The Realist Turn in the Philosophy of Science

Stathis Psillos

1. Introduction

The 'realist turn' in the philosophy of science occurred in the 1970s and marked a shift from empiricist views concerning scientific theories and their relation to the world to realist ones. It was associated with what came to be known as the *explanationist defense of realism*, viz., the strategy of showing that the basic realist tenets offer the best explanation of the empirical and predictive successes of scientific theories. It was motivated by a move from verification and issues in semantics (how do theoretical terms get their meaning?) to abduction (*aka* inference to the best explanation) and issues in epistemology (do we have reasons to take scientific theories, literally understood, as truthlike?). Realism initiated an era of epistemic optimism: science is in the truth-business. Soon enough however, this optimistic stance was challenged by rival views which aimed to show that, even after the collapse of instrumentalism, realism is not the only game in town concerning science (this was the key objective of van Fraassen's (1980) constructive empiricism); or that realism is at odds with the history of science and, in particular, with a track record of false and abandoned, but otherwise empirically successful, theories (this was Mary Hesse's and Larry Laudan's historical induction). In reply to the historical challenge, realists became more selective in what they are realists about.

This chapter offers a narrative of the basic twists and turns of the realism debate after the realist turn. I will start with what preceded and initiated the turn, viz., instrumentalist construals of scientific theories. I will then move on to discuss the basic lines of development of the realist stance to science, focusing on one of its main challenges: the historical challenge.

2. Semantic realism

The current phase of the scientific realism debate—what I call the *epistemic phase*—started in the middle 1960s and was based on an important consensus, viz., *semantic realism*. This is the view that the vocabulary of scientific theories should be treated in a uniform way on the basis of standard referential semantics. In the early 1950s, the

dominant empiricist view was that theories are only *partially* interpreted, meaning that they are divided into two parts, one theoretical and one observational, such that only their observational part (expressed by means of a theory-free observational vocabulary) is fully interpreted, while the theoretical part is only partially meaningful on the basis of its deductive relations to the observational part. Being only partially interpreted, theoretical terms were not taken to be putatively referring to anything in the world and concomitant theoretical assertions were not taken to have truth-conditions. At best, t-terms and t-assertions (that is, the core of theories *qua* theories) were taken to be ways to systematize deductively a set of observational assertions, which could be independently testable by means of direct observations.[1]

This empiricist approach to the semantics of theories started to crumble when Herbert Feigl (1950) argued that in thinking about the content of theories, we should separate the issue of what makes a theory true (if it is true) and the issue of the evidence for its truth. A scientific theory can then have truth-conditions which make an essential reference to unobservable entities and their properties and relations even if the evidence for these truth-conditions is, by and large, observational. What empiricists had come to call the 'excess content' of theoretical discourse is captured by the fact that t-discourse is about unobservable entities.

If semantic realism is taken for granted, it seems that the question of realism takes care of itself. Theories cannot be proved to be true; nonetheless, they can be confirmed by empirical evidence (as both realists and empiricists agree). Given semantic realism, if scientific theories are well-confirmed, there are reasons to believe in the reality of the theoretical entities they posit. To hold a theory as well-confirmed is to accept that the entities posited by the theory are part of the furniture of the world. This kind of view was captured very eloquently by Wilfrid Sellars when he said (1963: 97):

> To have a good reason for holding a theory is ipso facto to have good reasons for holding that the entities postulated by the theory exist.

3. For instrumentalism

But in the middle 1950s an argument became available to the effect that t-terms are dispensable. If this were true, semantic realism would become irrelevant. The very

---

[1] Earlier empiricist approaches understood the truth-conditions of theoretical assertions reductively, but this is a different story. For a detailed discussion, see my (1999, chapter 1).

idea that theories purport to describe the world as it is in its unobservable parts—a central realist intuition about science—would become a non-starter. The argument was based on a theorem proved by logician William Craig, the philosophical application of which led to the statement that came to be known as Craig's Theorem: for any scientific theory T, T is replaceable by another (axiomatisable) theory Craig(T), consisting of all and only the theorems of T which are formulated in terms of the *observational vocabulary* V_O (Craig 1956). The gist of Craig's theorem is that a theory is a conservative extension of the deductive systematization of its observational consequences. This theorem was taken to capture the canonical form of instrumentalism.

Though it's hard to find philosophers who explicitly characterized themselves as instrumentalists,[2] Graig's theorem offered a boost to instrumentalism—the view that theories should be seen as (useful) instruments for the organization, classification and prediction of observable phenomena; hence that the 'cash value' of scientific theories is fully captured by what theories say about the observable world. Craig's theorem was taken to show that the whole body of theoretical commitments in science—those expressed by the theoretical vocabulary—were dispensable, since theoretical terms could be eliminated *en bloc*, without loss in the deductive connections among the observable consequences of the theory.

At roughly the same time, Carnap (1958) re-invented the so-called Ramsey-sentence. The idea goes back to Frank Ramsey (1929): the content of a theory is captured by a single existential statement, in which the theoretical predicates are replaced by bound (second-order existential) quantifiers. The Ramsey-sentence ^RT that replaces theory T has exactly the same observational consequences as T; it can play the same role as T in reasoning; it is truth-evaluable if there are entities that satisfy it; but since it dispenses altogether with theoretical vocabulary and refers to whatever entities satisfy it only by means of quantifiers, it was taken to remove the issue of the reference of theoretical terms/predicates. Hence, it was taken to present a neutral ground between realism and instrumentalism. Carnap enthusiastically jumped on this idea since he thought he could deflate the debate between realism and

---

[2] A notable exception is Philipp Frank (1932), whose instrumentalism, in modern terminology, is a form of non-cognitivism: theories are symbolic tools that do not (aim to) represent anything which is not antecedently given in experience.

instrumentalism as being merely about a choice of language. At the same time, he thought he could secure the proper empirical content of theories against commitment to physical unobservable entities. What is more, as Carnap was first to note, the very theory T can be written down as a conjunction of two parts: the Ramsey-sentence ^RT of T and the conditional $^RT{\rightarrow}T$, which came to known as *Carnap-sentence* and was taken to be a meaning postulate with no empirical content.

