CrossMark

Towards a realistic success-to-truth inference for scientific realism

Peter Vickers[1]

Abstract A success-to-truth inference has always been at the heart of scientific realist positions. But all attempts to articulate the inference have met with very significant challenges. This paper reconstructs the evolution of this inference, and brings together a number of qualifications in an attempt to articulate a contemporary ('local') success-to-truth inference which is realistic. I argue that this contemporary version of the inference has a chance, at least, of overcoming the historical challenges which have been proffered to date (and without moving all the way to structural realism). However, there is a price to pay: the developments which help the realist answer the historical challenges also serve to increase the number of *non*-historical challenges.

Keywords Scientific realism · Confirmation · No miracles · Pessimistic induction · Structural realism · Divide et impera

1 Introduction

At the heart of any scientific realist position is a success-to-truth inference. As a first pass, the basic idea is this: where science is sufficiently successful—makes accurate predictions and/or exhibits significant explanatory power—the relevant theoretical hypotheses are (probably and/or approximately) true. However, both scientists and (realist) philosophers seem to be much too optimistic about the truth which can be inferred from successful predictions and explanations. There are umpteen examples of scientists making a doxastic commitment to a theory on the basis of its predictive/explanatory success, only to be forced to withdraw that commitment at a later

✉ Peter Vickers
    peter.vickers@durham.ac.uk

[1] University of Durham, 50/51 Old Elvet, Durham DH1 3HN, UK

date.[1] And even amongst philosophers of science who try to be much more nuanced vis-à-vis the doxastic commitments warranted by scientific success, there are historical cases seriously in tension with most of the (non-empty) success-to-truth inferences which have been proposed in the literature.

Given these facts, this paper considers the current status of the realist's success-to-truth inference. Is there a way to (once again) adjust the inference to accommodate the challenges from the history of science? Or should we now accept that we have here a degenerating research programme, and that the inference is fundamentally flawed? To help judge this, Sect. 2 is devoted to a history of the debate, starting with the flurry of activity in the 1970s and 1980s and the way in which the inference was adapted in response to objections. As further objections are presented in Sects. 3 and 4 further possible qualifications mount up including the 'divide et impera' move, the structural realist perspective, and an emphasis on (one or another) confirmation theory. In Sect. 5 various qualifications are brought together for the first time to present a realistic, contemporary, 'local' success-to-truth inference. I argue that this inference has a chance, at least, of resisting all known historical challenges (and without requiring a structural realist response to those challenges). However, various non-historical challenges remain, including several new challenges one inherits as a result of introducing the noted qualifications.

2 The trajectory of the inference

The success-to-truth inference can be traced back almost as far as you care to look. But a convenient starting point for the modern debate is Putnam's classic statement of the inference in 1975: "The positive argument for realism is that it is the only philosophy that doesn't make the success of science a miracle." (Putnam 1975, p. 73). Being a 'realist', for Putnam, simply meant committing to the truth of our best theories (in the correspondence-theory-of-truth sense).[2] And so we have here a success-to-truth inference as soon as we include the premise that to infer a miracle would be absurd.

Putnam's statement was immediately objectionable due to fairly straight forward historical 'counterexamples'.[3] Most famously, of course, Laudan explicitly targeted Putnam in his historical 'confutation'. For a counterexample to Putnam's success-to-truth inference all we need are examples where we have (significant) success, and yet what we infer certainly *isn't* true. As Laudan argues (1981, pp. 24–26), Putnam is happy to accept that, for example, 'aether' is a non-referring term, and thus theories employing this concept are not even approximately true. Laudan then argues that nineteenth century aether theories *were* successful theories (pp. 26–27). And he goes on to give his famous list of other possible/probable counterexamples (p. 33).

At this point it was fair to assume that Putnam's basic success-to-truth inference just needed refinement. Laudan (1981) worked quite hard to be charitable to Putnam,

---

[1] E.g. Einstein upon hearing of Bohr's successful prediction of the ionised helium spectral lines: "This is a tremendous result. The theory of Bohr must then be right." (cited in Pais 1991, p. 154).

[2] See e.g. Van Fraassen (1980, p. 8) and Laudan (1981) for discussion.

[3] Following standard practice in much of the literature, in this paper I use the locution 'counterexample to' in a loose sense, synonymous with 'in serious tension with'.

and pre-empted a few qualifications that Putnam endorsed but which didn't always come across in Putnam's writings. First and foremost, quite naturally, Putnam wanted to make an inference not to truth *simpliciter*, but rather to *approximate* truth.[4] Laudan knew this, but argued that (most of) the examples on his list were successful without being approximately true, on any reasonable account of 'approximate truth' known at the time (see especially Laudan 1981, Sect. 5). As he puts it (p. 33): "However the latter ['approximately true'] is understood, I take it that a realist would never want to say that a theory was approximately true if its central theoretical terms failed to refer." Thus nineteenth century aether theories apparently still stand as counterexamples to the success-to-truth inference.

A more promising qualification for the realist was to raise the bar for what counted as *sufficiently* successful science to warrant the success-to-truth inference. Again, Laudan (1981, p. 23) anticipated this move, noting that if the bar is raised too high it will turn out that most of science has *not* been successful after all. But he is happy to accept quite a stringent notion of success, writing that a theory counts as successful when "…it has functioned in a variety of explanatory contexts, has led to confirmed predictions and has been of broad explanatory scope." (ibid.). But here Laudan makes two mistakes. First, the realist *does* have the option to raise the bar for sufficient success. This might mean that we don't have warrant to make a success-to-truth inference in many cases where, previously, we thought we *did* have warrant. But the realist is not bound to the claim that all/most contemporary science is approximately true. Second, Laudan's construal of 'sufficient success' dramatically reduces the significance of his own list of proposed counterexamples to realism. In particular, success requires confirmed predictions, but (as has been repeatedly noted in the literature) most of the examples on Laudan's list don't appear to have enjoyed any predictive success.

