma: either CE has to buy into some substantive metaphysics (one that allows that there is a difference between empirical adequacy and truth), or CE collapses into realism. Note that it is not an option for CE to claim that SAT\* is not intended, because though the intended interpretation of the language is not fixed by the world, it is still fixed by the intentions and practices of the language users. This option, with which Putnam's internal realism flirted if it did not directly adopt it, is no less metaphysical than the straightforward realist enough. As van Fraassen (1997, 38) rightly notes, this option imputes unprecedented metaphysical powers to persons.

Van Fraassen claims that Putnam's model-theoretic argument can be dissolved, without adopting any suspicious looking metaphysical postulates. His central idea is captivatingly simple. Putnam, we have seen, equates truth with truth-in-an-interpretation, and since the latter is always available for a consistent theory, so is the former. Not so fast, van Fraassen replies. We know already that the equation between truth and truth-in-some-interpretation is illegitimate when it comes to our own language, the language we understand and use to make contact with the world. For our own language, we are not at liberty to pick any interpretation whatever. The interpretation is already fixed, as it were.

But which language *is* our own? This, van Fraassen says, is an indexical matter; it is the language we actually use and understand. If we lose sight of this pragmatic dimension of language use, van Fraassen (1997, 21) adds, we are tempted to think that the gap between truth and truth-in-an-interpretation 'might be filled by metaphysics'. But if we keep an eye on this dimension, there is no gap to be filled at all for our own language.

There is an obvious retort to this line of reasoning. Could we not consistently think that we might be wrong in interpreting our language the way we do? Could it not be the case that the extensions we assign to predicates in our language do not cut the world at its joints? If these are genuine possibilities, isn't there an opening for metaphysics to get in by virtue of the claim that it is the world that ultimately determines the *correct* interpretation of the language? This, you might recall, is in essence, Lewis's answer to Putnam's model-theoretic argument.

Van Fraassen's reply is most instructive. As I understand it, it goes like this: though it is indeed possible that our own language might have the wrong interpretation (it might fail to cut the world at its joints, as it were), we cannot (in our own language) coherently deny that it (the interpretation-in-use, as it were) is the right interpretation. This situation is supposed to be an instance of what van Fraassen calls pragmatic tautologies: propositions that can be false (i.e., they are not necessarily true) and yet are such that they cannot be coherently denied. Take the following two statements:

- (A) X believes that P, but it is not the case that P.
- (B) I believe that P, but it is not the case that P.

The person X in (A) might be myself. Hence, the content of (A) and (B) might well be the same. This shows that (B) is nothing like a formal contradiction. And yet (B) cannot be coherently asserted by me – it is a (pragmatic) contradiction. Van Fraassen's claim is that statements that fix the reference in one's own language (e.g., *cat* refers to cats) are pragmatic tautologies in one's own language: they cannot be coherently denied. As such, he thinks, they raise no metaphysical anxieties about how reference is fixed and what role the world might have in reference-fixing. When further metaphysical worries are raised, van Fraassen (1997, 39) says, they should be treated as needing 'philosophical therapy'.

I am not so sure. To fix our ideas let us make statements (A) and (B) more concrete:

- (A') X believes that X's language cuts the world at its joints, but it is not the case that X's language cuts the world at its joints.
- (B') I believe that my language cuts the world at its joints, but it is not the case that my language cuts the world at its joints.

(B') is supposed to be pragmatically incoherent, while (A') is not. But there is traffic between the two, as when I am in a reflective mode – when I treat myself as 'he', as it were. I know that (B') *might* be true, and I know that I can assert that it *is* actually true if I take the third-person perspective on myself – that is, to assert (A'). So I can raise the question whether *my* language cuts the world at its joints (equivalently, how is the reference of the linguistic items of *my* language fixed?), and I can at least attempt to answer it by letting the world do most of the work in answering it. That is, I can coherently let the world do most of the work in making the case that *natural* classes are the extensions of the predicates of my language. If indeed there is an objective criterion of rightness when it comes to reference-fixing – and both van Fraassen and Putnam think there is – this criterion has to hold for my language, too. But then it seems van Fraassen cannot so easily escape from metaphysics by going for the supposedly harmless pragmatic tautologies.

Ultimately, van Fraassen has had to turn against his former self and abandon the realist background that accompanied CE – in particular, the thought that truth is correspondence with reality. He says this much explicitly in his more recent work (2008, 252).

Once again we find ourselves with an idea akin to, of a piece with, the correspondence theory of truth, the idea that there is a user-independent relationship between words and things that determines whether a sentence is true or false. Such an idea cannot be carried through without postulating a good deal of ontological flora and fauna beyond concrete individuals. But we have discussed this issue sufficiently above, we don't need to repeat the argument against such presuppositions.

His thought is of a piece with the middle Putnam of internal realism, at least in so far as making the notion of truth-as-correspondence the culprit. But in the case of van Fraassen, we are not told with what to replace it.

