Synthese
DOI 10.1007/s11229-009-9463-6

Beyond Structural Realism: pluralist criteria for theory evaluation

Mark Newman

Received: 13 September 2007 / Accepted: 28 January 2009

Abstract In this paper I argue that singularist approaches to solving the Pessimistic Induction, such as Structural Realism, are unacceptable, but that when a pluralist account of methodological principles is adopted this anti-realist argument can be dissolved. The proposed view is a contextual methodological pluralism in the tradition of Normative Naturalism, and is justified by appeal to meta-methodological principles that are themselves justified via an externalist epistemology. Not only does this view provide an answer to the Pessimistic Induction, it can also accommodate our strongest intuitions regarding the progress of science.

Keywords Structural Realism · Pluralism · Theory choice · Pessimistic Induction · Preservativism · Epistemic desiderata · Methodology

1 Introduction

There is no doubt that fundamental scientific theories have undergone significant transformations through the history of science. This raises the question of whether such changes are due to revolutionary shifts in our knowledge base (and hence should undermine our faith in current science), or if they are revolutionary only in appearance such that we can be optimistic about our claims to theoretical knowledge. If our best scientific theories can be interpreted in such a way that the changes preserve content with at least some minimal continuity, then we may be able to retain the optimistic position that our best theories, when appropriately interpreted, are at least approximately true (Boyd 1981, 1984; Worrall 1989; Kitcher 1993; Psillos 1999; Ladyman 1998). If, on the other hand, we take the revolutionary transformations to be

M. Newman (✉)
Department of Philosophy, University of Minnesota, Duluth, MN 55812-3027, USA
e-mail: mnewman1@d.umn.edu

Published online: 12 February 2009

genuine, then there seems no option but to accept the pessimistic conclusion that further revolutionary change may lie ahead, and that current science could be far from even approximately true (Kuhn 1962; Laudan 1984).

A revolutionary interpretation of science gives rise to the Pessimistic Induction: past theories have been shown to be false, and there is no reason to suppose that current theories are in any more of an epistemically privileged position than their ancestors, so we ought to remain skeptical about current theories being true. This anti-realist argument was refined by Laudan (1981a), and with it he shifted our focus to the logical link between notions of ‘success’ and ‘truth’: past theories were successful yet false, so why think current successful theories entail theoretical truths?

Two responses to this challenge have been proposed. The first is to accept the Pessimistic Induction and reject scientific realism. The second response is to retain one form or another of scientific realism, and attempt to show that despite the historical record some of the theoretical content of science remains unaffected. Scientific realism comes in many forms but each must satisfy the requirement of selecting some content from past scientific theories that can explain their success and show that this content has been retained in current theories (or failing that, show that the methods of science have improved such that current theories are more reliable than past theories). If such continuity can be established, then maybe we really do have good reasons to think the relevant cognitive content of current successful science is true, and realism is secured while the progress of science lies in the continuing accumulation of just such cognitive content. This can be considered a ‘preservativist’ approach to saving science from the Pessimistic Induction.

Various forms of preservativist scientific realism are currently popular, but my focus in this paper will be with a specific variety that I shall call ‘Singularly Principled’. These versions are conspicuous in that they appeal to a *single* principled means by which we ought to interpret our best scientific theories so that we can establish the required historical continuity. The singularly principled scientific realist suggests that by appealing to a realistic interpretation of only one *kind* of theoretical notion in our theories we can rescue science from the Pessimistic Induction. Examples of this approach include Structural Realism and Entity Realism. For the Structural Realist this preferred content is the structure of the theory in question, and for the Entity Realist it is the existence of its theoretical entities. When it comes to problems regarding singularly principled theories, I assume that the difficulties for Entity Realism are similar to those for Structural Realism, and restrict my discussion to the latter.

In this paper I argue that singularist approaches to solving the Pessimistic Induction are unacceptable and that when a pluralist account of *methodological principles* is adopted we see the Pessimistic Induction is nothing to fear in the first place—it simply falls apart. In Sect. 2 I show why it is problematic to use just a single epistemic principle for saving scientific realism. In Sect. 3 I argue that there is an alternative, and much better, approach to answering the Pessimistic Induction; a contextual and pluralistic theory of epistemic desiderata. This view takes the methodology of science to be dynamic, and by studying the history of the methods and epistemic principles used in science we can show how the primary cases used to support the Pessimistic Induction are not really convincing—they never were particularly successful theories. This entails that the Pessimistic Induction is only an *apparent* and not a *genuine* problem for scientific realism. My theory of contextual methodological pluralism is in the tradition of Normative Naturalism, and in Sect. 4 I address some issues that this approach to scientific method raises. In particular, I argue that to defend such a contextual Normative Naturalism from falling into relativism one must appeal to meta-methodological principles that are themselves justified via an externalist epistemology. I will then attempt to show how one can be a contextual pluralist about scientific method and yet still accommodate our intuitions regarding the progress of science when realistically interpreted.

2 Structural Realism's singular epistemic principle

Structural Realism suggests we treat realistically only the structural content of theoretical claims made in our best, most mature scientific theories. Only the structural content is cognitively significant and epistemically justified. It is this structural content that is genuinely responsible for the success of our theories, and hence confirmation for theoretical content supports only the reality of theoretical structure, and not the reality of further claims regarding the underlying nature of the world (those regarding non-structural components). The Structural Realist argues that when we look to the history of science, although we see radical shifts in theoretical content from one theory to the next, there is nevertheless retention of theoretical structure, and hence we have no reason to be skeptical regarding the Pessimistic Induction—there is no rejection of past structure, so the Pessimistic Induction no longer has a basis.

Worrall's (1989) advanced a straightforward definition of structure as mathematical content, and famously appealed to the transition between Fresnel's ether theory of light propagation and Maxwell's vector field theory to illustrate that retained structure can be found in the Sine and Tangent laws of optical refraction and reflection. Others have followed Worrall's lead and there are now two general categories of Structural Realism; Epistemic and Ontic. The former comprise those theories that provide a definition of structure which is to be treated as an epistemic constraint—structure reflects the knowable genuine relational structure of the unobservable world. Ontic Structural Realism on the other hand suggests a definition of structure which is metaphysical in nature—the structure of our theories reflects the structure of the world, and there is nothing more to know; all is structural.[1] Given this reading, the Epistemic Structural Realist accepts that there is a great deal of reality about which we will remain forever ignorant. The Ontic Structural Realist takes structural knowledge to be limited only by our resources for discovery—if all is structure, there are no principled limits to our knowledge of the unobservable.

There are many interesting questions to be investigated regarding each of these positions (especially regarding how they are to define their respective notions of structure so as to avoid the Pessimistic Induction), but space limitations require we move on to consider a property common to both versions, against which I wish to argue. Put succinctly, I take it that both forms of Structural Realism are committed to the idea

[1] For the initial formulation and subsequent adjustments to ontic structural realism see Ladyman (1998), French and Ladyman (2003), French and Saatsi (2006), and Saatsi (2005).

that there is in principle a single definition of structure with which one can interpret the theoretical content of successful science, and with which one can show the Pessimistic Induction fails. If this is a sound assumption, then what the Structural Realist proposes is that we can look to a singular epistemic principle to divine both a defense of scientific realism and account for the intuitively obvious progress of science. This is however a poor policy, for as I will now argue, there is good reason to think that a realist interpretation of the history of science using a singular epistemic principle is myopic, and ineffective.

We will start with the hard cases for the scientific realist, those that form the basis for the Pessimistic Induction: Phlogiston, Caloric, and the Luminiferous Ether. I want to suggest that all three of these theories can be interpreted under a set of epistemic principles which are historically more accurate than those taken to hold under the current Structural Realist account, and by doing so show that the Pessimistic Induction is not a problem for scientific realism at all. Importantly we should do without singular epistemic principles for interpreting the causes of success in science. We will do better to adopt a dynamic set of epistemic desiderata with which to explain the success of science. In this way we not only avoid the unnecessarily restrictive constraint of a single principle dictated by a select definition of structure, we also accommodate more accurately the actual history of science and overcome the epistemologist's obsession with a definition of *justification*. Justification is replaced with a pluralist set of epistemic desiderata that reflect contextual weightings and which are justified within a naturalistic study of scientific methodology. I begin with general arguments against singular epistemic principles as a means for answering the Pessimistic Induction. In subsequent sections I will fill-out the details of the position sketched above.

2.1 Two concerns for Structural Realism

We need not belabor the myriad of specific criticisms hurled at Structural Realism[2] to date, but can rather focus on a couple of more general concerns. First, by its very methodology the Structural Realist's approach leaves the realist in a position only to argue for realism in a *post hoc* manner, illustrating which parts of the structure of our best theories we ought to believe-in only after history has long passed them by. If one wishes to know what we ought to believe about current and developing successful scientific theories, then you are out of luck. Scientific realism on this reading is entirely backward-looking, and passes judgment on theoretical content only to the degree that it can be shown to be retained as structure in subsequent successful theories. This leaves the realist in the uncomfortable position of failing to provide a solution to the Pessimistic Induction which tells us exactly what we are justified in believing about today's theories, as well as being subject to the additional accusation that he is reconstructing the historical cases to fit his specific epistemic principle.

Several versions of scientific realism (both Structural and non-structural) have attempted to avoid this post hoc difficulty by appealing to the *essential components* of

[2] For recent work criticizing various forms of Structural Realism see Psillos (1999), Cei (2006), and Chang (2003).

a theory—those components that are *indispensable* for a theory in making successful predictions (or otherwise gaining confirmatory support). Two examples are found in recent work by Psillos (1999) and Chakravartty (1998, 2003, 2004). We see Psillos arguing for the reality of all indispensable theoretical properties, relations, and processes that are responsible for the success of a theory. This clearly goes beyond the spirit of Structural Realism, and indeed Psillos rejects that position. Still, this is very close to Chakravartty's Structural Realism, which appeals to the reality of the minimal ontological interpretation of indispensable components responsible for the epistemic success of our theories. Chakravartty's difference with Psillos primarily resides in his definition of structure as being dispositional. Both approaches are plausible, but only, I suggest, to the degree that they diverge from a commitment to a singular epistemic principle for determining what is and what is not to be considered real. Psillos commits to a variety of epistemic principles (novel predictive success, causal mechanisms, reliable processes of belief formation, etc.), and Chakravartty considers himself a Structural Realist, but has a very broad notion of structure—so much so that he defines it in terms of the relations between first-order causal properties of theoretical entities:

> First order properties whose relations comprise these structures are what we might call *causal properties*: those that confer dispositions for relations and thus dispositions for behaviours on the objects that have them.[3]

If our epistemic principles are dictating that we should only believe in structure, yet this structure is broad enough to capture all first-order causal properties, it seems that we are committed to a great deal more than the austere notion of structure which was the motivation for Structural Realism in the first place. Still, Chakravartty does explain that we ought only commit to a minimal amount of structure—just that which is absolutely necessary for deriving predictions. He believes this is usually just a reflection of our best account of the causal structure of the world according to current detection devices. Here however, there is another concern. In some cases it is simply impossible to isolate-out exactly and only those theoretical structures that are responsible for empirical success—Leplin (1997, p. 182) argues for instance that isolation of particular entities is sometimes prevented in the theories that posit those entities, such as with the quark, the graviton, and magnetic monopoles. If we cannot isolate these entities, and as a consequence cannot show them to be detected, then how are we to seriously consider these entities, or even just their causal properties, detected?