This line of thinking led to a great emphasis on predictive success amongst realists from the 1980s right through to the present day.[5] So, for example, Musgrave (1985, p. 211) writes, "few, arguably none, of the theories cited [by Laudan] had any novel predictive success." However, whether it is 'few' or 'none' really matters to how the success-to-truth inference should be framed. If the inference is from novel predictive success to approximate truth, then just one example on Laudan's list is enough to undermine it. And since Musgrave it has gradually become widely accepted that certain examples on Laudan's list—caloric, phlogiston, nineteenth century aether theories— really *are* examples where novel predictive success was achieved.[6] Further, in recent years, a host of new examples of novel predictive success from theories generally

---

[4] Boyd had already made explicit reference to 'at least approximate' truth in 1973 (Boyd 1973, p. 1), and Putnam did so on several occasions (e.g. Putnam 1978, pp. 20–21). See also Laudan (1981), footnote 8. It should of course be noted here that by 'approximate truth' we really mean 'at least approximate truth, if not plain truth'. This qualification should be taken for granted in the following discussion.

[5] The correct way to think about 'predictions' has been a matter of much dispute—e.g. the debate between use-novel and temporally-novel predictive success—but that will not be a central theme here. See e.g. Chapters 2 and 3 of Psillos (1999).

[6] Lyons (2002) identifies various successful predictions enjoyed by the caloric theory of heat. Carrier (2004), Ladyman (2011), and Schurz (2011) have discussed the predictive successes of the phlogiston theory of oxygen. And most commentators have always accepted the predictive successes of nineteenth century aether theories (e.g. Worrall 1994; Psillos 1999, Chap. 6; Saatsi 2005).

considered to be *not* approximately true have been introduced to the literature (see below). With enough such examples—and absent a plethora of contrary examples—even an inference from novel predictive success to the *probable* approximate truth of the theory enjoying that success is undermined.

Establishing this result took some time. In the meantime, perhaps influenced by Laudan's focus on individual theories, the discussion gradually shifted away from a 'global' focus on the success of *science* (e.g. Putnam, Boyd) and towards a 'local' focus on the success of individual theories, or at least theories coming under a particular bracket.[7] This shift led some thinkers to consider whether realist commitment really ought to extend to the 'whole entire content' of an individual theory, and gradually an influential new qualification to the realist's success-to-truth inference emerged, sometimes described as the 'selective' or 'divide et impera' strategy. Instead of committing to the 'whole theory', the theory involved is divided into its 'working parts' (those parts actually responsible for the predictive successes), and its 'idle wheels' (those parts *not* involved in generating the predictions). With this distinction in place, the realist makes an inference from novel predictive success to the (probable) approximate truth of only the *working parts* of the theory enjoying that success. Writers often trace this idea back to Kitcher (1993), who introduced the 'working/idle' terminology. But we should go back at least as far as Worrall (1989), who made a distinction between the 'content' of a theoretical claim (which the realist need *not* be committed to given predictive success) and the 'structure' of a theoretical claim (which, according to Worrall, is what really does the work to bring about a prediction). Now there are various different ways to draw the distinction—and Worrall's particular proposal is just one (controversial) option—but nearly all modern realists accept *some* version of it in their success-to-truth inference.[8]

Does this help with the historical counterexamples to the realist's success-to-truth inference? It certainly isn't nearly so obvious that cases such as phlogiston, caloric, and nineteenth century aether theories present serious problems for the realist. And in addition, all of the realist's qualifications so far seem to be independently motivated (not just introduced ad hoc to answer an objection). So at this stage it looks like this research programme is still quite healthy, and not degenerating. However, we now reach some further historical challenges in the contemporary literature.

3 Contemporary challenges and the structural realist response

As it turns out, all of the qualifications introduced so far are *not* sufficient to answer the historical challenges. That is to say, further examples have come forward which apparently stand as firm counterexamples to the position we have reached: that given

---

[7] This shift of focus has been *very* gradual, and one still finds papers discussing realism from a 'global' perspective. However, a 'local' approach to the whole debate is today generally preferred. See for example Magnus and Callender (2004, especially Sect. 8), Turner (2004, especially Sect. 5), Psillos (2009, p. 65), Magnus (2010, especially Sect. 4), Saatsi and Vickers (2011, especially Sect. 5), Vickers (2013b, especially Sect. 8.3), Ruhmkorff (2013) and Peters (2014, especially Sect. 3).

[8] To his credit, Laudan also pre-empts this move by the realist—see e.g. Laudan (1981, p. 28). But he dismisses it too quickly. See e.g. Kitcher (1993, p. 143) for discussion.

novel predictive success we should infer the (approximate) truth of the working parts of the scientific theory enjoying such success. Or, better, we have counterexamples to nearly every interpretation of this claim—dependent, in particular, on how one understands 'working parts'.