By the end of the 1950s, then, the semantic realist project seemed to be short-lived. The realist conception of theories had to sail between the Scylla of Craig's theorem and the Charybdis of Ramsey-sentences. The argument that theoretical discourse possesses excess content over observational discourse and it is putatively referential came under severe pressure since by either Craig's theorem or Ramsey-sentences, theoretical vocabulary was rendered dispensable without loss of (empirical) content.

Carl Hempel (1958) expressed this pessimist sentiment in the form of 'the theoretician's dilemma'. If the theoretical terms and principles of a theory do not serve their purpose of a deductive systematization of the empirical consequences of the theory, they are dispensable. But, given Craig's theorem (and Ramsey-sentences), even if they do serve their purpose, they can be dispensed with. Hence, the theoretical terms and principles of any theory are dispensable.

4. Negative and positive arguments for realism

An otherwise plausible defense of semantic realism (via meaning holism and the denial of the distinction between theory-based and observation-based vocabulary) was turning against an(other) important realist intuition: viz., that subsequent theories, as a rule, do better than their predecessors in representing the world. This is the context in which the first thorough defense of realism takes place in the work of Hilary Putnam.

4.1 Against instrumentalism

In his writings in the 1960s, Putnam aimed to motivate and defend realism first by arguing systematically *against* instrumentalist approaches to scientific theories.

Two of his arguments stick out. The first relates to Craig's Theorem-based instrumentalism. Putnam (1965) mounted a formidable attack on the philosophical significance of Craig's theorem arguing that a) theoretical terms are meaningful, taking their meaning from the theories in which they feature and b) scientists employ

4

terms like 'electron', 'virus', 'spacetime curvature' and so on—and advance relevant theories—because they wish to *talk about* electrons, viruses, the curvature of spacetime and so on; that is scientists want to find out about the *unobservable* world. Theoretical terms provide scientists with the necessary linguistic tools for talking about things they want to talk about.

Putnam's second argument relates to the role of theories in the confirmation of observational statements. The idea is that theories are often necessary for the establishment of *inductive* connections between seemingly unrelated observational statements. Here is Putnam's (1963) own example. Consider the prediction *H*: 'When two subcritical masses of U_{235} are slammed together to form a supercritical mass, there will be a nuclear explosion'. *H* could be re-written in an observational language—that is without the t-term 'Uranium$_{235}$'—as O_1: 'When two particular rocks are slammed together, an explosion will happen'. Consider now the available evidence, namely O_2: 'Up to now, when two rocks were put together nothing happened'. Given this, it follows that $prob(O_1/O_2)$ is very low, (if it can be determined at all). But consider the posterior probability of O_1 given the past evidence and the atomic theory T which *entails* that the uranium rocks would explode if critical mass were attained quickly enough. It is obvious that $prob(O_1/O_2\&T)$ is now determined and is much greater than $prob(O_1/O_2)$.

To the challenge of semantic holism and the implication of radical reference-variance, Putnam replied by developing Saul Kripke's causal theory of reference. In a number of papers in the 1970s (1973; 1974; 1975a), he extended this theory to cover the reference of natural-kind terms, physical-magnitude terms and theoretical terms. A key consequence of this causal theory is that semantic incommensurability is disposed of and the possibility of referential continuity in theory-change is safeguarded. If, for instance, the referent of the term 'electricity' is fixed causally, all different theories of electricity refer to, and dispute over, the same 'existentially given' magnitude, viz. *electricity*. The causal theory makes available a way to compare theories and to allow claims to the effect that the successor theory is more truthlike than its predecessors. Besides, it tallied with Putnam's considered view that the positive defense of realism is, by and large, an empirical (naturalistic) endeavour. The way the world is constituted and causally interacts with the language-users is an indispensable constraint on the theory and practice of fixing the reference (and

meaning) of the language used to talk about the world: the conceptual and linguistic categories scientists use to talk about the world are tuned to accommodate the causal structure of the world.

Given these of arguments, the negative case for scientific realism—viz., that instrumentalism fails patently to account for the role, scope and aim of scientific theories—was hard to resist.

4.2 For Realism

Putnam went further by offering a positive argument for scientific realism. In his (1975: 73) he penned the most famous argument for scientific realism—which has become known as the 'no miracles argument' (NMA). Here is the argument in full:

> 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 even 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.

There has been heated debate about this argument (see my 1999, chapter 4). For now, I want to explain the reference in it to Richard Boyd. In his widely circulated and discussed, but (still) unpublished, manuscript *Realism and Scientific Epistemology*, Boyd tied the defense of scientific realism to the best (or "the only plausible") explanation of the fact that scientific methodology has succeeded in producing predictively reliable theories.

Boyd viewed scientific realism as an *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 about current science only; 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 an historical truth—viz., that scientific theories have been remarkably successful in the prediction and control of natural phenomena—the defense of scientific realism can only be *a posteriori* and broadly empirical. Boyd should in fact be credited with the move that what came to be known as the explanationist defense of realism should be conducted within a broadly naturalistic framework.