A second problem for the general strategy adopted by Structural Realists, and the one to which this section is primarily devoted, is that singular epistemic principles that attempt to delineate when we are and are not justified in holding a belief about the unobservable are unnecessarily restrictive. They appeal to formal explanatory properties which are supposed to be truth-conducive in some way, and which are supposed to enable us to infer from the presence of specific components in retained parts of theories to the conclusion that we are justified in believing in these components. I want to draw attention to the fact that there are a diverse set of desiderata which can all be used from

[3] Chakravartty (2004, p. 155).

time to time, depending on context and domain of application, to help guide us towards the truth, or at least likely towards the truth. We really do often think, for example, that unifying theories or providing accounts of causal mechanisms are truth tracking virtues. These approaches are generally left-out by Structural Realist accounts, and those that do try to accommodate them tend to stretch their notion of structure beyond what might reasonably be thought to be either formal or capable of avoiding the Pessimistic Induction. By diversifying our set of methodological principles my approach can appeal to the history of science to support contextual explanatory principles of epistemic value—and hence is both broad enough to capture all kinds of explanatory principles, and is also entirely naturalistic. The specification of which principles work and in which contexts is complicated of course, and I will review in the next section the most familiar. However, here it remains to be argued that appealing to a plurality of epistemic principles is preferable to the singularist approach adopted by Structural Realists.

2.2 Against singular epistemic principles

It is important to emphasize that structural realists may happily concede that there are a plurality of epistemic principles adopted by science in the process of theory confirmation, such as 'believe theories that *unify* a great deal of disparate phenomena', 'believe theories that make successful *novel* predictions', etc. The problem is, realists accept that these principles are all subject to the Pessimistic Induction—the history of science shows these principles operative in what turned out to be false theories. Newton's erroneous gravitational theory *unified* both celestial and terrestrial mechanics, while Fresnel's prediction of the white spot provides a clear example of a false theory that made a *novel* prediction. What I wish to urge is that the Pessimistic Induction provides only *apparent* counterexamples to these principles. We need not give up on the selection of such principles as epistemically probative if we acknowledge that *in problematic cases other principles were more valuable*. A naturalistically informed and dynamic account of epistemic principles which can reflect how in different domains, at different times, different principles track truth with different success rates, is far preferable to a singularist approach which narrows our resources unjustifiably.

I will begin here by first listing some of the principles one might specify as indicating when we are and are not justified in believing in the unobservable entities, mechanisms, laws, etc. of our best theories. These make up a diverse set of principles, all appealed to as constituting in some way part of a best explanation defense of scientific realism. They are merely examples of principles we see through the history of science being used as methodological rules for theory choice, and are not meant to be a definitive, exhaustive set of principles for scientific methodology. Neither do I claim the concepts being used in these principles are entirely precise or lacking ambiguity—even the notion of a best explanation itself is left unaddressed. These principles are however, commonly appealed to by methodologists of science, and they are supposed to each be weighted according to inductions that scientific methodologists can make regarding their relative value in a science, at a particular time. This will become clearer as we progress.

What constitutes a best explanation capable of surviving the Pessimistic Induction *might* be one that:

1. reveals *causal mechanisms* behind the phenomena,
2. maximally *unifies* the phenomenon to be explained with other phenomena,
3. is derived from a *reliable* method of investigating the unobservable,
4. uses a maximally *coherent* set of propositions that entail the phenomena,
5. is the *simplest* explanation,
6. is the most *fruitful* in terms of explaining new phenomena,
7. is the most *understandable*, *intelligible* explanation of the phenomena,
8. reveals *novel* predictions of new phenomena,
9. retains *correspondence* relations between prior and subsequent theories,
10. *explains past successes and failures* of predecessor theories as well as current phenomena, and
11. reveals the mathematical *structure* of the world.

The list could be extended I believe, perhaps to include requirements that the theory not be developed with ad hoc modifications, that it be falsifiable, that it use double-blind testing where possible, that it be internally consistent, etc. I take it that *prima facie*, each of these principles holds *some* epistemic virtue to be found in cases from the history of science. For instance, biological *causal* mechanisms used in genetics surely have plausibility from the historical record as contributing positively to our stock of true beliefs about the phenomenon of inheritance. The *unificatory* Newtonian synthesis of Galilean mechanics and Kepler's celestial laws of motion surely contributed to our knowledge of the laws of universal motion. The *reliability* of optical theories of the telescope to bestow on 17th century scientists true beliefs about the orbits of objects around Jupiter and the existence of mountains on the moon, again surely cannot be construed as failing to have epistemic worth. Likewise for the remainder of these candidates: the *coherence* of Hess' theory of seafloor spreading contributed to its overtaking Wegener's account of continental drift; the *simplicity* of adopting a non-Euclidean geometry over adjustments to physical theory made a positive epistemic contribution to the acceptance of the curvature of spacetime; the *fruitfulness* of Bohr's theory of the Hydrogen atom helped (through Sommerfeld's positing of elliptical orbits), to move us towards a more accurate theory of atomic structure; the *understandability* and *intelligibility* of the kinetic theory of gases undoubtedly brought us closer to our current knowledge of statistical mechanics; the *novel* predictive success of the curvature of light around a massive object contributed to our knowledge of the nature of space and time; the correspondence relations that hold between classical and special relativistic mechanics indicate a move in the direction of a more correct theory; the *explanation* given by general relativity *of the successes and failures* of Newtonian gravitational theory (ability to predict tides, failure to account for perihelion of Mercury) seems epistemically valuable; and finally, the *preservation of mathematical structure* between Fresnel's optical theory and Maxwell's electrodynamics is a clear indication of retained laws which are with us indefinitely.

Now of course these examples *assume* more recent theories to be advances towards the truth over earlier theories, and this itself is to beg the question against the anti-realist. However, the discussion I am trying to develop is one already within the realist camp, and no amount of realist wrangling here is likely to convert the empiricist. What I am trying to do is elucidate the diversity of epistemically probative principles still available to the realist, and I shall later try to illustrate how they can be justifiably adopted in a pluralist epistemology. For now, to return to the main thread of my argument, it is necessary to explain why our failures due to adopting only a single principle, license the adoption of many.

Initially it might be tempting simply to argue that since no one principle has been successful there is little hope of finding a single epistemically principled account by which to save realism. That is, one could make a case that past failures of realist theses support a straight induction to the failure of all such attempts. This would be far too quick though. Simply because a long train of very smart philosophers has failed so far to find a solution to this problem, it would be unfair to conclude such a solution unachievable—we still debate theories of explanation, laws of nature, etc. thinking there is hope for a solution to these issues. No need for pessimism on these grounds I think.

However there are other, positive, reasons for abandoning the single epistemically principled approach to saving scientific realism. The most obvious amongst these is the fact that (as illustrated in the examples above) there really are cases of divergent epistemic principles lending themselves positively to the advancement of our stock of truths about the world. The only step we really have to take here is to notice that our attempts to justify beliefs about unobservable entities reflect a divergent set of independent principles of epistemic justification that enable us to go beyond the empiricist's notion of what counts as evidential.

Yet this rather obvious tack of looking to successful methodologies to undermine Structural Realism and secure a broader notion of scientific realism is surely overwhelmed by the historical record. The whole point, after all, of the preservativist strategy is to isolate specific components of theories that are epistemically privileged by some principle which avoids committing us to entities, processes, laws, etc. that have subsequently been rejected from the corpus of scientific knowledge. How on earth are we justifiably to adopt a naïve pluralism about epistemic principles when the whole point of realist defenses has been to whittle-down commitments based on such principles so as to avoid the Pessimistic Induction?

The answer to this most pressing problem lies, I believe, in a naturalistic and historical account of the epistemology of science. In the next section I will illustrate why, by appreciating the contextual factors in scientific research such as specific disciplinary methodology and background theories, we have good reason to think that the reliable, contextually appropriate, and accepted principles operative during the periods when Phlogiston, Caloric, and the Luminiferous Ether were successful, were all abused. The positing of these entities was a mistake because at those times, in those contexts, such posits should have been rejected—if one follows the dynamic adoption of different principles for truth-tracking via a reliabilist epistemology.

3 A pluralist account of epistemic desiderata for science

So far I have argued that singular epistemic principles, such as those of the Structural Realists, are not able to accommodate the Pessimistic Induction because they are incapable of either providing forward-looking prescriptions for our beliefs regarding theoretical structure in science, or accommodating successful uses of methodological principles in the history of science. But how can my pluralistic account fare any better? The answer lies in the identification of *differential weightings* for different epistemic principles along a dynamic dimension. As context changes, the history of science indicates that the reliability for specific principles to provide epistemically probative information varies.[4] What we require is a careful analysis of this process as it occurs. This is the key to answering the Pessimistic Induction.