To start with, there are cases which apparently show that novel predictive success cannot possibly be enough to motivate a realist commitment. Vickers (2013a) discusses the case of Velikovsky, and in particular his successful prediction that the surface of Venus must be 'hot'. As Vickers writes (p. 195f.), "Given just how fantastic Velikovsky's ideas were, the realist has no chance of arguing that the 'working parts' of Velikovsky's theory were approximately true." So the success-to-truth inference fails for this case, if we think the success is sufficient. But intuitively the success is *not* sufficient, even though it is undoubtedly a case of novel predictive success (given how 'novel predictive success' is defined in the literature). It is not sufficient, because the prediction is so vague—Velikovsky says nothing more specific than that the surface temperature of Venus will be 'hot'. This is, certainly, a falsifiable prediction, but it is not a very *risky* prediction precisely because it is so vague, and thus compatible with a *very* large number of possible observations. The lesson seems to be that the realist needs to include in her success-to-truth inference some clause concerning (e.g.) 'sufficient degree of risk' of a novel prediction.[9]

And yet even with such a clause in place, there are still highly problematic cases. Saatsi and Vickers (2011) discuss the case of Kirchhoff's theory of diffraction. Here the predictions are quantitative (very specific and hence very risky) and highly accurate, and yet (the authors argue) at least some of the working parts of Kirchhoff's theory are definitely not approximately true (see especially Sect. 4.2). Another example is discussed in Vickers (2012): the case of Sommerfeld's prediction of the frequencies of the hydrogen fine structure spectral lines. Again, the predictions are exceptionally accurate, but Sommerfeld drew heavily on a theory of the hydrogen atom involving relativistically adjusted elliptical orbits of the electron. These orbits lie at the heart of Sommerfeld's theory, such that it is not at all feasible to describe them as 'idle wheels'. And in addition it does not seem at all feasible to describe these features of Sommerfeld's theory as approximately true. A final example concerns Dirac's prediction of the positron, as discussed in Pashby (2012). Again, we have a very impressive (risky) prediction which—Pashby argues—issues from hypotheses which cannot feasibly be described as 'approximately true'.

This list will certainly be extended: new cases are coming forward all the time.[10] This leads to the worry that the realist cannot respond by merely inserting the word 'probably' into her success-to-truth inference. And one may well start to worry that it is getting harder to see how else the realist *can* modify her success-to-truth inference without including *ad hoc* adjustments which are designed merely to preserve a degenerating project.

---

[9] This sort of consideration is hardly new: for example, in Popper (1983) the measure of corroboration includes the important term P(e|hb) – P(e|b) (cf. Rowbottom 2013), and Bayesian updating similarly accommodates such considerations. See below.

[10] See e.g. Díez and Carman (2015) and Lyons (2016). See also the list in Vickers (2013a, Sect. 2).

One option is to adopt a very narrow view on which parts of a theory are truly 'working'. As noted above, Worrall makes the distinction in terms of the 'content' of a theoretical claim, which is idle, and the 'structure' of a theoretical claim, which is working and thus merits our doxastic commitment.[11] In other words, we should not believe what our best current scientific theories *say*, but instead we should believe that they have the 'structure' of the world right (or approximately right). What should be meant by 'structure' has been a topic of much debate, and a lot hangs on it. But whatever the answer, our commitments become highly abstract, and are best expressed using some combination of logic and mathematics as opposed to the natural language of science.

This move has polarised the realist community. Many realists have taken issue with structural realism on the grounds that there are purely theoretical scientific claims (claims concerning unobservables), which are *not* purely structural claims, and which warrant realist commitment if anything does. Psillos (1999, Chap. 7) argues in this vein, focusing on Fresnel's theory of light and the transition to Maxwell's electromagnetic theory. For Psillos, two such examples are (a) Fresnel's claim that the energy associated with a light wave is a function of the square of the wave's amplitude, and (b) the principle of conservation of energy.[12] However, Votsis (2011b) dismisses Psillos's claims here with two strategies central to defending structural realism: (i) to claim that what appears to be non-structural is actually structural, or has a 'structural analogue', and (ii) to claim that what appears to be purely theoretical (unobservable), is actually observable (with observability construed sufficiently broadly). Psillos's example (a) is dealt with by Votsis via a combination of both (i) and (ii), with Votsis claiming that it reduces to a purely mathematical relation combined with a variable 'the amplitude of light', which, for Votsis, is "a broadly construed observable quantity" because "it is the kind of quantity that can be measured" (Votsis 2011b, Sect. 5, fn. 11). And Psillos's example (b) is dealt with by Votsis in a similar way, reducing the principle of conservation of energy to "a mathematical relation between masses and velocities, two measureable and hence broadly construed observable quantities." (ibid.).

One might immediately worry about the broadness of Votsis's construal of 'observable'. For example, is the amplitude of light really observable? Well, of course, if you define 'observable' in the right way it counts as observable, but then the same goes for just about *any* theoretical property/entity one could mention! Votsis's proposal perhaps becomes more reasonable if one is able to retain Fresnel's success whilst swapping amplitude-talk for intensity-talk (making use of the relationship between the two), since light intensities are straight-forwardly observable. However, Votsis does want to say that the *amplitudes* are observable. The reason is simply that, in Votsis's view, amplitudes merit realist commitment given their (non-idle) role within the theory. At the end of Sect. 5 of his paper Votsis lays down a challenge to the 'traditional realist', to identify a non-structural element which merits realist commitment (as Psillos

---

[11] 'Working'/'idle' is not Worrall's terminology. But Worrall (1994, p. 339) makes a distinction between playing/not playing an integral role in the theory's empirical success.