5.The three theses of realism

In light of the Putnam-Boyd understanding, scientific realism c. 1980 incorporated three theses:

a. REFERENCE: Theoretical terms refer to unobservable entities;

b. TRUTH: Theories are (approximately) true; and

c. CONTINUITY: There is referential continuity in theory change.

REFERENCE encapsulates semantic realism, and more specifically a certain non-verificationist reading of scientific theories—what came to be known as a 'literal or face-value understanding' of theories. But REFERENCE also implies a certain *metaphysical* image of the world: as being populated by unobservable entities. REFERENCE implies that (an essential part of) the subject matter of science is the unobservable world. By the same token, however, the metaphysical dimension of scientific realism is captured by not (much) more than the assertion that theoretical entities are real (viz., that theoretical terms genuinely refer).

TRUTH takes realism beyond REFERENCE in asserting that t-entities (at least those referred to by t-terms featuring in *true* theories) *are* real: they populate the world. For both Boyd (1971; 1981) and (the 1970s) Putnam, TRUTH implies a certain understanding of *truth*, viz., truth as correspondence: to say that a theory is true is to say that it corresponds to reality. The chief motivation for such a conception of truth was explanationist. Putnam and Boyd insisted that truth (and 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).

That truth has an explanatory function in science is the key idea behind the 'no miracles' argument. To be sure, it is *approximate* truth that at best can be attributed to

scientific theories.[3] But the *logical* point behind the 'no miracles argument' is that the success of scientific methodology is best explained by the fact that the theories that indispensably inform this methodology are relevantly true—that is true in the respects that inform the employment of these methodologies. Some philosophers (e.g., Ghins 2001) have argued that it is not the truth of theory X that explains its empirical success, but the fact that entities and properties posited by X are *real*. True enough! Yet all that is required to move from reality to truth is semantic ascent.

TRUTH has notable metaphysical implications, viz., that scientific theories are answerable to the world and are made true by the world. The most congenial to realism way to develop this insight is by what I have come to call THE POSSIBILITY OF DIVERGENCE. The notion of correspondence is meant to capture the *asymmetric dependence* of the theories on the world. This asymmetry implies that though empirical success (even empirical adequacy) is a sign of truth, 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, successful and well-confirmed theory *might* be false simply because the world might not conform to it. This POSSIBILITY OF DIVERGENCE is meant to capture a modal fact of the world and in particular a sense in which the world is *independent* of theories, beliefs, warrants, epistemic practices etc. It requires a conception of truth which distances truth from certain epistemic notions (even idealized ones) such as being *ideally warrantedly assertible*. Hence, TRUTH implies that realism is committed to a non-epistemic conception of truth.[4]

Taken together REFERENCE and TRUTH imply a certain way to view the metaphysics of scientific realism. It's not enough for realism to argue that certain theoretical entities posited by scientific theories are real. They and their properties should be (part of) the truth-makers of theoretical assertions and they should be mind-independent (in the way suggested by THE POSSIBILITY OF DIVERGENCE).

CONTINUITY takes scientific realism beyond REFERENCE and TRUTH by capturing the all-important notion of convergence in theory-change. 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

---

[3] The 'story' of approximate truth and truthlikeness (note: these are distinct concepts) is long and complex. For an account see my (1999, chapter 11). See also Niiniluoto (1987) and Kuipers (2000).

[4] For a different take on the relation between theory and truth, see Devitt (1984).

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 may be observational terms, if there any such) refers (1978, 25).

Then, REFERENCE and TRUTH go by the board too.

Putnam took it, correctly and insightfully I think, that this kind of pessimistic argument calls for a distinctively *philosophical* answer, viz. a theory of reference which allows for referential continuity on theory-change. So the key point is not that the premise of the inductive argument is false. Rather it is that this kind of argument relies on the implicit assumption that there is radical reference variance in theory change; that is that, a t-term that features in different theories necessarily refers to distinct unobservable entities. So Putnam's diagnosis was that the historical challenge to realism he envisaged was a golden opportunity to articulate realism in a better way: realism should avoid some descriptivist *and* holistic theory of reference. For it is only on such a theory of reference that, as we have already noted in section 3, it becomes inevitable that that every time the theory changes, the meanings of *all* terms change, too; and given that reference is supposed to be fixed by descriptions, meaning change is taken to lead reference variance. It transpires then, that adopting a theory of reference, such as the causal theory, which allows for referential stability in theory-change is indispensable for CONTINUITY and scientific realism.[5]

6. Looking for a role for history

6.1 The principle of no privilege

Things did not turn out to be very easy for realism. If realism is an historical thesis, the history of science should be called in to support it or undermine it. In her (1976), Hesse advanced what she called "a principle of no privilege," according to which

---

[5] The pure causal theory of reference fails for various reasons, as I noted in my (1999, chapter 12). There, I articulated a causal-descriptive theory of reference as part of the realist toolbox. For more on this, see my (2012).

our own scientific theories are held to be as much subject to radical conceptual change as past theories are seen to be.

Hesse (1976: 266) put forward an argument that all theories are *false*.

> Every scientific system implies a conceptual classification of the world into an ontology of fundamental entities and properties - it is an attempt to answer the question "What is the world really made of?" But it is exactly these ontologies that are most subject to radical change throughout the history of science. Therefore in the spirit of the principle of no privilege, it seems that we must say either that all these ontologies are true, ie: we must give a realistic interpretation of all of them or we must say they are all false. But they cannot all be true in the same world, because they contain conflicting answers to the question 'What is the world made of?' Therefore they must all be false.

This argument, it should be clear, implies a substantial role for history of science. For unless there is a recognizable pattern of change in the 'ontology of fundamental entities and properties', it can always be argued that our current scientific theories are not subject to radical change. The rationale for the Principle of No Privilege is predominantly historical and hence its defense should be historical. As Hesse admitted, the Principle arises "from accepting the induction from the history of science" (1976: 271).