Let me make this a broadly two step argument. In the first step I list some of the epistemic principles that *seem* to be what I shall call 'primary' for a scientific discipline in a particular context, as reflected by the history of science. I shall also list some principles that appear to be 'secondary' in the sense that they may have been used by a community at a particular time, but were not at that time really justified as rules which secured reliable beliefs. Just as with the primary principles, secondary principles are deemed so by the scientific community at hand, implicitly or explicitly, and may change epistemic status as science progresses. A secondary principle may turn out in time to show its worth as a reliable means of acquiring beliefs about the world and in doing so may be justifiably promoted to the rank of a primary principle—but remember that this is relative to a context of investigation and does not apply to all scientific investigations across the board. Similarly a primary principle, thought to be a reliable rule for investigating the world may turn out as time goes by to be unreliable for that particular context of investigation, and may find itself demoted to the status of secondary for further investigations. This might happen when a rule which works adequately for a while becomes unreliable as the domain of inquiry broadens to include new phenomena.

A principle deemed 'primary' then, acquires this status from the scientific community using it, and as such the division between primary and secondary theoretical principles is contingent upon the success these principles achieve in producing reliable consequences. I shall suggest that those principles in the primary camp for a science at some particular juncture in the history of science show themselves to be reliably capable of generating true beliefs by themselves, and those in the secondary camp lead to true beliefs perhaps as a consequence of implicitly depending upon primary principles.

In the second step of the argument I illustrate how the cases of Phlogiston, Caloric, and the Luminiferous Ether *seemed* to provide successful theories that required belief in non-existent entities, but in fact *given the primary epistemic principles operative in those domains, at the time in question, there was little reason to think these entities*

[4] Truth-probative principles are hard to specify without assuming, as above, that we are justified in thinking today's theories are at least approximately true. We have to make this assumption if we are to show that some of our principles are more truth-conducive than others in different contexts, but there should be no worry in this claim since the very plausibility of the Pessimistic Induction itself hangs on the conflict between the truth of past and that of current theories. The conclusion of the Pessimistic Induction, which asserts the falsity of the realist's assumption, doesn't defeat the need to make the optimistic assumption for the purposes of argument. Similarly, for the purposes of my argument, we can here assume the approximate truth of current mature science and point to divergent epistemic principles as indicating that we sometimes get at the truth in very different ways, not by merely Singular Epistemically Principled strategies.

actually existed. The principles used to support the existence of Phlogiston, Caloric, and the Luminiferous ether were not 'primary' at the time in question—they were at best deceptively so. These principles may reasonably have been adopted as heuristic devices for theory development, or for instrumentally deriving predictions, but it was not reasonable to believe in their reliability at revealing theoretical truths. It may have been necessary in theorizing to use some particular entities, or properties of entities (such as the repulsive force of Caloric particles), but given the reliable methodology of the time, these were assumptions, not established facts. This latter argument requires I show that although there may have been principles indicating belief in the non-existent entities, these were trumped by alternative principles appropriate to the science of the time. The epistemic principles of scientific justification ought to be appropriately contextualized and weighted, and when they are, the 'hard cases' of the Pessimistic Induction actually evaporate. Explaining away the hard cases in this reliabilist manner—by appealing to the illegitimate use scientists made of secondary principles as primary—alleviates the realist's need to adopt a singular epistemic principle for science. We can be '*contextual realists*' using a pluralist epistemic desiderata by showing that at some moments in the history of science it is appropriate in a particular domain to appeal to some principles, but it is inappropriate to use others. It is in this way—by looking at the success rates of specific rules of investigation at a particular time in a particular context in science—that we can generate a ranking and weighting scheme for methodological principles which informs a realist interpretation of theories based on something other than a static singular epistemic principle, such as that suggested by Structural Realists.

3.1 Primary and secondary methodological principles

In order to illustrate the distinction between primary and secondary methodological principles we really need a context from science. Without a context it makes no sense to speak of a principle or rule of investigation as being primary since this status can only be attributed on the grounds of evidence of reliability, which obviously requires some episode of investigation. Even so, some principles are primary in more than one field of investigation, and I list below just a few very common and highly reliable principles of scientific investigation which are quite general and are used in a wide range of disciplines. Each rule has been adopted based on the high degree of reliability it bestows on conclusions accurately drawn from it, and hence is taken to confer positive epistemic status on any theory constructed using it:

1. independently test variables of a theory—control for confounding variables,
2. use double-blind test procedures when possible,
3. use highly reliable measurement instruments and procedures, and
4. use randomly sampled, large populations for experimental testing.

Now these are clearly principles of empirical support and comparative theory assessment which most realists would happily endorse as truth-conducive throughout the history of science. To justify our confidence in these rules we can point to the specific reasons we think they are aids to deriving truths about the world. The first principle is of course designed to rule-out the successful results of an experiment being identified as due to an incorrect causal factor—to prevent us from believing for example that yellow stained fingers cause cancer (obviously smoking causes both the staining and the cancer). The second principle is well known to be an effective means of preventing both placebo effects and experimenter bias, while the third suggests that we take adequate measures to rule out instrument bias. Both of these principles then attempt to rule-out forms of bias of one kind or another—a factor especially important in medical and biological trials. Similarly, bias is addressed by the fourth principle, although here it is a selection bias and has proven quite difficult to combat in many instances—consequently the use of statistical techniques to inform appropriate confidence intervals for inductions, as well as the use of stratified and systematic sampling, have been useful developments in achieving the aim of this primary rule.

Of course these four rules are just a tiny selection of general primary principles of modern scientific investigation, and many, many more could be listed without even digging into particular domains of inquiry. Still, these apparently obvious research tools weren't always in the repertoire of early scientists (and consequently played no role in early theory confirmation). In particular the confirmation of the existence of unobservable theoretical entities was derived from less selective procedures. This means we'll see such reliable rules missing from many episodes in the history of science, and in those cases other rules of investigation will have been followed.[5]

None of this is news to the philosopher of science, who has studied the historicist theories of Kuhn, Lakatos, Laudan, etc. Their theories of science attempt to provide us with a story which accurately reflects the divergence of such methodologies throughout the history of science. But putting the pessimistic conclusions of these historicists aside, if we are assuming (as we must) that current science is approximately true, then what is it about these principles that leads us to the truth? For the first four principles this is fairly obvious (as I have indicated), and we can easily unpack an explanation for why they work to prevent error. What is perhaps more illuminating is why more controversial alternative principles often adopted by scientists and philosophers should receive the epistemic weight they do. In many instances of theory selection it is less than clear why the principles used for theory choice were thought to lead to the truth. Examples of such principles might include:

5. seek novel predictions,
6. seek explanations for a broad range of phenomena,
7. avoid ad hoc modifications to theory in light of new evidence,
8. seek multiple routes to deriving the same phenomenon,
9. seek models for underlying mechanisms, and
10. preserve locality of causal interactions.

These are principles at a slightly higher, more general level than (1–4), and represent not properties of investigative procedures, but properties of theories themselves. It is unclear even today why such rules are thought to lead to theoretical truths, and

5 Even with agreement on a set of rules of confirmation such as those listed above, it is unlikely that we can expect equal weighting to be transmitted to each rule by every member of a scientific community. Some scientists may prefer (1) over the others, so much as to weigh it heavier than the sum of all remaining principles. Other scientists will be willing to distribute equal epistemic worth to the satisfaction of each principle.

philosophers continue to debate what underlying causal factors might be responsible for the reliable success they do sometimes have. It might be possible that the achievement of these more general principles can be found in the use of more particular, specific principles. That is, perhaps one can argue that the secondary principles used for theory selection are quite often just the result of more reliable primary principles. For example, to achieve a very broad-reaching theory, one might find it necessary to adopt (4)—for instance in developing a theory that explains the cancerous as well as dietary effects of some drug on a species we may well require the experimentally controlled testing of large groups of randomly selected rats or mice. This is however a rather tenuous conjecture, and requires a great deal more investigation than we have room for here.

But even if there is a clearly explainable link between primary and secondary principles, how does this help the main thread of my argument, which is to show that some of these principles are more clearly truth-conducive than others? Well, one simple way to proceed is to point to the development of rigorous experimental procedures during the last hundred years, and the correlated fact that when used the scientific theories they play a part in confirming are still taken to be true. We could then compare these primary methodological principles with the secondary principles, which may correlate with historically successful theories in some cases, but fail in others.

For example, our discovery of the Placebo effect has played a prominent role in *ruling-out possible confounders in studies*, so that nowadays an experimental design that fails to accommodate this effect or account for it with at least single-blind (and ideally double-blind), methods is considered suspicious. Similarly, we discovered that to test for the existence and properties of theoretical entities, such as the structure of DNA molecules, we needed to independently test our measuring instruments (X-ray devices) against independently developed theory (X-ray diffraction). Without such independent support the images produced by early theorists, which were suggestive of an “X” structure, couldn’t plausibly be claimed to indicate a three dimensional double-helical structure.

On the other hand, secondary epistemic principles, such as *appeal to novel predictive success*, may appear on the surface to indicate we have discovered the true nature of some unobservable entity responsible for a phenomenon, but nevertheless be misleading all the same. For example, although the positron was a novel predictive success—Dirac posited the particle as responsible for excess energy expulsion in high energy particle collisions—such novelty is also reminiscent of Poisson’s white spot, which was famously a surprising prediction made as a consequence of Fresnel’s wave theory of light. We now think the wave theory defunct—although white spots still appear, we now know that it is for a quite different reason than that given by Fresnel. So sometimes novelty seems to guide us to the truth, and sometimes it does not.

Another example: The *ability to mechanically model the underlying mechanism of a theory* was famously Lord Kelvin’s criteria for belief in a theory. This principle seems to have worked quite nicely for explaining the diffusion of gas through a room along the lines of kinetic theory. Still, the mechanical models of the Luminiferous Ether developed by the likes of McCullough, Green, and Stokes, all failed to pick-out an entity that would persist into the twentieth century. As such, the ether, as a mechanism was abandoned because it was seen to be otiose in light of the development of field theory. Mechanical modeling therefore also seems to be a fickle friend.

I suggest that we conclude from such examples that not only do methodological principles come and go, but that as epistemic desiderata some are decidedly more reliable in some contexts than in others. To accommodate this fact we ought to look to the history of science and discriminate the appropriate weightings of primary and secondary principles on the grounds of their success up to the current period.

Additionally, there are likely to be examples in different domains where different principles work with different effect. For example, functional or teleological explanations for a phenomenon, such as those given for the evolution of the human eye, will find little help from the primary methodology found in optical experiments on prisms. The two frameworks are so entirely divergent in their approaches to their respective fields of study that it would be absurd to expect anything less than very different methodologies. I don't think this undermines my position, since a domain of methodological principles is really established via a posteriori interaction with the world, so there is no a priori reason to think that what goes for one domain ought to go for another (which is another perfectly compelling reason to reject the idea of singular preservativist epistemic strategies).