[12] Fresnel actually expressed example (a) by making reference to the displacement of ether molecules, but he was committed to (a) in virtue of his explicit commitments. And it is only (a) that really plays a role in the success of Fresnel's theory, as Psillos notes (p. 159).

1999 attempts). For Votsis, the amplitudes are non-structural and do merit realist commitment, so the only way he can accommodate them is to regard them as (broadly speaking) 'observable'.

Other structural realists would surely try to regard amplitudes as idle, and this would then allow for a less controversial definition of 'observable' where amplitudes do count as *un*observables, but do not present a counterexample to structural realism because they do not merit realist commitment. Such a consideration exposes a problem with the overall debate here: the sheer flexibility of structural realism when it comes to avoiding counterexamples. Indeed, there is fierce debate concerning all of the key concepts: 'structure', 'observable', 'working/idle', etc. The word 'structure' itself is used in so many significantly different ways that advocates of structural realism find it hard to agree on anything at all. This is accordingly to Votsis himself![13] A symptom of this is that even academics heavily involved in the scientific realism debate struggle to communicate with each other about 'structural realism'. Psillos's examples (a) and (b), above, seem to Psillos clearly non-structural, and yet for Votsis they should be interpreted structurally.

Consider another possible counterexample to structural realism, courtesy of Saatsi (2005), again focusing on the Fresnel–Maxwell theory shift. Saatsi identifies certain spatiotemporal properties of light, which in his view are clearly non-structural in their content. He writes: "[I]t should be *obvious* that the minimal derivation [Saatsi's reconstruction of Fresnel's reasoning after idle wheels are eliminated] appeals to crucial unobservable properties and theoretical principles besides formal, logico-mathematical structure" (Sect. 3.4, my emphasis). But is it obvious? No doubt Votsis would dismiss this example too on the grounds that Saatsi's unobservable properties and theoretical principles can be substituted for 'structural analogues', consisting of mathematical relations between quantities which are 'observable' (broadly speaking). Saatsi's unobservable properties are, after all, expressed using equations. Saatsi insists that the equations are aids for expressing *non*-structural theoretical hypotheses concerning the *nature* of light. But these examples seem no more or less structural than the examples Psillos (1999) puts forward, and which Votsis dismisses in his (2011b).[14]

Why in any case would realists make such a strong claim, that in *any case whatsoever* all we can *ever* have knowledge of (concerning unobservables) is 'structure'? Why could it not be the case that, say, *often* we can only acquire structural knowledge, but *sometimes* our knowledge can go (modestly) beyond the purely mathematical and logical? Peters (2014) takes issue with all realist accounts which insist on there only being one 'special' *type* of theoretical element which can be deserving of realist commitment, instead favouring a 'neutral' realism which allows for realist commitment to all different types of theoretical entity. Thus he criticises structural realism, entity realism, and phenomenological realism. Peters puts forward two main arguments (Sect. 3.3): (i) these 'special' realisms are usually motivated by historical cases which are cherry-picked to favour them—e.g. structural realism is usually motivated

---

[13] See Votsis (2011c, fn. 2).

[14] Peters (2014, Sect. 3.3), suggests that theories in modern molecular biology that rely on lock-and-key models of molecular interaction would present a problem for structural realists. But again, I'm certain Votsis would argue for 'structural analogues'.

using theories which are substantially mathematical in their content, and (ii) the criteria used to argue for realist commitment to structure will (at least sometimes) also commit realists to non-structure (mutatis mutandis for other 'special' accounts).

No doubt structural realists would at this point bring in the 'upward path to structural realism' (e.g. Votsis 2005). According to this argument, structure is all we can ever hope to know on basic epistemological grounds, thinking now about the fundamental gap between the truly fundamental nature of reality and the limited perceptual capacities of the human being (the limited information which actually reaches us from that fundamental reality). However, even if this is right, the beauty of a 'neutral' realism is that it is consistent with our knowledge of unobservables being purely structural. It is also consistent, however, with our knowledge going beyond that; in other words it is silent on this issue. One may worry that this necessarily means that a 'neutral' realism is too abstract and non-committal to successfully respond to the contemporary historical challenges. But in fact—even laying structural realism to one side—the realist has some significant cards left to play.[15]

4 A confirmationist turn

The Velikovsky case provides a clue to a useful development of the realist's success-to-truth inference. In the previous section I dismissed the Velikovsky case on the grounds that the prediction is not risky enough to merit a realist commitment (even though it is novel). However, although perhaps the prediction is not *very* risky, it is risky *to an extent*. Indeed, in general, there are all possible degrees of riskiness of prediction, with corresponding degrees of significance for the realist. But how can the realist's success-to-truth inference accommodate such degrees of significance?

The realism debate is so often conducted in stark black and white terms when it comes to realist commitment: under what circumstances should/shouldn't we make a realist commitment? At what degree of success have we 'crossed a threshold' such that a realist commitment is warranted?[16] These questions would make no sense to a Bayesian. The obvious answer to the Velikovsky case is to incorporate degrees of belief into the account.[17] Velikovsky's prediction is not *very* risky, but, when it turns

---

[15] Of course, in this section I have focused only on Votsis, as a prominent example of a structural realist. Other structural realists—e.g. French, Ladyman, Worrall—hold different positions and offer different arguments. My intention here is not to tackle all such positions and arguments, but rather to give an indication for why realists might want to answer the historical challenge to realism without having to adopt a structural realist position.