But this is precisely the problem with this Principle: it should be borne out by the history of theory-change in science that all these 'ontologies' have been incompatible with each other; hence they cannot all be true. Showing incompatibility presupposes a theory of reference of t-terms which does not allow that same or different terms featuring in different theories can nonetheless refer to the same entity in the world. And this is precisely the position already challenged by Putnam: it is simply question-begging to adopt a theory of reference which makes it inevitable that there is radical-reference variance in theory-change.[6]

---

[6] Hesse, like almost everyone else at that time, made a connection between reference and ontology in that the ontological commitments of the theory are reflected in the (putative) reference of its theoretical terms. Hence, whether or not there is continuity in ontology among successive theories was taken to be the same as the existence or not of referential continuity. In her argument, Hesse relied precisely on the possibility, "emphasized by revolutionaries," that, as she put it "all our theoretical terms will, in the natural course of scientific development, share the demise of phlogiston" (1976: 271).

To be sure, Hesse, like almost anyone else in this debate, shares the *intuition* that falsity cannot genuinely explain the successes of science. Hence she goes on to argue that there is some continuity in theory-change which is not restricted to the "accumulation of true observation sentences," but includes

> some theoretical sentences which are carried over fairly directly from a past theoretical framework to our own, that is, which do not depend for their truth on the existence and classification of particular hypothetical entities, but are nearer to pragmatic predictive test.

Interestingly, these statements include that "water is composed of discrete molecules of hydrogen and oxygen in definite proportions." This, she says, "is true, though we are not able to specify in ultimate terms what exactly molecules and atoms of water, hydrogen, and oxygen are (Newtonian, Daltonian, quantum, and relativistic field theories tell different stories about them)."

The issue then is this. Is there a sense in which "the revolutionary induction from the history of science about theory change" (Hesse 1976: 268) can be blocked by admitting that the continuity in theory change is *substantial*? Differently put, Hesse's argument says nothing about false theories being such that some of them are *truer* (in their theoretical assertions) than others. And this is precisely the option realists came to exploit.

6.2 Getting nearer to the truth

William Newton-Smith (1981) was perhaps the first to think that the history of science (better: the past track-record of science) could be used in *defence* of realism. He took realism to be committed to two theses

(1) theories are true or false in virtue of how the world is, and
(2) the point of the scientific enterprise is to discover explanatory truths about the world.

He then noted that (2) is under threat "if we reflect on the fact that all physical theories in the past have had their heyday and have eventually been rejected as false." And he added (ibid):

Indeed, there is inductive support for a pessimistic induction: any theory will be discovered to be false within, say 200 years of being propounded. We may think of some of our current theories as being true. But modesty requires us to assume that they are not so. For what is so special about the present? We have good inductive grounds for concluding that current theories—even our most favorite ones—will come to be seen to be false. Indeed the evidence might even be held to support the conclusion that no theory that will ever be discovered by the human race is strictly speaking true. So how can it be rational to pursue that which we have evidence for thinking can never be reached?

It should be obvious that part of the argument that Newton-Smith aimed to neutralize is Hesse's Principle of No Privilege, cast as a question: "*what is so special about the present?*" His reply to this argument was that realists should posit "an interim goal for the scientific enterprise. This is the goal of getting nearer the truth." If this is the goal, Newton-Smith argued, there is no reason to bother with the preceding induction: "its sting is removed." Accepting the pessimistic induction "is compatible with maintaining that current theories, while strictly speaking false, are getting nearer the truth."

But the role of the history of science in the defense of realism was suitably restricted to motivating what Newton-Smith called 'the animal farm move', viz., that though all theories are false, some are truer than others. He took it that what was needed to be defended was the thesis that if a theory T_2 has greater verisimilitude than a theory T_1, T_2 is likely to have greater observational success than T_1. And he advanced what he called a *transcendental* strategy in its defense, which, for all practical purposes I think, is a 'best explanation' strategy. The key argument was that there is an "undeniable fact" to be reckoned with, viz., that "in a mature science like physics, contemporary theories provided us with better predictions about the world than their predecessors and have placed us in a better position to manipulate that world." The reckoning came with the claim that if the 'greater verisimilitude' thesis is correct (that is, if theories "are increasing in truth-content without increasing in falsity-content"), then the increase in predictive power would be explained and be rendered expectable. This increase in predictive power "would be totally mystifying (…) if it were not for the fact that theories are capturing more and more truth about the world."

This kind of argument, plausible though it may be, dismisses the force of the pessimistic induction all too quickly. Not because Newton-Smith is wrong about the

need to focus on near or approximate truth rather than on (full and exact) truth. But because the pessimistic induction, if forceful at all, undercuts the explanatory link between success and approximate truth. Hence the realists needed to do some more work to restore this link.

6.3 A confutation of convergent realism

That more work was needed became obvious after the publication of Laudan's (1981). His history-based argument against realism was precisely meant to show how the link between success and truth is undermined by taking into account the history of science. Laudan formulated his argument via *reference*—a point alluded to in Putnam's formulation of realism. But he did aim to block the claim that there is an explanatory connection between (approximate) truth and success—a point raised by Newton-Smith's argument.

Laudan started with granting, "for the sake of argument" that if a theory is approximately true, then it will be successful. He then aimed to show that even if we granted this, "explanatory success" cannot be taken "as a rational warrant for a judgment of approximate truth." So his aim was to show that the realist thesis is not rationally warranted.