If the argument above establishes that epistemic principles can vary over context, and ought to be given different weightings in different domains, then how can we use this idea to answer the Pessimistic Induction? The trick here is to show that the primary epistemic standards of the day did *not* really license belief in those 'hard cases' delivered by the anti-realist—anything less than this (for instance, claiming realism can be satisfied alone by contextually relative methodological principles), will fail to show that these historical cases should not be treated as counterexamples to the realist thesis.

3.2 Dissolving the Pessimistic Induction—phlogiston

So now it is time to move to the second part of my argument. In this second step I illustrate how the cases of Phlogiston, Caloric, and the Luminiferous Ether *seemed* to provide successful theories that required belief in non-existent entities, but in fact *given the reliability of the epistemic principles operative in those domains, at the time in question, there was no reliable reason to think these entities actually existed.* Additionally, I want to argue that there is a single principle which played an importantly *misleading* role in all of these theories, inclining natural philosophers to infer the existence of non-existent entities. That principle I will call 'Intelligibility' and it can be characterized by the constraint: 'Believe only in those theoretical entities that make the theory most *understandable* when providing explanations for specific phenomena'. The problem here is that although intelligibility has been retained even today as an epistemically valuable desideratum, its intuitive use nowadays is far less epistemically compelling than in previous centuries. The common error that led many theorists of the past to adopt entities like Phlogiston, Caloric, or the Luminiferous Ether is that they granted excessive epistemic weight to the notion of intelligibility—elevating it to the level of what I am calling a 'primary' principle—even though the successful sciences prior to these episodes sometimes recommended a rejection of epistemic principles that praised the status of intelligibility. In fact I would argue that the reason some think we can use intelligibility even now as a reliable guide to truth is precisely because the history of science shows that when nature forces what appear to be unintelligible interpretations of the unobservable upon us, *continued use* of these interpretations, and hence *familiarity* with them, breeds a *sense* of intelligibility that only success can induce. This is why some might suggest that quantum mechanics really is intelligible despite its counterintuitive interpretations. I want to suggest however that an initially intelligible account of the underlying phenomena, even in the 18th century, was an unreliable method of inference to unobservables. If this can be shown, then these cases should not be taken as counterexamples to the realist thesis at all, because it was a clear mistake for theorists at the time, *by their own lights*, to adopt Phlogiston, Caloric, or the Ether.[6]

Starting with Phlogiston, is there reason to think that this research program was developed on faulty methodology, even by the lights of scientists of the day? I think so, on the grounds that chemists of the day recognized the importance of *quantitative* methods of analysis for substances, and yet Phlogistonians overlooked the glaring disparity between predictions their theory made and the empirical facts regarding the weight change of metals upon calcination. These Phlogistonians were focusing especially on the intelligibility of the theory (a secondary principle) rather than on its quantitative rigor (a primary principle), even though they had good evidence that quantitative rigor was a more epistemically respectable principle—in fact some of these Phlogistonians are traditionally considered forerunners in the shift to quantitative methods in chemistry.

The Phlogiston theory[7] is generally taken as developing from Georg Stahl's work in the 1720–1730s. The theory was capable of providing explanations for all manner of phenomena, but most notably those of combustion—that in the process of burning, a substance is releasing another substance (Phlogiston) into the air, and that when a flame dies out in a container, it is due to the air becoming saturated with Phlogiston. Besides combustion, Phlogiston also provides an explanation for the process of smelting whereby ore is refined—heating ore with charcoal releases Phlogiston from the charcoal, and it enters the ore to generate the respective metal. If one continues the heating process the metal returns to an ore state since Phlogiston is now heated right out of the metal. Phlogiston can also account for calcination and respiration: the surface of a metal changes color due to the release of Phlogiston (rusting), and when we breathe we absorb and release Phlogiston—again, when the air is Phlogiston-saturated we are unable to breathe any longer (suffocation). The theory was therefore considered to be a quite general and unifying account of several very important, and up until then problematic, chemical phenomena.

[6] It is important to emphasize that the realists who posited these false entities were not following reliable epistemological principles, otherwise the reliability of current epistemological principles will fall to a Pessimistic Induction on reliable principles itself. What follows is a sketch of the historical situation. For more detailed accounts references are provided.

[7] For these details on phlogiston I have used Giere (1997), Perrin (1990), Thagard (1992), Hudson (1992), Partington (1965), Brown (1920), and Mason (1962).

But Stahl's theorizing about Phlogiston was based very much on his appeal to one amongst several possible general chemical frameworks—that of 'essential chemical principles' (Perrin 1990). This rubric suggested that substances are composed of various element-like components, some defined as substances, others more functionally. Simple principles divide into three categories: vitrifiable principles, liquefiable principles, and inflammable principles. Properties of substances, and especially compounds can be explained in this framework by appealing to the properties of their constituent principles. Phlogiston, on Stahl's theory, was the inflammable principle. Yet other rubrics were available under which Stahl could have investigated these phenomena. First there was the traditional Aristotelian approach which appealed to four elements (earth, air, fire, water). Second, there was the framework of chemical affinities—a view propounded by Etienn-Francois Geoffory around 1718, whereby substances are categorized according to the reagents with which they react.

Now although Stahl's Phlogiston theory was advanced two decades earlier, according to Gabriel Venel's encyclopedic article 'Chymie'(1753), by mid-century most French chemists still accepted the view that natural bodies were composed of the Aristotelian four elements.[8] This is an important piece of evidence against the view that Phlogiston was accepted across the entire chemical community since the four elements view conflicts with Stahl's approach—Phlogiston was one of three possible earthy substances, but on the Aristotelian account earth was a simple substance.

More compelling evidence against the general acceptance of Phlogiston can be found if we go further back, to the mid 17th century, where we find something like an early version of Phlogiston theory and see even there significant opposition to it on quantitative grounds by leading men of science.[9] To trace Stahl's theory to its origins, we must start with his teacher Johann Becher (1635–1682). Becher wrote a book in 1669 titled *Physica Subterranea* in which he advocated the view that combustion is the decomposition of a substance into its constituent parts, and that when this burning occurs its cause is the release of one of the earthy principles (*terra pinguis*). When metals calcinate it is due to the release of *terra pinguis*, and when combustibles burn it is due to the expulsion of this earthy substance. It was in Stahl's (1697) *Zymotechnia Fundamentalis sive Fermentationis Theoria Generalis* that he advanced Becher's views, and in particular adopted the term Phlogiston for what had previously been *terra pinguis*. The term Phlogiston was not new at this time, and had at least been used previously by Boyle—which is ironic because we see Boyle arguing *against* just such a principle even in 1673. In his *Experimenta Nova*, Boyle showed that heating tin increases its weight, and concluded that this is due to the metal *absorbing* (not expelling) some form of matter. This is of course in direct conflict with the notion that when calx is formed via calcination it is due to the release of Phlogiston (or *terra pinguis*).

8 According to Perrin (1990, p. 266) this was significantly due to Hale's discovery in the 1720s that air could be extracted from animal, vegetable, or mineral substances and trapped in an apparatus. This empirical discovery overwhelmed the weak predictive success enjoyed by the chemical affinities approach—success mostly derived from analogies between reagents and their associated substances.

9 The account given in the following section draws variously on Brown (1920), Levere (2001), Leicester (1956), Davis (1966), and Partington (1965).

But Boyle was by no means the only natural philosopher to take issue with Phlogiston. Nicolas Lemery argued in his *Cours de Chemie* (1675) that when lead gains weight upon burning it is due to the *absorption* of corpuscles of fire that insinuate themselves into the pores of the metal. Boerhaave argued for a similar view as late as 1732, while Van Helmont had argued in 1640 that burning is not a separation as Stahl and Becher were to propose, but is rather a special glowing condition of volatile bodies. Newton in 1701 similarly suggested burning was merely the emission of light, while perhaps more impressive is the work of Jean Rey who suggested as early as 1630 that the increase in weight upon heating for both lead and tin is due to the condensing of air; air develops an adhesive nature, he thought, and in doing so *attaches* itself to the tiniest particles of metals. As J.C. Brown points out in his *A History of Chemistry*, it is remarkable that Rey could have come so close to the true theory of combustion and have just missed it. This is true also of Robert Hooke, who in his *Micrographia* (1664) provided a theory of combustion tantalizingly close to that developed over a hundred years later by Lavoisier. For Hooke combustion was a process whereby a part of the air, not all of it, was required for a body to burn—and importantly this part of the air was not the same as that part required for respiration.

These early opinions are all in some form or another opposed to the idea that burning is the *release* of a substance from matter, but there are many more leading, influential chemists that from the 17th and into the 18th century also argued against principles operating along the lines of Phlogiston. John Mayow argued in the late 1660s that burning is due to only some subtle part of the air which combines with the combusting substance. Importantly, this process is the *absorption* of particles, and it is they that account for the weight difference before and after burning. This process, he noted, also occurs with metals as they calcinate, and he even gave these particles the name *spiritus nitro-aereus*. This is surely as close a discovery of oxygen as that by Rey, yet significantly both works were overlooked. We can also add to our list of antiphlogistonists Stephen Hales, who in 1727 suggested that the increase in weight after calcination was due to some *addition* of substance to the metal.

Not only was there significant divergence over the fundamental constituents of matter (and hence disagreement over whether Phlogiston could exist as a category of substance at all), nor merely was there rampant disagreement over the mechanics of burning, it is also not even clear how important the phenomena explained by Phlogiston theory were to most chemists of the period—these antiphlogiston opinions were after all overlooked. Why should it be that the views of leading scientists were ignored by the Phlogistonians? Well, in favor of Phlogiston, Venel indicates a growing surge behind the theory due to the *explanatory* power it seems to have enjoyed (explained above), and also since it was capable of *unifying* (as was Geoffray's account) disparate phenomena (charcoal, calcination, metals) through the chemist's ability to *manipulate* Phlogiston in experiments on combustion. That is, Phlogiston provided a means by which chemists could explain (by appeal to analogy with notions like fluids, repulsive forces, and saturation) a number of diverse phenomena in terms of an unobservable fluid-like substance that behaved in many ways like visualizable, more familiar substances and entities. But despite these benefits of the theory, this is quite probably where the notion of *intelligibility* played too strong a role for Phlogistonians.