[16] Stanford (2009, p. 384) uses this question to introduce a 'threshold problem' for scientific realism. But the noted 'problem' assumes from the start a yes/no picture of realist commitment. Cf. Vickers (2013a, p. 197).

[17] Some may prefer to handle the Velikovsky case by dismissing it as 'immature' or 'pseudoscience' which shouldn't be taken seriously by the scientific community regardless of how many successes it has. But this approach has two serious (overlapping) problems: (i) it puts a lot of weight on the problematic distinction between science and pseudoscience (or the distinction between 'immature' and 'mature' science), and (ii) it is apparently subject to counterexamples from the history of science, where a theory which was initially 'immature' or appeared to be 'pseudoscience' came gradually to be taken very seriously as evidence accumulated.

out to be successful, that is still a confirmation of Velikovsky's theory that cannot be completely dismissed. And it is precisely such an emphasis on *confirmation* that will enable the realist to answer some of the most serious historical challenges.[18]

There are various theories of confirmation the realist could tap into. My goal here is not to decide between them; just for the purposes of illustration I will discuss how the realist could draw on Bayesian confirmation theory. As is well known, according to this account belief in (the approximate truth of) a theory T is not a black-and-white affair, but rather one has a *degree* of belief in T. And one updates that degree of belief in accordance with the usual formula, interpreting probabilities as degrees of belief. However, in the context of the present discussion the realist of course wants to talk about *working parts* of theories, and not simply theories. Thus I swap the usual 'T' for 'W' in the Bayesian formula, where 'W' stands for 'the working parts of T':

$$P(W|E) = \frac{P(E|W)P(W)}{P(E)} = \frac{P(E|W)P(W)}{P(E|W)P(W) + P(E|\neg W)P(\neg W)}$$

Putting to one side some potentially awkward questions (see next section) this formula can be a very useful tool for a realist. In particular the realist can make use of the fact that our degree of belief should increase very significantly only when $P(E)$ is small. In other words, we increase our degree of belief significantly only when the empirical result in question would be deemed (highly) unlikely were the working parts in question false: $P(E|\neg W)$ is small.[19] But in the case of Velikovsky's prediction that the surface temperature of Venus would be 'hot', the probability of this being correct in the event that Velikovsky's working hypotheses are false is not very small. Just to illustrate, suppose we put this probability at 0.45: on balance we don't expect the surface to be hot, but it's pretty 50:50 as we don't have much evidence either way.[20] And our initial degree of belief in the working parts is 0.2 (we are initially quite sceptical of the relevant hypotheses). Further, the result is not entailed by Velikovsky's theory, so suppose we put $P(E|W)$ equal to 0.8. Crunching the numbers, our degree of belief increases from 0.2 to only 0.31. In other words the realist accepts that in this case our degree of belief *should* increase, but that the increase falls short of warranting any success-to-truth inference. Instead this modest success entails a modest increase in one's degree of belief.[21]

What does the realist's success-to-truth inference now look like? Consider the following as a 'first try':

---

[18] Some actors in the realism debate are already on board with this proposal: see e.g. Howson (2000, 2013) and Psillos (2009, p. 56ff). But most are not.

[19] Needless to say, the Bayesian realist would need to offer some solution to the 'problem of old evidence', as well as taking a stance on various other issues concerning Bayesian confirmation theory.

[20] In 1950 some scientists had suggested that the surface temperature of Venus would be relatively cold, reasoning that the thick cloud cover would reflect a significant proportion of the sun's radiation.

[21] Howson (2000, p. 57, 2013) and Magnus and Callender (2004) make the further important point that if P(T) starts out very small, then even if the prediction *is* very risky/impressive, and $P(E|\neg T)$ is very small, then $P(T|E)$ will still be quite small. In the case of Velikovsky, if we are absolutely convinced that the theory is radically false from the start, then even an 'amazing' novel predictive success will not bring us to believe the theory (nor even, probably, its 'working parts').

> One should believe the *working parts* of a theory to be approximately true just when one's degree of belief in them becomes very high, and one's degree of belief should be proportioned to the evidence in the following way: whenever a theory achieves novel predictive success, one's degree of belief should increase to the degree that the prediction is risky (the empirical result predicted has a low prior probability).[22]

This leaves out the issue of whether the theory actually entails the empirical result in question: $P(E|T)$ in the Bayes formula. But I assume for the time being that we're talking about cases where this is equal to 1 (the theory entails the prediction), or where it is at least very high.

Dramatically, this new take on the success-to-truth inference still doesn't answer the contemporary challenge from the history of science as it stands. Prima facie (at least) the cases of Kirchhoff's theory of diffraction, Sommerfeld's theory of the hydrogen fine structure, and Dirac's prediction of the positron, all turn out to be cases where one's degree of belief should increase *very* dramatically following the predictive success. In other words, the prior probability of the relevant empirical result is *very* low (e.g. because the prediction is to several significant figures) and, according to the noted criteria, this means that one's degree of belief should become high enough for us to describe one's commitment as *plain belief*. In other words, one comes to believe that the theory's working posits are approximately true. But in these cases that would be a mistake, at least for most accounts of 'working'.