What is the structure of Laudan's argument? There is some controversy concerning this issue, but the thought has been that if we are to take seriously Laudan's "plethora" of theories that were "both successful and (so far as we can judge) non-referential with respect to many of their central explanatory concepts", then the argument is *inductive*. In particular:

(I)
There is a plethora of theories (ratio 6 to 1)[7] which were successful and yet not approximately true.
Therefore, it is highly probable that current theories will not be approximately true (despite their success).

---

[7] Laudan (1981: 35) noted the famous 6 to 1 ratio: "I daresay that for every highly successful theory in the past of science which we now believe to be a genuinely referring theory, one could find half a dozen once successful theories which we now regard as substantially non-referring."

Yet, this kind of argument has obvious flaws. Two are the most important, I think. The first is that the basis for induction is hard to assess. This does not just concern the 6:1 ratio—where does it come from? It also concerns the issue of how we individuate and count theories as well as how we judge success and referential failure. Unless we are clear on all these issues in advance of the inductive argument, we cannot even start putting together the inductive evidence for its conclusion.

The second flaw of (I) is that the conclusion is *too* strong. It does not just undercut the connection between success and approximate truth; it yields as a conclusion that it is more likely than not that current successful theories are not approximately true. Hence it makes it the case that there is rational warrant for the judgment that current theories are *not* approximately true. The flaw with this kind a of sweeping generalisation is precisely that it disregards totally the strong evidence there is for current theories—it renders it totally irrelevant to the issue of their likelihood of being true. Surely this is unwarranted. Not only because it disregards potentially important differences in the quality and quantity of evidence there is for current theories (differences that would justify treating current theories as more supported by available evidence than past theories were by the then available evidence); but also because it makes a mockery of looking for evidence for scientific theories! If I know that X is more likely than Y and that this relation cannot change by doing Z, there is no point in doing Z.

If the "plethora" of theories cannot warrant an inductive conclusion, what is its role in Laudan's argument? Note the stated aim of the argument, viz., to show "explanatory success" cannot be taken "as a rational warrant for a judgment of approximate truth".

For X to be a rational warrant for Y, X must offer good reasons to accept Y. Past experience of X being correlated with Y is a good reason to accept a future correlation (for non inductive skeptics, anyway). And conversely, if X and Y have not been correlated in the past, we are not warranted in expecting that they will be correlated currently or in the future. Note that this kind of reasoning does not render it false that X may go with Y currently or in the future. It just undermines the warrant for this kind of judgment or expectation. An alternative way to see the issue is this. Y is supposed to explain X (approximate truth is supposed to explain success). But if X and Y have not been correlated in the past, (if X has not been associated with Y, or if

14

(more strongly) X has been associated with not-Y), then the warrant for accepting Y as the (best) explanation of X is undercut.

6.4 The *divide et impera* strategy

If we think of Laudan's argument as a *warrant-remover argument* and if we also think that the fate of (past) theories should have a bearing on what we are warranted in accepting now, we should think differently. In my (1996; 1999 chapter 5) I argued that we should think of Laudan's argument as a *kind* of reductio. And by this, (somewhat confusingly I must now admit), I meant to imply that it is not a *proper* reductio. As I noted, Laudan's argument aimed to "discredit the claim that there is an explanatory connection between empirical success and truth-likeness" which would warrant the realist view that current successful theories are approximately true. If we view the argument this way, as a potential warrant-remover argument, then the past record of science does play a role in it, since it is meant to offer this warrant-remover. But Laudan was careful to be using the qualifier "so far as we can judge" repeatedly. Past theories are non-referential "so far as we can judge," that is by our own lights. This implied that past theories were false, "so far as we can judge." This means that if we accept current theories to be true, then "so far as we can judge" past theories cannot be true. All this is consistent with leaving it open that current theories are true or false. It just requires that it cannot be the case that both past theories and current ones are true.

So my (1996) reconstruction of Laudan's argument was as follows:

(P)
(A) Currently successful theories are approximately true.
(B) If currently successful theories are truth-like, then past theories cannot have been.
(C) These characteristically false theories were, nonetheless, empirically successful. (the 'historical gambit')
Hence, empirical success is not connected with truth-likeness and truth-likeness cannot explain success: the realist's potential warrant for (A) is defeated.

(B) is critical for the argument. It is meant to capture discontinuity in theory-change, which I put it thus (stated in the material mode): "Past theories are deemed not to

have been truth-like because the entities they posited are no longer believed to exist and/or because the laws and mechanisms they postulated are not part of our current theoretical description of the world."

In this setting, Laudan's 'historical gambit' (C) makes perfect sense. For unless there are past successful theories which are *warrantedly* deemed not to be truth-like "so far as we can judge," the previous premise cannot be sustained and the warrant-removing reductio fails. If premise (C) can be substantiated, success cannot be used to warrant the claim that current theories are true. And there is no way that this premise can be substantiated apart from looking at past successful theories and their fate. History of science is thereby essentially engaged.

I still think this is the best way to make sense of the challenge Laudan had in mind in a way that a) the fate of past theories is seriously taken into account and b) the argument is seen as warrant-removing. To respond then to this argument, realists needed to be selective in their commitments. This response has come to be known as the *divide et impera* strategy to refute Laudan's argument (see my 1996). The focus of this strategy was on rebutting the claim that the truth of current theories implies that past theories cannot be deemed truth-like.

