

# Expanding Our Grasp: Causal Knowledge and the Problem of Unconceived Alternatives

Matthias Egg, University of Lausanne

November 21, 2013

(forthcoming in *The British Journal for the Philosophy of Science*)

**Abstract** I argue that scientific realism, insofar as it is only committed to those scientific posits of which we have causal knowledge, is immune to Kyle Stanford's argument from unconceived alternatives. This causal strategy (previously introduced, but not worked out in detail, by Anjan Chakravartty) is shown not to repeat the shortcomings of previous realist responses to Stanford's argument. Furthermore, I show that the notion of causal knowledge underlying it can be made sufficiently precise by means of conceptual tools recently introduced into the debate on scientific realism. Finally, I apply this strategy to the case of Jean Perrin's experimental work on the atomic hypothesis, disputing Stanford's claim that the problem of unconceived alternatives invalidates a realist interpretation of this historical episode.

## 1 Stanford's Argument from Unconceived Alternatives

Kyle Stanford (2006) argues against scientific realism by claiming that much of our scientific knowledge is subject to what he calls 'the problem of unconceived alternatives' (PUA). The importance of this problem lies in its detrimental effect on eliminative inferences, that is, inferences which proceed by formulating different hypotheses, testing them and, if only one of them survives all the tests, concluding that it must be true. As Stanford notes, such a procedure is only reliable 'when we can be reasonably sure that we have considered all of the most likely, plausible, or reasonable alternatives before we proceed to eliminate all but one of them' (2006, 29). Therefore:

**(PUA)** Many eliminative inferences in science are unreliable, because there are plausible and sufficiently distinct alternative hypotheses that are not taken into consideration.

This formulation of the PUA is intended to show that Stanford is not attempting to undermine eliminative inference in general, but only some instances of it. The interesting question then is, of course, which instances these are. Stanford's answer is that the PUA threatens 'our efforts to theorize about the most fundamental aspects of the constitution and dynamics of the various domains of nature' (2006, 33). In order to argue for this claim, he identifies in the historical record of fundamental theorizing a pattern of 'recurrent, transient underdetermination' (RTU) and then performs what he calls a 'new induction' (NI)<sup>1</sup> so as to generalize from past theorizing to all (including present) theorizing. In more detail, the argument goes as follows:

---

<sup>1</sup> By calling his argument a 'new induction', Stanford emphasizes the contrast (but also the connection) between his argument and the pessimistic (meta-) induction

(RTU) Past theorists failed to consider relevant and radically distinct alternatives to their own theories, namely the theories that came to be accepted in the course of later inquiry. These theories were plausible, since they were at least as well confirmed by the available evidence as the previous theories. This shows that past theorists were subject to the PUA in their fundamental theorizing.

(NI) Since present theorists are not relevantly different from past theorists, it follows by induction that they, too, are subject to the PUA in their fundamental theorizing.

The immediate consequence of (NI) is that realism about the claims of present fundamental theories should be rejected insofar as these claims result from eliminative inferences. Stanford (2006, chap. 8) thus advocates an epistemic instrumentalism with respect to these claims.

Equipped with this rough sketch of Stanford's argument, I will in the following section give an overview of the most important objections raised by his critics, grouping them according to which link of Stanford's argumentative chain they are attacking. I cannot discuss these objections in detail, but I will highlight some of their shortcomings, in order to show that more work is required to refute Stanford's argument. Section 3 then attempts to do the necessary work by fleshing out a response to Stanford's challenge, based on the notion of causal knowledge. Finally, section 4 applies these considerations to a concrete case from the history of physics, on which Stanford has commented in some detail: Jean Perrin's work on Brownian motion and the atomic hypothesis.

## 2 Previous Attempts to Undermine the PUA

Realist criticisms of Stanford's reasoning fall into three camps. The first denies that the historical cases of unconceived alternatives mentioned in (RTU) are plausible in the relevant sense, the second denies that these alternatives are sufficiently distinct from the theories available at the time, and the third seeks to block the induction described in (NI) by pointing out differences between the past and the present development of science. I will discuss these three camps in turn, arguing that none of them has so far produced a completely convincing response to the PUA.

### 2.1 The plausibility of unconceived alternatives

Unconceived alternatives can only generate a problem for scientific realism if they have some plausibility. Stanford (2006, sec. 1.2) is therefore very careful to distance the PUA from what he calls 'Cartesian fantasies', that is, far-fetched sceptical scenarios which, if taken seriously, would call into question any knowledge claim whatsoever. One line of criticism against Stanford is that the notion of plausibility employed in (RTU) threatens to collapse this distinction between serious alternatives and sceptical fantasies. To see how, notice first that it is not part of this notion that the alternative theories would have seemed plausible to the scientists who had failed to conceive of them: Although special relativity seems plausible to us, it is unlikely that 18th century physicists would have found it plausible, had they conceived of it (Magnus 2006; 2010, sec. 4).<sup>2</sup> Stanford (2009a, 380–1) admits this, but he insists on calling such theories 'plausible' simply because they actually came to be accepted by

---

usually attributed to Laudan (1981). For a detailed analysis of how the NI differs from the 'old' pessimistic induction, see Magnus (2010).