Can the realist hope to adjust her success-to-truth inference even further, without it looking like a desperate attempt to overcome the historical challenge in any way however ad hoc? In fact the emphasis on confirmation provides the realist with another very important card to play. The success-to-truth inference in the realism debate has always rested upon the 'no miracles argument'. The argument—in its 'local' or 'retail' form[23]—states that a novel predictive success issuing from a completely false theory would be tantamount to a miracle. Therefore, if a theory achieves novel predictive success, we ought to believe that theory (or at least its 'working parts') to be approximately true. But as it stands this allows for the following: we ought to believe in the working parts of a theory which makes hundreds of totally false (even bizarre) predictions so long as it has achieved one novel predictive success. This is absolutely anathema to the Bayesian updating approach. According to this approach we settle our degree of belief taking into account all of the successes *and failures* of a given theory. The formula works iteratively: each *posterior* degree of belief $P(W|E)$ is plugged back into the formula as a *prior* degree of belief $P(W)$ when we consider the next relevant empirical result. A single novel predictive success taken on its own might increase our degree of belief dramatically, but several disconfirmations of the

---

[22] Depending on the account of confirmation favoured, some realists could simply include what Hájek and Joyce (2008) call 'Surprisingness': 'For fixed values of P(E|H) and P(H) with P(E | H) > P(E), the degree to which E confirms H decreases with increases in P(E).'

[23] The 'global' or 'wholesale' form of the no miracles argument states that the *overall* success of science would be a miracle if our scientific theories were not approximately true (generally speaking). As already noted above, 'global' perspectives on the realism debate have gradually gone out of favour, replaced by 'local' perspectives.

theory will immediately bring that degree of belief back down (assuming for present purposes that the Duhem–Quine thesis is appropriately taken into account). Indeed, if one avoids a 'knee-jerk' account of doxastic updating, and instead considers a body of evidence (evidence *against* as well as evidence *for*) over a period of time, then the realist never believes the hypotheses in question despite the novel predictive successes they give rise to.

The case of Velikovsky is again instructive here. Somebody might protest against my discussion, above, where I suggested that Velikovsky's prediction of the surface temperature of Venus should only marginally increase our degree of belief in the theory (e.g. from 0.2 to 0.31). Perhaps my numbers were unrepresentative: $P(E|\neg T)$ should have been 0.3 (say), and $P(E|T)$ should have been 1 (perhaps). If so, our degree of belief would increase from 0.2 to 0.45, and if we started at 0.5 our degree of belief would increase to 0.77. This makes the success look more significant, but it still pales into insignificance when one considers all the empirical *failures* of the theory. Just to give one example, Velikovsky (1950) claims not only that Venus should be hot, but also (and for the same reasons) that Mars should be hot (pp. 367–368). As Sagan (1979, p. 113) points out, Mars is not 'hot' (on any reasonable interpretation of 'hot'), and this was even known before 1950. How should this affect our degree of belief in the theory? Suppose, following the Venus success, we start with $P(T) = 0.77$. Now, the probability of Mars *not* being hot if the theory is true is very small, because the theory predicts that Mars *will* be hot. Suppose it doesn't entail it, but the probability is 0.9. This means that $P(E|T) = 0.1$, where the empirical result is that Mars is *not* hot. To generate a result with the Bayes formula we also need to know $P(E|\neg T)$; in other words, we need to know the probability that Mars will not be hot on the assumption that Velikovsky's theory is false. This must be very high, since without the theory there is no reason to assume that Mars will be 'hot', and every reason to assume that it will be precisely the temperature it should be given its distance from the sun. Thus it makes sense to set $P(E|\neg T)$ at 0.9. Crunching the numbers, we find that, following the Mars result, our degree of belief in Velikovsky's theory drops dramatically from 0.77 to 0.27.

The point here is not that we should follow these numbers precisely, nor even that scientific realists should be Bayesians. What the above calculations illustrate, however, is that a single novel predictive success should not, on its own, convince us of the approximate truth of a theory's working posits. Instead it is crucial to consider also any dramatic *dis*confirmations of those same working posits, and adjust one's credence in the working posits accordingly. The case of Velikovsky is an easy case, but this same thinking can be applied to at least some of the remaining serious historical challenges to contemporary realism. E.g. Sommerfeld's predictive success concerning the fine structure of hydrogen no doubt should increase our degree of belief dramatically, but the failure of his theory when it comes to helium (or indeed *any* element heavier than hydrogen) should surely decrease our degree of belief. How much isn't clear, and Bayes' formula might not be able to help us much here. But the fact remains that the successes of Sommerfeld's theory are half the story at most, and a simple inference from success (however impressive) to truth cannot be part of a defensible realist position.

5 Taking stock

It's time to take stock and consider what a realistic success-to-truth inference might look like. The discussion so far has led to the following *Qualified Realist Statement* (fine tuning aside):

> One should believe the *working posits* of a theory to be (approximately) true just when one's degree of belief in those working posits becomes very high, and one's degree of belief should be proportioned to the evidence in the following way: whenever a theory achieves novel predictive success, one's degree of belief should increase to the degree that the prediction is risky (the empirical result predicted has a low prior probability), but at the same time *dis*confirmations should *decrease* one's degree of belief: (a) to the extent that the *negation* of the empirical result in question is expected if the working posits are approximately true (other things being equal, the greater this expectation the greater the disconfirmation), and also (b) to the extent that the actual empirical result is expected if the working posits are false (other things being equal, the greater this value the greater the disconfirmation).[24]

We've come a long way from the original 'no miracles' formulation of Boyd, Putnam, Leplin and others in the 1970s and 80s. But each qualification is independently motivated *even if* each qualification was in fact introduced in response to an anti-realist objection:

(i)   Success can't feasibly indicate absolute truth, but at best approximate truth.

(i)   There is no warrant for belief in the truly 'idle' parts of a theory which has been successful.