Requiring an explanation in terms of a discrete particulate fluid substance that flows from matter under combustion may have superficially seemed good methodology, but this preference overlooked not only the difficulties of mechanically accounting for how particles of Phlogiston repel out of a substance, but more importantly, it overlooked an established epistemic desiderata: quantitative measurement.

To establish my intended conclusion, that the Pessimistic Induction shouldn't appeal to Phlogiston as one of its 'hard cases', we need to establish not only that there was disagreement in the community over Phlogiston's existence, but also that the epistemic principle of intelligibility was already known at the time by Phlogistonians to be misleading, and that the adoption of Phlogiston as a real entity was the result of inadequately appreciating prior lessons from scientific investigation—overlooking the need for well-established methods of investigation (like quantitative measurement) before adopting unnecessary metaphysical commitments.

One way to do this is simply to direct our attention to earlier accounts of past successful but rejected scientific theories that hinged on epistemic principles such as unification, mechanical analogy, and manipulability, and show that the 18th century theorists knew better than to ignore these lessons. A problem with this approach is that it would invite the 18th century realist about Phlogiston to claim, as we would wish to do in respect of *his* methodology, that his methods are far more reliable than those of his predecessors—effectively that he has learned from past mistakes, and hence is justified in his appeal to Phlogiston.

A more amenable, and plausible approach is I think to show that even at its height of success and 'acceptance' Phlogiston should not have been licensed by even the chemical community's own lights, and this is indicated by that fact that many of the leading Phlogistonians who were fully aware of the theory's quantitative inadequacies, *also advocated for the primacy of quantitative rigor*. If this can be shown, then we ought to conclude that not only was there never consensus on Phlogiston's existence, but even those who did accept it failed to follow their own methodological rules.

If this view is correct it provides us with two reasons to dismiss Phlogiston as a significant threat to scientific realism. First, without historical evidence of consensus by a scientific community on the existence of an entity, realists should not today assume to place that theory in their store of successful science—so it doesn't help the anti-realist to use Phlogiston as a weapon in the first place. Second, even for those who did adopt the theory, there is good reason to think they should not have done so by even their own scientific methodology. I think the historical evidence shows that Phlogiston was a promising program, but never reached the point of consensus, and even if it had, it failed to sit consistently with the epistemic principles of the day.

Let me start this argument—that Phlogistonians had good reasons to reject Phlogiston on quantitative grounds—by briefly noting the growth of quantitative rigor outside chemistry long before Phlogiston came on the scene. The 17th century had already seen a dramatic dispute between Newton (Clarke) and the Cartesians over the plausibility of Newton's theory of gravity, which had left many natural philosophers weary of positing unobservable mechanisms to account for empirical observations. A significant mark of physics at the end of this century was of course mathematical rigor, primarily due to the work of natural philosophers like Galileo, Torricelli, and Newton. With this mathematization came the notion of careful quantitative analysis of a system, and this was now seen as a primary epistemic desideratum for any theory in that field.

As the 18th century wore on it became quite common for scientists and philosophers to take the mark of an acceptable theory to be its mathematical representation rather than its possessing an intuitively plausible underlying mechanism.[10] The Phlogiston theory is supposed to have reached its pinnacle of popularity just when such epistemic desiderata were dominant in the physics community, so it is curious that one who thinks there are permanent singular epistemic principles for science (like Structural Realists) should also think that Phlogiston would have been justifiably licensed by the acceptable principles of the day.

More important than these considerations however, is the apparent fact that at the time when Phlogiston was most popular, there was a significant tension between the coherence of the theory and a primary epistemic principle for chemistry. This is as we know the tension between the epistemic principle that 'one ought to achieve significant quantitative measures for the entities in a theory', and the discrepancy that sat between the empirical measure of the weight after burning a substance and the theory's prediction for that process. That is, notable chemists of the day, as well as their predecessors, advocated for theory choice only after the careful and accurate measurement of all of its components, yet this was amongst the primary problem areas for the Phlogiston theory.[11]

To establish this claim let's first look at some Phlogistonians who were explicitly advocating quantitative rigor while the theory was at its height of popularity, and then turn to some recent work by historians which indicates a strong commitment to quantitative methods by leading chemists even in the early 17th century.

Returning to Brown's *A History of Chemistry* we see what he calls the 'quantitative period' begin around 1775—and this is quite a common view in traditional history of science. It is also clear however that this date is only a rough approximation since there are clear cases of quantitative methodology being primary for chemists long before that year. What is a little surprising is that some of those primary examples of quantitative chemists are themselves Phlogistonians. For example, we see on this list the likes of Cullen, Black, Cavendish, Priestley, Bergman, Scheele, and Kirwin. Each of these scientists adhered strongly to the Phlogiston hypothesis despite having a serious commitment to quantitative rigor. Black for example is well known for his attention to experimental accuracy, and played a key role in the eventual downfall of Phlogistion due to his (re-)discovery of what we now call carbon dioxide. His efforts with the balance to establish the nature of air released from carbonates are well known, but success at identifying the gas released as accounting for weight loss was not sufficient for him to reject Phlogiston. Similarly, Cavendish using quantitative methods, famously discovered that gas released from the application of acids to metals (hydrogen)—which is responsible for some weight loss—failed to recognize it for what it

[10] For accessible accounts supporting this interpretation see Dear (2006), De Regt and Dieks (2005).

[11] For details of this case, and evidence supporting the principle, see Newman and Principe (2005).

was, instead maintaining this to be a particular metal form of Phlogiston. And the story is similar for others on our list.

What is somewhat confusing then is that if these chemists had a serious concern for quantitative rigor, why did they adopt a theory with a glaring quantitative anomaly? Brown himself conjectures that this was due to the intelligibility of the theory (1920, p. 276), and highlights the simple error of mistaking a property for a substance (combustability is a property not a substance). However, following Brown, it is quite common to believe that with the arrival of the weight balance, and accurate measurements, Phlogiston was quickly rejected and a more accurate explanation was sought. This is incorrect. We now have good evidence that for over a century before 1775 there was ample use of rigorous quantitative methodology in chemistry, and that chemists of the time were fully aware that Phlogiston (or some precursor) did not quantitatively add up, and should have been rejected—at least treated instrumentally instead of realistically.

For this evidence we can turn to recent work by William Newman and Lawrence Principe, who show not only that quantitative methodology was present in very early chemistry, but that it was in fact rampant throughout the field. This thesis, if correct, strongly supports the view that although chemists of the 18th century had diverse methodological principles, the role of quantitative analysis in the construction and acceptance of theories should not have been overlooked the way some history portrays it. The significant length of time during which leading figures in the field advocated quantitative rigor reveals that this principle was not merely one among many equally weighted competing rules of investigation. Nor was it an inferior principle for chemistry, as some who wish to justify the Pessimistic Induction might suggest. Quantitative rigor was highly respected from the very earliest days of chemistry, and as such should be taken by scientific realists and anti-realists alike as providing very good reason to be suspicious that Phlogiston ever was as successful a theory as they suppose.

So where is the evidence for this rather strong claim? Well, Newman and Principe have argued in their (2002, 2005) that much of history's overlooking the role of quantitative analysis in early chemistry is due to insufficiently appreciating the role of several key influential chemists, and primary amongst them is Joan Baptista Van Helmont. By tracing his influence on subsequent chemists they claim to show that the role of quantitative analysis—especially by use of the mass balance—was significant even in the 17th century (2005, p. 85). This lends support to my rather general claim that such methods were primary in the community because Newman and Principe also contend that "Van Helmont's writings constituted what was probably the most wide-ranging and influential chymical theory of the second half of the seventeenth century" (2002, p. 296). Since quantitative analysis is primary for Van Helmont, if he really was as influential as our historians claim, we have good reason to think mathematical rigor a primary methodological principle for chemists in general in the 18th century, and consequently their acceptance of Phlogiston as a successful theory, as problematic.

We don't have time to enter into the details of Newman and Principe's argument, but a brief sketch of their story will help lend plausibility to my thesis. They focus in their (2005) on the use of assaying processes in 12th and 13th century alchemy, which required determination of the specific weight of substances, and they move on to Geber's work at the end of the 13th century, which advanced the analysis of alloys through cupellation and cementation methods and required careful weighing before and after testing. The subsequent adoption of fire-driven analysis and synthesis methods by Paracelsus, and its adoption and quantification via gravimetric measurement by Van Helmont in the 17th century, are important steps in the quantification of chemistry. By 1644 the latter was advocating conservation of weight in chemical reactions (as a consequence of careful measurements) and this was so important as to become a *named principle* for the community—the 'mass balance'. This principle reflects Van Helmont's commitment to quantitative rigor, and Newman and Principe quote him as claiming that "a mathematical demonstration [is] stronger than any syllogism" (2005, p. 81).

This may be so for Van Helmont, but what of the rest of the chemistry community through the 17th and into the 18th century? On this question Newman and Principe spend a great deal of time explaining how primary figures in chemistry were heavily influenced by, and adopted for themselves, the Helmontian model of using gravimetric techniques, and hence quantitative analysis, in determining chemical truths. Helmont's approach reaches Alexander von Suchten, who himself heavily influences George Starkey, who's techniques themselves are further developed by Wilhelm Homberg into the 18th century. This is a lineage significant for its influence not just across Europe, but also into Britain and across to America. Without here going into further detail, Newman and Principe can best summarize their findings regarding the methods of these chemists in the following way:

> Yet all of these experiments further refined and elaborated the analytical methods which had been developing for centuries within the alchemical tradition. This cumulative work thus produced by the early modern period a generalized technique for chymical study, and perhaps more importantly, an operative mindset towards analysis/synthesis and weight-determination as the basic implements for the investigation and control of material substances. These techniques—worked out in the context of preparing grand chymical arcana—would be inherited and deployed famously by Lavoisier and his successors (like Liebig and other 19th century analytical chemists) down to the present day. (2005, pp. 85–86)

Here is a definitive claim to support my thesis, that during the entire period that Phlogiston theory was developing, and through its peak, right up until its fall, quantitative methods were primary methodological principles for chemists, and as such, it seems questionable to consider the theory nearly as much of a success—and certainly not as much of a 'mature' theory—as those who advocate the Pessimistic Induction take it to be.