Empiricism has always been anti-metaphysics. Yet, logical empiricism – which van Fraassen has attacked – aimed to occupy a position such that the critique of metaphysics left the world as described by science intact (see Psillos 2011b). Van Fraassen's CE has taken it to be the case that the world as described by science was too metaphysical for the empiricist to feel at home in it. Hence, CE marked a revisionist stance towards science, a stance according to which belief in unobservables – and not just

in a theory-free standpoint to view the world – was optional. Traditional empiricism took it that the critique of metaphysics was tied to a critique of language and of the limits of meaningful discourse. Van Fraassen disdained verificationism and its dependence on language. The irony is that the rescue from metaphysics comes again from, and through, language, though with a pragmatic account of it.

#### 4. Concluding thoughts

Complaining against the criticism that his middle views flirted with idealism, Putnam (1994, 462) noted that he never denied that our practices were ‘world-involving’. Van Fraassen, too, never doubted that the rightness of opinion depends on what the world is like (cf. 1989, 177). But for the world to be involved in any way whatever with our practices or for the world to be a certain way and not another, it is required that the world – even if understood in a relatively minimal fashion as whatever resists our theorizing – must have *some* structure; it must be, to some extent at least, ready made. Both Putnam and van Fraassen have resisted this image of the ‘ready made’ world. They have both seen this image as too metaphysical. Their resistance to realism has been motivated, at least to a considerable extent, by the thought that realism is wedded to inflationary metaphysics. And their recoil from realism has been motivated, at least partially, by the thought that ‘proper’ views of language and science should deliver us from metaphysics. It is still an open question – to me at least – exactly how much metaphysics (scientific) realism requires or implies, apart from whatever commitments are necessary for securing the Possibility of Divergence. If this is thin metaphysics, so be it!

#### References

- Boyd, Richard (1971), ‘Realism and Scientific Epistemology’ (unpublished typescript).  
Glymour, Clark (1982), ‘Conceptual Scheming or Confessions of a Metaphysical Realist’, *Synthese*, 51: 169–80.  
Jardine, Nick (1986), *The Fortunes of Inquiry*. Oxford: Clarendon Press.  
Lewis, David (1984), ‘Putnam’s Paradox’, *Australasian Journal of Philosophy*, 62: 221–36.



- Merrill, G. H. (1980), 'The Model-Theoretic Argument Against Realism', *Philosophy of Science*, 47: 69–81.
- Psillos, Stathis (1999), *Scientific Realism: How Science Tracks Truth*. London: Routledge.
- (2007), 'Putting a Bridle on Irrationality: An Appraisal of van Fraassen's New Epistemology', in B. Monton (ed.), *Images of Empiricism*. Oxford University Press, pp. 134–64.
- (2011a), 'Scientific Realism with a Humean Face', in Juha Saatsi and Steven French (eds), *The Continuum Companion to the Philosophy of Science*. London: Continuum, pp. 75–95.
- (2011b), 'Choosing the Realist Framework', *Synthese*, 190: 301–16.
- Putnam, Hilary (1962), 'What Theories Are Not', in E. Nagel, P. Suppes and A. Tarski (eds), *Logic, Methodology, and Philosophy of Science*. Palo Alto, CA: Stanford University Press. Reprinted in Putnam (1975), *Mathematics, Matter and Method*.
- (1963), '"Degree of Confirmation" and Inductive Logic', in P. Schilpp (ed.), *The Philosophy of Rudolf Carnap*. La Salle, IL: Open Court. Reprinted in Putnam (1975), *Mathematics, Matter and Method*.
- (1965), 'Craig's Theorem', *Journal of Philosophy*, 62. Reprinted in Putnam (1975), *Mathematics, Matter and Method*.
- (1971), *Philosophy of Logic*. London: George Allen and Unwin.
- (1975), *Mathematics, Matter and Method*, Philosophical Papers, vol. 1. Cambridge: Cambridge University Press.
- (1978), *Meaning and the Moral Sciences*. London: Routledge and Kegan Paul.
- (1980), 'Models and Reality', *Journal of Symbolic Logic*, 45: 464–82.
- (1981), *Reason, Truth and History*. Cambridge: Cambridge University Press.
- (1982), 'Three Kinds of Scientific Realism', *Philosophical Quarterly*, 32: 195–200.
- (1983), *Realism and Reason*, Philosophical Papers, vol. 3. Cambridge: Cambridge University Press.
- (1994), 'Sense, Nonsense and the Senses: An Inquiry into the Powers of the Human Mind', *Journal of Philosophy*, 91: 445–517.
- (1995), *Words and Life*. Cambridge, MA: Harvard University Press.
- Sankey, Howard (2008), *Scientific Realism and the Rationality of Science*. Aldershot: Ashgate.
- van Fraassen, Bas C. (1980), *The Scientific Image*. Oxford: Clarendon Press.
- (1989), *Laws and Symmetry*. Oxford: Clarendon Press.
- (1997), 'Putnam's Paradox: Metaphysical Realism Revamped and Evaded', *Philosophical Perspectives*, 11: 17–42.
- (2008), *Scientific Representation: Paradoxes of Perspective*. Oxford: Clarendon Press.