(iii) Success, even novel predictive success, comes in degrees (cf. Fahrbach 2011); for example, the realist should be influenced to the extent that a prediction is 'risky'.

(iv) Disconfirmations of a theory's working posits should reduce our confidence in those working posits (even if those same working posits are responsible for dramatic predictive successes).

Each of these qualifications is 'out there' in the literature, but realists have thus-far failed to bring it all together to fully articulate their position. No surprise, then, that anti-realists continue to bring historical evidence to bear on realism which would indeed be problematic for early versions of the success-to-truth inference, but which poses no problem when the noted qualifications are added.[25]

---

[24] Howson's (2013) criticisms of the 'no miracles argument' fail to threaten this formulation of the success-to-truth inference. Howson makes no attempt to address the selective realist move, and doesn't consider cases where several successes accumulate. Nor does he consider the fact that in real science, to be considered by the scientific community at all, a theory must have a significant prior probability much(!) higher than the hypothetical example Howson discusses of a theory with a prior probability of 0.001.

[25] The general move from a 'strong' realist statement to a 'weak' realist statement has been popular amongst realists over the past 40 years or so. As Devitt (1997, p. 109) puts it, "the weaker the realism, the more easy it is to defend." However, Devitt defends a weakened form of realism "for tactical reasons" (ibid.). In my opinion realists should embrace a weak form of realism not at all for *tactical* reasons, but rather because of the strength of the arguments against strong versions of realism (including the historical evidence).

Without question, major challenges lie ahead concerning the above Qualified Realist Statement. The qualifications do help to defend against attacks, but they also introduce new challenges. There continues a heated debate concerning how to identify the 'working posits', and indeed whether it is even possible to identify working posits except retrospectively following a period of scientific change.[26] The antirealist will no doubt be concerned, too, that no historical challenge *can* threaten the realist given how abstract and vague her commitments become when 'working posits' is interpreted structurally, and the specific definition of 'structure' is left vague. Then there is the long-standing concern that appeal to 'approximate truth' is empty so long as the realist cannot offer a theory of 'approximate truth' which has not met with severe criticism. In addition the confirmationist turn means that the realist cannot help but get embroiled in the complexities of debates concerning confirmation theories. If the realist takes the Bayesian approach there are all the usual challenges to consider: the question of subjective versus objective Bayesianism, the problem of old evidence, the problem of the priors, whether we are talking about group-level degrees of belief or degrees of belief for individual scientists, how to accommodate the Duhem-Quine thesis, and so on.

There is also the added difficulty of articulating exactly how Bayesian epistemology can accommodate a focus on 'working posits' as opposed to theories. Swapping 'T' for 'W', as above, is just a quick fix for a potentially thorny problem. What happens, for example, when the set of working posits varies depending on the specific prediction we are considering, with some posits featuring more often than others?[27] This strikes me as quite realistic, and will result in all different degrees of belief for all different hypotheses at a given time (even for a single scientist). For any individual prediction of 'E', $P(W)$ must then apparently refer to the product of one's credences in the working posits in play. And any increase from $P(W)$ to $P(W, E)$ will then need to be distributed fairly across each of the working posits. There are significant complexities lurking here.

This leaves the realist with a lot of work to do. But the historical challenge to realism has been a major problem for a very long time. The Qualified Realist Statement, above, might just be consistent with all the historical challenges which have been brought to bear. And this isn't because it is impervious to historical counter-example: if all the conditions were satisfied the realist *could* end up committing doxastically to something that was not even approximately true (on any reasonable interpretation of 'approximate truth'). For example, in light of the Qualified Realist Statement the case of Dirac's prediction of the positron needs to be re-investigated, and it *could* turn out that this case remains highly problematic. But if not, and all the historical cases noted above are consistent with the Qualified Realist Statement, then perhaps the philosophical price to pay for adopting the position is outweighed by the gains.

---

[26] For recent optimism see Votsis (2011a) and Vickers (2013a, Sect. 6). For scepticism from an antirealist see Stanford (2003, p. 915; 2006, pp. 173–180; 2009, p. 385); for scepticism from a realist see Saatsi (2015).

[27] Thanks to an anonymous referee for pushing me on this point.

References

Boyd, R. (1973). Realism, underdetermination, and a causal theory of evidence. *Nous*, *7*, 1–12.

Carrier, M. (2004). Experimental success and the revelation of reality: The miracle argument for scientific realism. In M. Carrier, et al. (Eds.), *Knowledge and the world: Challenges beyond the science wars* (pp. 137–161). Heidelberg: Springer.

Devitt, M. (1997). *Realism and truth* (2nd ed.). Princeton, NJ: Princeton University Press.

Díez, J. A., & Carman, C. (2015). Did Ptolemy make novel predictions? Launching Ptolemaic astronomy into the scientific realism debate. *Studies in the History and Philosophy of Science*, *52*, 20–34.

Fahrbach, L. (2011). Theory change and degrees of success. *Philosophy of Science*, *78*(5), 1283–1292.

Hájek, A., & Joyce, J. M. (2008). Confirmation. In S. Psillos & M. Curd (Eds.), *The Routledge companion to the philosophy of science* (pp. 115–128). London: Routledge.

Howson, C. (2000). *Hume's problem: Induction and the justification of belief*. Oxford: Clarendon Press.

Howson, C. (2013). Exhuming the no miracles argument. *Analysis*, *73*(2), 205–2011.

Kitcher, P. (1993). *The advancement of science: Science without legend, objectivity without illusions*. Oxford: Oxford University Press.