3.3 Caloric

Next we consider Caloric, and the question to answer is: was there good reason at the time to accept the view that heat was a fluid entity? Were there more compelling reliable epistemic principles at work which should have made the scientific community suspicious of positing this entity?

This case is on the surface not as easily made as that for Phlogiston since Caloric was capable of explaining a far broader range of phenomena. We see however that the theory suffered from a similar quantitative weakness as Phlogiston—perhaps not entirely surprising since many of the scientists involved were the same people.

Caloric was itself considered a theoretical fluid identified with heat. As a material substance it accounted for the flow of heat from hotter to colder substances, the expansion of substances under heating, the latent heat of substances, radiant heat, heat released from chemical reactions, the springiness of gases, the fluidity of liquids, and it was used in the derivation of the gas laws, as well as Carnot's heat engine experiments. Caloric consisted of material particles, mutually repelling but attracted to ordinary matter, which would flow like water to where there was less 'pressure' from other repulsive particles of Caloric—hence the flow to cooler locations. The simplicity of explanations for heat flow that appealed to Caloric accounted for much of its success, but it was also used to derive the adiabatic gas law and gave more accurate predictions for the speed of sound in air.

This is quite an impressive assortment of accomplishments, but the theory was not without a competitor. Advocates for the alternative vibratory theory of heat went as far back as Bacon, Boyle and Locke, and although typically thought to be averse to the position, even Laplace and Lavoisier sometimes aired their concerns about the material view—although Lavoisier's extremely influential *Traité élementaire de Chimie* (1789) relied heavily on Caloric. However, the most significant opponent to Caloric was Count Rumford who famously showed that experiments boring cannons directly conflicted with a material theory of heat. In particular his experiments on boring cannons with a blunt tool demonstrated that heat was created for as long as the boring continued, which seemed to indicate that Caloric was either an infinitely generated material, or an incorrect explanation. The small amount of metal collected from the boring process was supposed on the material view to explain heating to boiling point 2,300 g of ice-cold water—the amount of Caloric needed to achieve this seemed far out of line with such a small amount of metal. Importantly, the rate at which the Caloric must be generating this heat did not diminish one bit as the boring continued—and as theory would suggest. Rumford even made an estimate of the amount of work that went into generating heat energy and came up with 1 calorie for 5.5 J—pretty close to the correct value of 4.18 J. He urged others to take this as evidence for the vibratory theory, and was followed by Thomas Young and Humphry Davy (who performed diverse experiments in this field) in recommending against Caloric. Joule persisted with experiments illustrating the conversion of mechanical energy into heat, and devised an interesting experiment (1844) whose results were inexplicable on the material approach, yet seemed quite easily derivable from the vibratory view (Torretti 1999, p. 183). It wasn't until 1847 however, and Helmholtz's establishing the conservation of energy that the material theory conclusively fell.

So, what can we select as the epistemic principles by which the Caloric theory was established and retained? Clearly, again, we see the appeal to 'intelligibility' as grounds for justification. The familiarity and clarity of presentation of a fluid description for otherwise confusing phenomena is understandably a strong psychological appeal of the theory. Still, there were clearly empirical, not merely explanatory, problems with the theory. The problem of accounting for all the Caloric operating in the cannon boring experiments is reminiscent of the reluctance of Phlogistonians to pay adequate attention to the quantitative difficulties of calcination. Just as Phlogiston didn't make analytical sense, just as the numbers didn't add up for the weight of a substance after burning, so too in the case of Caloric the quantitative analysis leaves unacceptable lacunas. And perhaps this latter case is the worse since there didn't seem available any (even ad hoc) explanations (unlike the Phlogiston case where it was suggested Phlogiston was of negative weight).

I would suggest then that just as with its predecessor, the Caloric theory of heat was susceptible to overconfidence in the truth-tracking ability of intelligible explanations of the phenomena. Again, it is not that intelligibility ought play no role in the epistemic evaluation of an explanation, but given that the physicists and chemists of the time had seen just this error having been made with Phlogiston, and given that they were also quick to claim the necessity of rigorous quantitative methods (after all some of these experimenters were the very same in both episodes) then it seems that *by their own standards of the day* the choice of a material fluid theory in preference to at least agnosticism seems peculiar.

3.4 Luminiferous ether

Predictively, I wish now to suggest that the last 'hard case' raised by the Pessimistic Induction is also due to a situation where the scientists of the day were quick to overlook their own methodological prescriptions. Although Laudan (1981a) is primarily the culprit we look to for pushing the Pessimistic Induction on us (and in particular the use of the ether as an example of a failed but successful research program), it is also his (1981b) that provides excellent support for my analysis of the differential epistemic value of divergent methodologies. In his paper 'The Epistemology of Light: Some Methodological Issues in the Subtle Fluids Debate' Laudan provides a clear explication of how, when first adopted, ether theories were at odds with the reliable epistemic principles of the day. Starting with the early 18th century he says,

> Induction and analogical reasoning were the rage and Newton's doctrine of *vera causae*… was thought to exclude the postulation of any entity or process not strictly observable…speculative systems and hypotheses are otiose; scientific theories must deal exclusively with entities which can be observed or measured. (1981b, p. 112)

Nevertheless, by the 1760s, he reports, there were a plethora of theories appealing to one form of ether or another in their attempts to explain everything from light and heat, to the workings of the mind and our perceptual system. All of these generally popular approaches, Laudan claims, clearly broke the scientific epistemic strictures of the time, which were undoubtedly empiricist (1981b, p. 113).

In particular, of the early ether theorists, Laudan singles-out Hartley and LeSage as guilty of advocating theories that even they knew clearly went against the reliable empiricist epistemology of their day. Hartley clearly adopted an abductive approach, somewhat akin to inference to the best explanation, supplemented with an appeal to explanatory breadth and unification to bolster his cases—'confirmed explanatory scope' as Laudan calls it (1981b, p. 117). Importantly for us, it is pointed out that,

> Hartley immediately conceded that this criterion does not guarantee that the hypotheses it licenses will be true or that they will even stand up to further testing. They will possess none of the *reliability* (then) associated with the methods of induction and analogy. (1981b, p. 117)[12]

Moving on to LeSage we see a similar story, although here it is emphasized that the reactions to LeSage's ethereal model of gravitational attractions were "not only largely negative; *the grounds for criticism were generally epistemological rather than substantive*" (1981b, p. 119).[13] And it is perfectly clear that LeSage's attempts at appealing to unobservables to explain empirical phenomena were universally rejected on empiricist grounds. Despite his insightful attempts to argue for an important distinction between respectable hypotheses which could be verified, and illegitimate hypotheses which were mere fanciful conjectures incapable of confirmation, LeSage failed to rally against the reliable epistemology of his age. Paradoxically, although scorn on the method of conjecture persisted well into the 19th century[14] ether theories became a popular means by which many scientists attempted to explain phenomena. One of the reasons for this is perhaps that there was an ongoing dispute over whether perhaps there might be epistemological desiderata that identified more than mere empirical adequacy as confirmatory for a theory. In the early to mid decades of the 19th century, Laudan argues, there was a change in methodology—in particular *novel predictive success* became a newly adopted epistemic desiderata (in part due to the discussions of Herschel, Whewell, and Mill).

The attempts by LeSage to dignify hypotheses were now being vindicated, in particular by the 1820s, and it was the notion of novel prediction (and also that of accommodating unexpected domains of phenomena), which provided the necessary distinction between the respectable and the unrespectable hypotheses. This, as Laudan points out, is to give a redefinition of what constitutes *evidence*. And this move occurred fairly quickly. In fact Herschel and Whewell were advocating for this redefinition in the 1830s and 1840s, which was early enough to worry that perhaps such methodology itself required some testing before being taken as a *reliable* indicator of tracking the truth. In fact I think this a plausible interpretation of the error made in the mid-19th century of adopting too hastily ether theories that made novel predictions or exemplified Whewell's *consilience of inductions*. This applies in particular to Fresnel's wave theory of optics, which had the remarkable result of generating the Poisson white spot prediction, yet its apparent ability to satisfy epistemic standards of the day reflects in this case the immaturity of those standards being adopted for this area of science.

Now, if the above historical arguments are plausible then we can conclude that the Pessimistic Induction really isn't a significant hurdle for scientific realism. Given

[12] Emphasis added.

[13] Emphasis in original.

[14] Especially from Scottish philosophers of Reid's 'Common Sense School'.

that the primary methodological principles were not followed in all three most highly contested historical cases that comprise the basis for the Pessimistic Induction—Phlogiston, Caloric, and the Luminiferous Ether—we can reject the anti-realist's claim that our current methodologies are subject to the same unreliability as past methodological principles. We can therefore stop worrying about these cases as counterexamples to the optimistic induction on science. Well, almost. More work has to be done, and of a less historical nature—we need to provide a full justification for how methodological principles ought to be selected and justified, how this can be put into service to defend realism against possible anti-realist objections, and show how it can also provide a plausible account of the progress of science. All of which I turn to next.

4 Justifying epistemic pluralism and accommodating the advancement of science

I have been arguing that the hard cases used as a basis for the Pessimistic Induction fail to be problematic if we look at the epistemic principles that were accepted as reliable at the time in question. The accepted theories should never have been taken as anything more than conjectural, according to methodological rules and epistemic principles thought to be reliable at the time. If this argument is correct, the realist doesn't have to worry about the Pessimistic Induction in the first place. However, there are some significant questions that arise for my way of viewing the methodology of science, and in this section I want to clarify the interpretation I give to the relation between scientific theories and methodological principles, as well as the justification of each. I then wish to argue that my approach, although entirely naturalistic, is still capable of accounting for the genuine phenomenon of the progress of science, and why it is an advance over Structural Realism.

Scientific theories are evaluated based on specific epistemic and methodological principles that scientists use for theory choice. Such principles may include requirements like empirical adequacy, novel predictive success, explanatory breadth, etc. We are very familiar with these notions. But the justifications for selecting members of a set of principles (even if it is just one principle) have to be given. If scientific theories form the bottom of a linear hierarchy then the methodology of science forms the second, more abstract level. In the same way that a scientific theory derives justification from a select methodology, that methodology itself presumably requires justification. This we can visualize as the third level of the hierarchy, where we find meta-methodology. Meta-methodology itself comprises rules or principles of theory choice—methodological theory choice. Our meta-methodological principles tell us which methodological principles should be used for choices regarding first level scientific theory selection. We are climbing a ladder upward, and one might wonder if the meta-methodological principles themselves have justification from above. Here however we reach a point of significant dispute amongst methodologists of science. Some suggest that there are no meta-methodological principles in the first place, and that methodology is self-justifying. Some suggest that meta-methodology is itself justified from below by its subservient methodology. Still others suggest meta-methodology is self-evident, or a priori. And there are variants on each of these approaches.