Philip Kitcher (1993) and myself (1996; 1999) have argued that there are ways to distinguish between the 'good' and the 'bad' *parts* of past abandoned theories and to show that the 'good' parts—those that enjoyed evidential support, were not idle components and the like—were retained in subsequent theories. This kind of response suggests that there has been enough theoretical continuity in theory-change to warrant the realist claim that science is 'on the right track'. To be more precise, the realist strategy proceeds in two steps. The first is to make the claim of continuity (or convergence) plausible, viz., to show that there is continuity in theory-change and that this is not merely empirical continuity: substantive theoretical claims that featured in past theories and played a key role in their successes (especially novel predictions) have been incorporated in subsequent theories and continue to play an important role in making them empirically successful. But this first step does not establish that the convergence is to the *truth*. For this claim to be made plausible a second argument is needed, viz., that the emergence of this evolving-but-convergent network of theoretical assertions is best explained by the assumption that it is, by and large,

approximately true.[8]

7. Structural Realism

The selective realist trend started with the position that John Worrall (1989) dubbed 'structural realism'. This was an attempt to capitalize on the fact that despite the radical changes at the theoretical level, successor theories have tended to retain the mathematical structure of their predecessors. Worrall's thought was that theories can successfully represent the structure of the world, although they tend to be wrong in their claims about the entities they posit. As Worrall put it: the structural realist "insists that it is a mistake to think that we can ever 'understand' the nature of the basic furniture of the universe" (1989: 122). Then, in opposition to scientific realism, structural realism restricts the cognitive content of scientific theories to their mathematical structure together with their empirical consequences. But, in opposition to instrumentalism, structural realism suggests that the mathematical structure of a theory represents the structure of the world (real relations between things).

Unsurprisingly, the chief argument for structural realism is a (weak) version of the 'no miracles' argument. The key idea is that though successful novel predictions suggest that the theory has latched onto the world, it is only the *structure* of the world (as this is expressed by the mathematical structure of the theory) that the theory latches onto. Against the pessimistic induction, structural realism contends that there is continuity in theory-change, but this continuity is (again) at the level of mathematical structure. Hence, the 'carried over' mathematical structure of the theory correctly represents the structure of the world and this best explains the predictive success of a theory.

Now, if this kind of argument is to lend any credence to structural realism, it must be the case that the mathematical structure of a theory is somehow exclusively responsible for the predictive success of the theory. But, as I have argued in detail in my (1995), it is not true that the mathematical equations alone—devoid of their physical content—can give rise to any predictions.

If structural realism is to employ a version (no matter how weak) of the no-miracles argument in order to claim that retained mathematical equations reveal real relations in the world, it should also admit that some physical content—not

---

[8] The *divide et impera* strategy has generated considerable discussion. For some recent takes on it see Cordero (2011) and Vickers (2013).

necessarily empirical and low-level—is also retained. But such an admission would undercut the claim that the predictive success vindicates only the mathematical structure of a theory; by the same token, it would undercut the epistemic dichotomy between the structure and the content of a physical theory.

Structural realism was independently developed in the 1970s by Grover Maxwell (1970; 1970a) in an attempt to show that the Ramsey-sentence approach to theories need not lead to instrumentalism. He called 'Structural realism' the view that: i) scientific theories issue in existential commitments to unobservable entities and ii) all non-observational knowledge of unobservables is *structural knowledge*, i.e., knowledge not of their first-order (or intrinsic) properties, but rather of their higher-order (or structural) properties. The key idea here was that a Ramsey-sentence satisfies both conditions (i) and (ii). So we might say that, if true, the Ramsey-sentence ^RT gives us knowledge of the structure of the world: there is a certain structure which satisfies the Ramsey-sentence and the structure of the world (or of the relevant worldly domain) is isomorphic to this structure. It should be noted that Maxwell's point against Carnap was that the Ramsey-sentence approach to theories was amenable to a realist construal more than to an instrumentalist one.

Though initially Worrall's version of structural realism was different from Maxwell's, being focused on—and motivated by—Henri Poincaré's argument for structural continuity in theory-change[9], in later work Worrall came to adopt the Ramsey-sentence version of structural realism (see appendix IV of Zahar 2001). So what I (2006) have called Maxwellian-Worrallian structural realism asserts that the world has excess structure over the appearances, but this excess structure can be captured (hypothetico-deductively) by the Ramsey-sentence of an empirically adequate theory.

Recall from section 3, that Carnap's insight in the 1950s was that a scientific theory T is logically equivalent to the following conjunction: ^RT & ($^RT \rightarrow T$), where the Ramsey-sentence captures the factual content of the theory, and the conditional $^RT \rightarrow T$ captures its analytic content (it is a *meaning postulate*). Precisely because the so-called Carnap-sentence is analytic, Carnap thought that characterizing the excess content of a theory over its Ramsey-sentence as talking about certain unobservable entities (e.g., electrons) or as talking indifferently about *whatever* satisfied the theory

---

[9] For a detailed analysis of Poincaré's structural realism in relation to his conventionalism, see my (2014).

(even if these were taken to be numbers and sets thereof) was a matter of linguistic choice. But for a realist this cannot be a matter of choice of language. In any case, it turns out that if the Ramsey-sentence ^{R}T is true, the theory T must be true: it cannot *fail* to be true. Is there a sense in which ^{R}T can be false? A Ramsey-sentence may be empirically inadequate. Then it is false. But if it *is* empirically adequate (if, that is, the structure of observable phenomena is embedded in one of its models), then it is bound to be true. For, as Max Newman (1928) first noted in relation to Russell's structuralism, given some cardinality constraints, it is guaranteed that there is an interpretation of the variables of ^{R}T in the theory's intended domain.[10]

Though Carnap felt at home with this result since empiricism could thus accommodate the claim that theories are true, without going much beyond empirical adequacy, reducing truth to empirical adequacy *is* a problem for those who want to be realists, even if just about structure. For, it is no longer clear what has been left for someone to be realist about.