Ladyman, J. (2011). Structural realism versus standard scientific realism: The case of phlogiston and dephlogisticated air. *Synthese*, *180*(2), 87–101.

Laudan, L. (1981). A confutation of convergent realism. *Philosophy of Science*, *48*, 19–48.

Lyons, T. D. (2002). Scientific realism and the pessimistic meta-modus tollens. In S. Clarke & T. D. Lyons (Eds.), *Recent themes in the philosophy of science: Scientific realism and commonsense* (pp. 63–90). Dordrecht: Kluwer.

Lyons, T. D. (2016). Structural realism versus deployment realism: A comparative evaluation. *Studies in History and Philosophy of Science*. doi:10.1016/j.shpsa.2016.06.006.

Magnus, P. D., & Callender, C. (2004). Realist Ennui and the Base Rate Fallacy. *Philosophy of Science*, *71*(3), 320–338.

Magnus, P. D. (2010). Inductions, red herrings, and the best explanation for the mixed record of science. *British Journal for the Philosophy of Science*, *61*(4), 803–819.

Musgrave, A. (1985). Realism versus constructive empiricism. In P. M. Churchland & C. A. Hooker (Eds.), *Images of Science* (pp. 197–221). Chicago, IL: Chicago University Press.

Pais, A. (1991). *Niels Bohr's times*. Oxford: Oxford University Press.

Pashby, T. (2012). Dirac's prediction of the positron: A case study for the current scientific realism debate. *Perspectives on Science*, *20*(4), 440–475.

Peters, D. (2014). What elements of successful scientific theories are the correct targets for selective scientific realism? *Philosophy of Science*, *81*, 377–397.

Psillos, S. (1999). *Scientific realism: How Science tracks truth*. London: Routledge.

Psillos, S. (2009). *Knowing the structure of nature*. New York: Palgrave Macmillan.

Putnam, H. (1975). What is mathematical truth? In H. Putnam (Ed.), *Mathematics, matter and method* (pp. 60–78). Cambridge: CUP.

Putnam, H. (1978). *Meaning and the moral sciences*. London: Routledge & Kegan Paul.

Popper, K. R. (1983). *Realism and the aim of science*. London: Routledge.

Rowbottom, D. P. (2013). Popper's measure of corroboration and $P(h|b)$. *British Journal for the Philosophy of Science*, *64*, 739–745.

Ruhmkorff, S. (2013). Global and local pessimistic meta-inductions. *International Studies in the Philosophy of Science*, *27*(4), 409–428.

Saatsi, J. (2005). Reconsidering the Fresnel–Maxwell case study. *Studies in History and Philosophy of Science*, *36*, 509–538.

Saatsi, J. (2015) Historical inductions, old and new. *Synthese*. doi:10.1007/s11229-015-0855-5.

Saatsi, J., & Vickers, P. (2011). Miraculous Success? Inconsistency and untruth in Kirchhoff's diffraction theory. *British Journal for the Philosophy of Science*, *62*(1), 29–46.

Sagan, C. (1979). *Broca's Brain: Reflections on the romance of science*. Manhattan: Random House.

Schurz, G. (2011). Structural correspondence, indirect reference, and partial truth: Phlogiston theory and Newtonian mechanics. *Synthese*, *180*(2), 103–120.

Stanford, P. K. (2003). No Refuge for Realism: Selective Confirmation and the History of Science. *Philosophy of Science*, *70*(5), 913–925. **PSA Proceedings**.

Stanford, P. K. (2006). *Exceeding our grasp*. Oxford: OUP.

Stanford, P. K. (2009). Author's response, in grasping at realist straws, a review symposium of Stanford (2006). *Metascience*, *18*, 379–390.

Turner, D. (2004). Local underdetermination in historical science. *Philosophy of Science*, *72*, 209–230.

Van Fraassen, B. C. (1980). *The scientific image*. Oxford: Clarendon Press.

Velikovsky, I. (1950). *Worlds in collision*. Garden City, NY: Doubleday & Co.

Vickers, P. (2012). Historical magic in old quantum theory? *European Journal for Philosophy of Science*, *2*(1), 1–19.

Vickers, P. (2013a). A confrontation of convergent realism. *Philosophy of Science*, *80*(2), 189–211.

Vickers, P. (2013b). *Understanding inconsistent science*. Oxford: OUP.

Votsis, I. (2005). The upward path to structural realism. *Philosophy of Science*, *72*(5), 1361–1372.

Votsis, I. (2011a). The prospective stance in realism. *Philosophy of Science*, *78*(5), 1223–1234.

Votsis, I. (2011b). How not to be a realist or why we ought to make it safe for closet structural realists to come out. In D. Rickles (Ed.), *Structure, objects and causality (Western Ontario Series in Philosophy of Science), chap. 3*. Dordrecht: Kluwer.

Votsis, I. (2011c). Structural realism: Continuity and its limits. In A. Bokulich & P. Bokulich (Eds.), *Scientific Structuralism (Boston Studies in the Philosophy and History of Science)* (pp. 105–117). New York: Springer.

Worrall, J. (1989). Structural realism: The best of both worlds? *Dialectica*, *43*, 99–124.

Worrall, J. (1994). How to remain (reasonably) optimistic: scientific realism and the luminiferous ether. In *PSA: Proceedings of the biennial meeting of the Philosophy of Science Association*, Vol. 1994, Volume One: Contributed Papers (1994), pp. 334–342.