I recommend we adopt a view on meta-methodology that is as pluralist as my recommendations regarding the second (methodological), level of this hierarchical schema. Where methodological principles in the history of science can be seen to take different weightings within different domains (dependent upon the performance of those principles in those domains at that time), I think we should also see meta-methodological principles as diverse and differentially weighted. What might such a differential weighting of abstract principles look like, and how will each justify the adoption of particular methodological principles below? Actually, the picture is quite simple, it merely requires we relinquish the assumption that our notions of justification are themselves singular epistemic principles. The general approach towards this meta-methodological level is not as purely naturalistic as that at the methodological level proper, but still requires the cooperation of the world. We look to our intuitions about notions like epistemic justification and accept that this notion itself is comprised of diverse values or principles. Justification is not merely a matter of using reliable processes for belief formation, as many suspect, it may also include other epistemic desiderata such as our having adequate evidence (reasons) for adopting some particular methodology (which will presumably come from the kind of naturalist examination of the history of science proposed in the previous section). Another epistemic desideratum for determining which methodological principles to adopt, and how heavily to weigh them, might come from a prior commitment to a set of intellectual virtues pushed by virtue theorists in epistemology. Other candidates might include desiderata we find at the methodological level itself, such as principles that select theories based on their ability to unify phenomena, their coherence, their plausibility, etc. All of these methodological principles (and many others besides) which operate with different degrees of success (and hence should be accorded different weights given their context) can be applied at the meta-level for selecting methodological principles.

The problem immediately arises however of how we are to evaluate the weights we are to assign to principles at the meta-level. If we think we have difficulty at the methodological level, assigning weights based on such vague notions as ‘success’, ‘domain’ and ‘at a time’, how on earth is one to accomplish the same task at the even more abstract level of the meta-methodology of science? Furthermore, a naturalistic pluralism regarding methodology, such as is promoted here, seems to face the problem of still having to provide an account of something most of us take as obvious about science—its progressive nature. How, given a contextual pluralism regarding scientific methodology, can we account for progress in science?

These are difficult questions for any account of methodology which appeals to naturalistic methods, as this one does, and so what follows is somewhat tentative. Still, it seems plausible that we can make some headway in answering these challenges by looking to the very notion, or concept, of science itself. If science is said to instantiate *any* set of properties, reliability surely must be primary. Even those who take a pragmatic approach to science can be interpreted as accepting that by its very definition science is supposedly reliable.[15] What, after all, would it mean to have a scientific discipline whose claims are unreliable? That would be no kind of science at all.

[15] For example, Laudan (1996) clearly takes science to require reliable procedures.

If reliability is the hallmark of science, then although it need not be our only meta-methodological rule for choice between methodologies, it would appear primary amongst them. We can now ask ourselves the same question but now at the meta-level, what would it mean to have a rule that selected methodological principles that was itself unreliable? This would be no effective strategy for selecting rules that generate successful scientific theories. So, I take it that however we determine a set of meta-methodological principles for choice of methodology, they should jointly accomplish the task of reliably selecting methodological principles that themselves reliably generate successful scientific theories. A naturalistic analysis of how to carry out such an achievement might follow the strategy above for the selection of weightings for methodological rules itself, but I leave that open here.

Assuming that this appeal to reliability as the primary meta-methodological principle is not implausible, we might bolster its cause by providing a reliabilist defense of the principle. This tactic might appeal to rule-circularity to defend the meta-methodological principle itself. One could argue that the meta-methodological principle for reliability says something like (R): 'Select only those methodological principles of theory choice which are reliably successful at producing successful scientific theories'. To defend this principle one might suggest the rule-circular argument:

R is a reliable rule for selection of methodological principles, (Application of R), therefore, R. This inductive argument secures the inference to R only by using the premise that R genuinely is a reliable rule for selecting methodological principles. This is in principle determinable by looking at the historical record itself. However, such arguments are unconvincing to those who reject anything but an externalist justification for one's inference rules. An epistemic externalist claims that to be justified in adopting a belief we do not need *positive* evidence for that belief, we only need to be using a reliable process in forming that belief. Given this epistemology, one does not need to have transparent access to positive reasons for adopting a rule of inference. If the world cooperates by not contradicting our inferences, that is reason enough to think them reliable. One can therefore defend just such an argument as that for (R) above.

Unfortunately, the plausibility of epistemic externalism is still up in the air, and absent an adequately developed epistemology, one's appeal to Normative Naturalism seems shaky. Still, if externalist reliabilism in epistemology is one's position (and it is for many philosophers of science), then Normative Naturalism of the variety I advocate here should be plausible. In particular, although I am not advocating solutions to the problems facing epistemic externalism, for those who already adopt the position it seems reasonable to adopt a contextual pluralism regarding methodology, and use this liberally to dissolve the Pessimistic Induction.

This appeal to externalism raises another potential objection: why not simply appeal to reliabilist epistemology to underwrite one's inferences at the methodological level to start with? If we have no good reason to doubt our methodological rules at this level then realism is saved without having to mess about with appeals to epistemic desiderata at the meta-methodological level. In fact this strategy is adopted by Psillos in his (1999), but what is crucially overlooked in that case is the vicious circularity of appealing to the reliability of inference to the best explanation when one seems to have evidence from the history of science to the contrary, in the form of the Pessimistic

Induction. I do suggest that for those who adopt such an externalism, appeals to reliabilism cannot secure a defense of scientific realism against the Pessimistic Induction by using something like inference to the best explanation. To assume reliability of inference to the best explanation to answer the anti-realist is to beg the question.

With this vicious circularity threatening uses of reliabilism to secure inference rules, why doesn't it similarly affect my appeal to reliability of inferential rules like (R) at the meta-level? Surely there is a methodological Pessimistic Induction to be had here, where we can see through the history of science that different rules have taken different weights of epistemic reliability—that after all is precisely what contextual pluralism teaches us.

This much seems correct, but the pluralism I recommend may have the resources to look at the history of science and claim that when there have been failures of particular methodological principles (such as I argued there were for the intelligibility principle in the cases of Phlogiston, Caloric, and the Luminiferous Ether) then that is where the *metaphysical content of our theories has had to change.* The argument would be unique to inference rules at the meta-methodological level only, and not those like inference to the best explanation, which reside at the methodological level below. The case would have to be made historically, and I have no space to do so here, but it doesn't seem implausible from a glance at the history of science, to suggest that where cases of theory choice have proved problematic (and we see the failure of a previously reliable methodological rule of theory choice in some specific domain), then that failure can be explained by showing that science reached a point of remarkable discovery—for example that of physically instantiated and causally efficacious vector fields, of quanta, or whatnot. That is, when methodological rules have broken-down, and failed to maintain their previous reliable ability to pick-out successful theories, then that is because our science is running-up against a part of the world that requires we change some metaphysical presuppositions.

Now this is not much more than a promissory note for how one might defend meta-methodological principles via an externalist analysis, but this project is in its infancy and already the historical record looks promising.[16] As John Losee notes in regards to the methodological principle of 'explanatory completeness' and its failures when facing the Quantum revolution:

> *Prima facie*, abandonment of the earlier standard of explanatory completeness is evidence against the inductive principle. However, one may choose to salvage the inductive principle by claiming that evaluative situations in the quantum domain are *different in kind* from previously encountered evaluative situations. (2004, p. 139)

One may find this kind of 'running-to-new-metaphysics-defense' for epistemic principles to just reek of *ad-hocness*. However, by appeal to the success of subsequent scientific theories that abandoned the principle in question, and at the same time

[16] Donovan et al. (1988) made a first step along these lines, but made no effort to defend their adoption of meta-methodologial principles. Losee (2004) and Nola and Sankey (2007) have made some progress on this issue, but in their accounts the defense of externalism is left vulnerable to the vicious circularity objection which threatens from a methodological Pessimistic Induction.

adopted new metaphysical properties previously unavailable to earlier theories, one may generate support for just such an appeal. I leave further details of this potential defense for another time.

One benefit that certainly follows from finding a rescue from a Methodological Pessimistic Induction via appeal to changing ontology in our theories is that we can turn this account directly into an explanation for the apparent progress of science. A contextual pluralism about matters methodological appears on the surface to stymie any hope of generating an account of scientific progress. This is because with scientific theory-selection-principles diverging throughout the history of science we seem to lose the ability to argue for any methodologically privileged accumulation of scientific knowledge. The Structural Realist at least held out the hope of accounting for progress despite radical theoretical discontinuity by appealing to structural continuity. Progress is accomplished on that approach in virtue of the accumulation of theoretical truths regarding the structure of the world. Without a single epistemic principle to privilege particular kinds of theoretical components how can this contextual approach maintain any claim to the effect that science is advancing?

The answer to this problem is found buried in the arguments above; for any form of scientific realism (which takes the primary aim of science to be the attainment of theoretical truths), justifying the progress of science requires justifying one's claim that it really is accumulating more and more truths, and in particular, truths that are significant (rather than merely trivial).[17]

My account of contextual Normative Naturalism is capable of achieving this because it advocates the adoption of primary methodological principles *for a specific domain at a specific time*, by appeal to higher-level meta-methodological principles. I suggest that primary among such meta-methodological principles is the notion of reliability, and hence our adoption of any single methodological principle had better be reliable in order to satisfy the meta-level conditions for being primary. We find in the history of science that where previously reliable methodological principles have come to fail to guide us towards the truth, this is because either the principle was mistakenly taken to be primary when in reality it was at best secondary, or the metaphysical presuppositions of the domain in question have had to be adjusted. In the former case we have seen that historians can be invaluable in clarifying whether specific principles are genuinely reliable, and hence primary—as we saw is problematic with Phlogiston, Caloric, and the Luminiferous Ether. For these hard cases where history shows us the scientific communities were acting against their own methodological prescriptions we can conclude the Pessimistic Induction fails because the theoretical entities which proved to be problematic never should have been treated realistically. In the latter case, where the genuinely primary methodological principles have been reliable for a domain but fail to continue to be reliable in that same domain, I suggest history teaches us that this is when genuinely revolutionary changes in underlying ontology are occurring—we are unveiling a new *kind* of ontology at the unobservable level.