This is a pretty damaging objection to structural *realism*. The only way out is for structural realism to abandon pure structuralism and to treat structure as being defined by real or natural *relations*. Having first specified these natural relations, one may abstract away their content and study their structure. But if one begins with the structure, then one is in no position to tell *which* relations one studies and *whether* they are natural or not.[11]

8. Concluding thoughts

Four decades after the 'no miracles argument' and the 'pessimistic induction', where does the realism debate stand? It seems fair to say that a key realist claim, viz., that science does offer knowledge of the unobservable part of nature, has been vindicated. Currently, all sides of the debate—with the exception of constructive empiricism—admit that science does offer epistemic access to some unobservable parts of reality. Hence, the unobservable is not, *ipso facto*, epistemically inaccessible. Old empiricism-motivated claims that scientific knowledge is restricted to whatever is

---

[10] For more on this see my (1999; 2001 & 2006). See also Demopoulos (2003).

[11] For more on this see Psillos (2009); Ainsworth (2009); Cruse (2005); Cruse & Papineau (2002). Partly because of the failures of the standard (so-called 'epistemic') version of structural realism and partly because of independent reasons, an 'ontic' version of structuralism has acquired currency. I won't go into the debates around ontic structural realism. For an overview and recent developments see Ladyman & Ross (2007); French (2014) and Psillos (2009; 2016a).

given in immediate experience and observation hold no weight any more.[12]

By the same token, however, the rivalry to scientific realism has now shifted to the general point that there is a sharp epistemic division to be drawn *within* the unobservable; that is, between those aspects of the unobservable that are epistemically accessible and those that are not. Structural realists, for instance, draw the division between the knowable (unobservable) structure of nature and whatever is left to 'fill in' the structure—objects, entities, natures and the like. Contextual instrumentalists (Stanford 2006) draw the division between those entities to which there is an independent route of epistemic access (mediated by theories that cannot be subjected to serious doubt) and those entities to which all supposed epistemic access is mediated by high-level theories. The former are epistemically accessible, while the latter are said to be impenetrable. Semi-realists (cf. Chakravartty 2007) draw the division between detection properties and auxiliary properties of particulars; *and so on*.

The common denominator of all these dichotomous positions is this: there is a principled limit to the scientific knowledge of the world. (The limit is different in different positions, but it is always principled and definite). The realist victory is that this division is *within* the realm of the unobservable. But the realist defeat is that some aspect of the unobservable is, for principled reasons, inaccessible.

In my own work I have tried to argue that there is no good reason (either a priori or a posteriori) to think that there is a principled epistemic division between what can be known of nature and what cannot. There might be parts of nature that science might never be able to map out, but these do not fall nicely within a conceptual category which captures one side of a sharp epistemic dichotomy (the unknown X: the noumena; the non-structure; the intrinsic properties; the auxiliary properties; whatever-there-is-only-thin-epistemic access-to; whatever-there-is-only-theory-mediated-access-to; and the like).

Though the epistemic debate still goes on, the focus of attention has been shifting from epistemology towards metaphysics and ontology. The key question seems to be the following: if we take science seriously and if we take scientific theories as true, or approximately true, are we thereby committed to a certain way to understand the deep structure of the world? Are we committed to substantive accounts of causation, laws, necessity, properties and other key metaphysical categories? Or are deflationary

---

[12] It must be stressed, however, that logical empiricists defended some anti-metaphysical version of scientific realism—see my (2011).

accounts good enough? As I have put it in my (2013), the key contrast is between a neo-Aristotelian scientific realism and a neo-Humean one. I have personally sided with the neo-Humeans, but currently lots of interesting work is done on this front. The current enthusiasm for structuralism is a case in point.

References

Ainsworth, P. (2009) "Newman's Objection," *The British Journal for the Philosophy of Science* **60**: 135-171.

Boyd, R. (1981) "Scientific Realism and Naturalistic Epistemology," in P.D. Asquith & T. Nickles (eds) *PSA 1980*, Vol.2, East Lansing: Philosophy of Science Association.

----------. (1971) *Realism and Scientific Epistemology*, unpublished typescript.

Carnap, R. (1958) "Beobachtungssprache und Theoretische Sprache," *Dialectica*, **12**: 236-248. Trans as "Observation Language and Theoretical Language," in J. Hintikka (ed.) *Rudolf Carnap, Logical Empiricist* (1975), Dordrecht: Reidel

Chakravartty, A. (2007) *A Metaphysics for Scientific Realism: Knowing the Unobservable*, Cambridge: Cambridge University Press.

Cordero, A. (2011). "Scientific Realism and the Divide et Impera Strategy: The Ether Saga Revisited," *Philosophy of Science* **78**: 1120-1130.

Craig, William (1956) "Replacements of Auxiliary Assumptions," *The Philosophical Review* **65**: 38-55.

Cruse, P. (2005) "Ramsey-sentences, Structural Realism and Trivial Realisation," *Studies in History and Philosophy of Science* **36**: 557-576.

Cruse, P. & Papineau, D. (2002) "Scientific Realism Without Reference," in M. Marsonet (ed.) *The Problem of Realism*, Aldershot: Ashgate.

Demopoulos, W. (2003) "On the Rational Reconstruction of Our Theoretical Knowledge," *The British Journal for the Philosophy of Science* **54**: 371-403.

Devitt, M. (1984) *Realism and Truth*, (Second Revised edition 1991), Oxford: Blackwell

Feigl, H. (1950), "Existential hypotheses: Realistic versus Phenomenalistic Interpretations," *Philosophy of Science* **17**:35-62.

Frank, P. (1932) *The Law of Causality and its Limits*, (M. Neurath & R. S. Cohen trans) Dordrecht: Kluwer.

French, S. (2014) *The Structure of the World: Metaphysics and Representation*, Oxford: Oxford University Press.

Ghins, M. (2001) "Putnam's No-Miracle Argument: A Critique," in S. P. Clarke and T. D. Lyons (eds) *Recent Themes in the Philosophy of Science*, Kluwer Academic Publishers.

Hempel, C. (1958) "The Theoretician's Dilemma: A study in the Logic of Theory Construction," *Minnesota Studies in the Philosophy of Science*, **2**, Minneapolis: University of Minnesota Press.

Hesse, M. B. (1976) "Truth and Growth of Knowledge," in F. Suppe & P.D. Asquith (eds) PSA 1976 Vol.2, East Lansing: Philosophy of Science Association.

Kitcher, P. (1993) *The Advancement of Science*, Oxford: Oxford University Press.

Kuipers, T. A. F. (2000) *From Instrumentalism to Constructive Realism*, Dordrecht: Reidel.

Ladyman, James& Ross, Don (2007) *Every Thing Must Go*: *Metaphysics Naturalised*, Oxford: Oxford University Press.

Laudan, L. (1981), "A Confutation of Convergent Realism," *Philosophy of Science* **48**: 19-49.

Maxwell, G. (1970) "Theories, Perception and Structural Realism," in R. Colodny (ed.) *The Nature and Function of Scientific Theories*, Pittsburgh: University of Pittsburgh Press.

-----------. (1970a) "Structural Realism and the Meaning of Theoretical Terms," in *Analyses of Theories and Methods of Physics and Psychology*, Minnesota Studies in the Philosophy of Science, 4, Minneapolis: University of Minnesota Press.

Newman, M. H. A. (1928) "Mr. Russell's "Causal Theory of Perception"," *Mind* **37**: 137-148.

Newton-Smith, W.H. (1981) *The Rationality of Science*, London: RKP.

Niiniluoto, I. (1987) *Truthlikeness*, Dordrecht: D. Reidel Publishing Company.

Psillos, S. (2016a) "Broken Structuralism," *Metascience*.

-----------. (2014) "Conventions and Relations in *Poincaré's* Philosophy of Science," *Methode-Analytic Perspect*ives **3**: 98-140

-----------. (2013) "Semirealism *or* Neo-Aristotelianism?," *Erkenntnis*, **78**: 29-38

-----------. (2012) "Causal-descriptivism and the reference of theoretical terms," in A. Raftopoulos & P. Machamer (eds) *Perception, Realism and the Problem of Reference*, Cambridge University Press, pp. 212-238

-----------. (2011) "Choosing the Realist Framework," *Synthese* **190**: 301-316.

-----------. (2009). *Knowing the Structure of Nature*, London: Palgrave/MacMillan.

-----------. (2006) "Ramsey's Ramsey-sentences," in *Cambridge and Vienna: Frank P Ramsey and the Vienna Circle* (Vienna Circle Institute Yearbook 12), (ed. M. C. Galavotti), Springer, pp. 67-90

-----------. (2001) "Is Structural Realism Possible?," *Philosophy of Science* **68**: S13-24.

-----------. (1999) *Scientific Realism: How Science Tracks Truth*, London & New York: Routledge.

-----------. (1996) "Scientific Realism and the "Pessimistic Induction"," *Philosophy of Science* **63**: S306-14

-----------. (1995) "Is Structural Realism the Best of Both Worlds?," *Dialectica* **49**: 15-46

Putnam, H. (1963) ""Degree of Confirmation" and Inductive Logic," in P. Schilpp (ed.) *The Philosophy of Rudolf Carnap*, La Salle IL: Open Court — reprinted in *Mathematics, Matter and Method*.

-----------. (1965) "Craig's Theorem," *Journal of Philosophy*, **62** —reprinted in *Mathematics, Matter and Method.*

-----------. (1973) "Explanation and Reference," in G. Pearce & P. Maynard (eds) *Conceptual Change*, Dordrecht, Reidel.

-----------. (1974) "Philosophy of Language and Philosophy of Science," in R. S. Cohen et al. (eds) PSA1974, pp. 603-10.

-----------. (1975) *Mathematics, Matter and Method*, Philosophical Papers Vol.1, Cambridge: Cambridge University Press.

----------- (1975a) "The meaning of "Meaning"," in K. Gunderson (ed.), *Language, Mind and Knowledge*, Minnesota Studies in the Philosophy of Science, Vol. VII, Minneapolis: University of Minnesota Press, pp. 131-93.

-----------. (1978) *Meaning and the Moral Sciences*, Routledge and Kegan Paul.

Ramsey, F. (1929) "Theories," in *The Foundations of Mathematics and Other Essays*, (R. B. Braithwaite ed.) (1931) London: Routledge and Kegan Paul.

Sellars, W. (1963) *Science, Perception and Reality*, (re-issued 1991) Atascadero CA: Ridgeview Publishing Company.

Smart J. J. C. (1963) *Philosophy and Scientific Realism*, London: Routledge and Kegan Paul.

Stanford, P. K. (2006) *Exceeding Our Grasp: Science, History, and the Problem of Unconceived Alternatives*, Oxford: Oxford University Press.

van Fraassen, B. C. (1980) *The Scientific Image,* Oxford: Clarendon Press.

Vickers, P. (2013) "A Confrontation of Convergent Realism," *Philosophy of Science* **80**: 189-211.

Worrall, J. (1989) "Structural Realism: The Best of Both Worlds?," *Dialectica* **43**: 99-124.

Zahar, E. 2001. *Poincaré's Philosophy: From Conventionalism to Phenomenology*, La Salle IL: Open Court.