[17] Bird points out in his (2007) that it is really accumulation of theoretical knowledge, and not merely truths, that the realist must defend if he is to defend the progress of science thesis. The difference between the two being that knowledge requires we not only attain truths, but that we do so with the right sort of justifications as well.

This is certainly progressive, but does raise concerns about how contextual Normative Naturalism can advocate a realist position. That is, how can such revolutionary changes in ontology not justify anti-realism towards even reliable methodology? This is a very important point for Normative Naturalism—a Pessimistic Induction on the methodology of science would appear devastating to the position. Still, here one must remember that on an externalist epistemology one is justified in a realist interpretation of theoretical entities if one has no reason to doubt their existence. With the hard cases from the history of science dissolved, we no longer have significant worries about primary principles being refuted. Even where such worries linger (for example with explanatory completeness and the transition from classical to quantum mechanics), as Losee has pointed out, we might well consider the reliability of the principle to be just as good as before, but now restrict its use to the former domain. In the quantum domain explanatory completeness is not a reliable principle, so inferences made using it as a methodological principle would be illegitimate. In the classical domain such inferences may still be legitimate. So, contextual Normative Naturalism is capable of justifying realism, but it is a qualified realism which requires a well tested set of methodological principles for any given domain of science and only for the period of time for which those principles are reliable. This requires we be willing to characterize past science (such as classical mechanics) as true, but true relative to a methodology. This is fully within the traditional fallibilist approach scientific realists usually adopt—after all no one expects certainty for science, but realist inferences derived via well-tested reliable methodology provide us with good reasons for optimism, and undermine pessimistic anti-realist worries derived from the history of science. So, the approach I am pushing advances beyond Structural Realism because it is not limited to an overly restricted epistemic principle—reliable methodologies of any kind are acceptable for any specific domain, and hence we are not constrained by non-explanatory methods. For example, if intelligibility turns out to be reliable for a domain, then that is perfectly alright—so long as it has proved itself genuinely reliable and we are careful to delineate exactly which domain within which it is operative.

This framework also has the pleasant result of answering the problem burdening preservative realism—its being entirely backward-looking. Recall that Structural Realism suffers from the limitation of only sanctioning realism about retained structure, and hence has to wait for theory transitions to occur before being capable of prescribing what we ought to believe about our best theories. Contextual Normative Naturalism suffers no such defect though because when one has evidence that a methodological rule is reliable in a domain, and there are no reasons to think such primary rules are vulnerable to a Pessimistic Induction, then on an externalist epistemology one has no reason to be skeptical of inferences made using such a rule. We can go even further than this. Not only are we justified in believing our inferences about the unobservable are legitimate for today's science, we can go so far as to claim that so long as our domain of inquiry remains constant then all well reasoned inferences using primary principles will continue to be justified. So, contextual realism can solve both the Pessimistic Induction and the projectability issues for scientific realism.

There is however at least one more concern regarding contextual Normative Naturalism. I have appealed to the primacy of the meta-methodological principle of reliable rule selection, while leaving open the possibility of a pluralism of principles at even this level with which to make choices between methodologies. It seems that this might usher-in a rampant relativism regarding progress in science, yet I wish to resist this conclusion. If, as I suspect, reliability is empirically discovered to be the most effective rule of selection (overtaking other rules such as 'select the principles that are internally maximally coherent' or 'select rules that select theories which explain all of a previous theory's mistakes'), then this rule of selection should be most heavily weighted. In this case, one can simply infer that the progress of science results from the steady accumulation of more and more reliable theories. Now, a die-in-the-wool anti-realist might balk at this, suggesting that reliability is now just being used in place of success, and we saw that the latter couldn't be used to infer truth so why think the former any different? However, again, such criticisms hang on the rejection of a reliabilist epistemology, and with application to scientific method this has not yet been established.

On the other hand, if one rejects the primacy of reliability at the methodological level, seeing in the historical record much greater success attributable to methodologies that select scientific theories under different rules (perhaps maximal coherence, or simplicity) then one is still appealing to the meta-methodological rule (R) above, and the reliabilist argument can be run again.

Progress on this contextual approach therefore is more complicated than traditionally assumed because it is measured according to the changes over time of the appropriate methodologies in a domain. Reliable selection of methodological rules over time could mean that as science changes we adopt new standards of theory appraisal, and this is fine so long as they reflect accompanying changes in metaphysics for our scientific theories. Progress is achieved with the continued reliability of our selection rules (methodology) no matter whether it is itself calling for predictive novelty, maximal coherence, or what. We saw that intelligibility failed to sustain progressive research programs through the history of science, and we now have good evidence that this was because the ontology of our theories required transformation. Progress was achieved because the principles being used to select successful theories had to be updated, and with them came new ontological discoveries. All of which can be characterized broadly as the accumulation of knowledge, but with the caveat that such knowledge accumulation will almost certainly not remain representable via some static epistemically privileged principle like that pushed by Structural Realists.

References

Bird, A. (2007). What is scientific progress? *Nous (Detroit, Mich.), 41*, 64–89. doi:10.1111/j.1468-0068.2007.00638.x.

Boyd, R. (1981). Scientific realism and naturalistic epistemology. In P. D. Asquith & T. Nickles (Eds.), *PSA 1980* (Vol. 2). East Lansing, MI: Philosophy of Science Association.

Boyd, R. (1984). The current status of the realism debate. In J. Leplin (Ed.), *Scientific realism*. Berkeley: University of California Press.

Brown, J. C. (1920). *A history of chemistry from the earliest times to the present day*. London: J&A Churchill.

Buchwald, J. (1989). *The rise of the wave theory of light*. Chicago: University of Chicago Press.

Buchwald, J., & Franklin, A. (Eds.). (2005). *Wrong for the right reasons*. Dordrecht: Springer.

Cei, A. (2006). Structural distinctions: Entities, structures, and changes in science. *Philosophy of Science*, Supplement: Proceedings of PSA.

Chakravartty, A. (1998). Semirealism. *Studies in History and Philosophy of Science, 29*, 391–408. doi:10.1016/S0039-3681(98)00013-2.

Chakravartty, A. (2003). The structuralist conception of objects. *Philosophy of Science, 70*, 867–878. doi:10.1086/377373.

Chakravartty, A. (2004). Structuralism as a form of scientific realism. *International Studies in the Philosophy of Science, 18*, 151–171. doi:10.1080/0269859042000296503.

Chang, H. (2003). Preservative realism and its discontents: Revisiting caloric. *Philosophy of Science, 70*, 902–912. doi:10.1086/377376.

Davis, K. (1966). *The cautionary scientists*. New York: Putnam.

De Regt, H., & Dieks, D. (2005). A contextual approach to scientific understanding. *Synthese, 144*, 137–170. doi:10.1007/s11229-005-5000-4.

Dear, P. (2006). *The intelligibility of nature*. Chicago: University of Chicago Press.

Donovan, A., Laudan, L., & Laudan, R. (Eds.). (1988). *Scrutinizing science: Empirical studies of scientific change*. Baltimore: Johns Hopkins University Press.

French, S., & Ladyman, J. (2003). Remodelling structural realism: Quantum physics and the metaphysics of structure. *Synthese, 136*, 31–56. doi:10.1023/A:1024156116636.

French, S., & Saatsi, J. (2006). Realism about structure: The semantic view and non-linguistic representations. *Philosophy of Science*, Supplement: Proceedings of PSA.

Giere, R. (1997). *Understanding scientific reasonsing* (4th ed.). Belmont: Wadsworth.

Hudson, J. (1992). *The history of chemistry*. London: Macmillan.

Kitcher, P. (1993). *The advancement of science*. Oxford: Oxford University Press.

Kuhn, T. (1962). *The structure of scientific revolutions*. Chicago: University of Chicago Press.

Ladyman, J. (1998). What is structural realism? *Studies in History and Philosophy of Science, 29*, 409–424. doi:10.1016/S0039-3681(98)80129-5.

Laudan, L. (1981a). A confutation of convergent realism. *Philosophy of Science, 48*, 19–49. doi:10.1086/288975.

Laudan, L. (1981b). *Science and hypothesis: historical essays on scientific methodology*. Dordrecht: D. Reidel Publishing Company.

Laudan, L. (1984). *Science and values*. Berkeley: University of California Press.

Laudan, L. (1996). *Beyond positivism and relativism*. Boulder: Westview Press.

Leicester, H. M. (1956). *The historical background of chemistry*. New York: John Wiley and Sons.

Leplin, J. (1997). *A novel defense of scientific realism*. New York: Oxford University Press.

Levere, T. (2001). *Transforming matter: A history of chemistry from Alchemy to the Buckyball*. Baltimore: John Hopkins.

Losee, J. (2004). *Theories of scientific progress*. New York: Routledge.

Mason, S. (1962). *A history of the sciences*. New York: Macmillan.

Newman, W., & Principe, L. (2002). *Alchemy tried in the fire*. Chicago: University of Chicago.

Newman, W., & Principe, L. (2005). Alchemy and the changing significance of analysis. In J. Buchwald & A. Franklin (Eds.), *Wrong for the right reasons*. Dordrecht: Springer.

Nola, R., & Sankey, H. (2007). *Theories of scientific method*. Ithaca: McGill-Queens University Press.

Partington, J. R. (1965). *A short history of chemistry*. London: Macmillan.

Perrin, C. E. (1990). The chemical revolution. In R. C. Olby, et al. (Eds.), *Companion to the history of modern science*. London: Routledge.

Psillos, S. (1999). *Scientific realism: How science tracks truth*. New York: Routledge.

Saatsi, J. (2005). Reconsidering the Fresnel–Maxwell theory shift: How the realist can have her cake and EAT it too. *Studies in History and Philosophy of Science, 36*, 509–538. doi:10.1016/j.shpsa.2005.07.007.

Thagard, P. (1992). *Conceptual revolutions*. Princeton: Princeton University Press.

Torretti, R. (1999). *The philosophy of physics*. Cambridge: Cambridge University Press.

Worrall, J. (1989). Structural realism: The best of both worlds. *Dialectica, 43*, 99–124. doi:10.1111/j.1746-8361.1989.tb00933.x.