<sup>2</sup> Indeed, it has been argued that special relativity was not just implausible, but even inconceivable for these physicists (French 2011, sec. 2).

later scientific communities. But then, the critic continues, we should also regard some sceptical scenarios as ‘plausible’ in this sense. For example, we presently find the possibility that all our sensory impressions are generated by a giant computer simulation wildly implausible, in the same way as Newtonian physicists would have found relativity theory implausible. But if someone in the future discovered that the simulation scenario was actually true, then we would have to conclude that it had been a serious possibility all along, hence it should count as ‘plausible’ in Stanford’s sense (Magnus 2010, 810). If this were so, then Stanford’s cherished distinction between Cartesian fantasies and scientifically serious possibilities would collapse, and the realist could conclude that we need not worry about the PUA any more than about Cartesian scepticism.

The problem with this objection is that it ignores the role of RTU in delimiting the scope of the PUA. The history of science shows that there were unconceived alternatives to past scientific theories and that they are plausible in the above-mentioned sense of being accepted by a later scientific community. By contrast, the history of pondering on sceptical scenarios of the computer simulation type does not show any such development and therefore does not give us any reason to take these possibilities seriously. Therefore, the PUA is a distinctive challenge which should worry the scientific realist, quite apart from general sceptical considerations.

## 2.2 The distinctness of unconceived alternatives

The second way to attack the premise (RTU) is to deny that the alternative theories which past scientists failed to consider are in fact ‘radically distinct’ from the theories they did consider. If there is sufficient continuity across scientific theories over time, then the existence of unconceived alternatives is no longer a threat to realism, because these alternatives, although unconceived, can be expected to be relevantly similar to the theories at hand. The PUA is then no longer a reason to doubt that these theories are at least approximately or partially true. Since realists have employed similar strategies to counter the conventional pessimistic (meta-)induction, Stanford devotes two chapters (2006, chaps. 6 and 7) to criticizing such attempts.<sup>3</sup> His central claim is that the appeal to continuity is vacuous, because the realist can only retrospectively identify those parts or aspects of earlier theories that were retained in later theories, and consequently, this strategy ‘allows us to trust only *some* of what current theories tell us about the natural world ... while leaving us completely unable to be confident in our ability to discern just which parts of our theories *actually* constitute this privileged class of theoretical claims’ (2006, 153).

In response, realists have insisted on our ability to identify the trustworthy parts not only of past theories, but also of our current theories. Our strategies for performing this difficult task may not yet be perfect, but there is no principled reason to think that it cannot be done (Psillos 2009, sec. 4.2). Furthermore, continuity between successive theories does not necessarily require that some parts of the earlier theory are explicitly retained in the later theory, it is sufficient that the earlier theory survives as a limiting case of the later theory (Enfield 2008). But in either of these cases, the claim that it is in principle possible to identify some aspects of current theories which are likely to survive future theory change does not go to the heart of Stanford’s challenge. A convincing response to his critique of previous realist

---

<sup>3</sup> The fact that these realist strategies are equally effective against the (old) pessimistic induction as against the NI (if they are effective) has prompted Chakravartty (2008, 153) to call the NI ‘a novel red herring’. This verdict is justified as long as one is only concerned with this particular type of realist strategy. However, as Magnus (2010, sec. 5) points out, there are other realist strategies which are effective against the pessimistic induction, but not against the NI.

strategies must do more than simply assert that one can do better; it must show explicitly how one can do better.

Several ways to do so have been proposed in the recent literature. One group of strategies seeks to identify those parts of theories which are in some sense responsible for their empirical successes, and claims that we can expect these to be retained in later theories. The focus here can be either on successful novel predictions (Saatsi 2009) or on the comparative success of a theory relative to its predecessor (Harker 2013). One problem for this approach is that the fixing of criteria for what is to count as genuine success leads to a dilemma, as Stanford (2009a, 384) points out: On the one hand, if the criteria are too strict, then much of contemporary science will be excluded from realist treatment. For example, exclusive concern with novel predictive success at the expense of explanatory success will result in antirealism about large parts of some scientific fields, such as geology or evolutionary biology. On the other hand, even a very restricted realism seems to be confronted with some counterexamples, such as Poisson's spectacularly successful novel prediction of a bright spot at the centre of the shadow of a circular disc, which rested on an assumption later abandoned, namely the ether hypothesis.

Another problem for this type of strategy is that it is often hard to tell which theoretical assumptions are really responsible for a particular empirical success. For illustration, consider the example of Rutherford's atomic model, as discussed by Harker (2013). Rutherford's model differed from Thomson's previous model in two respects, by postulating (i) a positively charged nucleus and (ii) a uniformly distributed negative charge surrounding it. Harker (2013, 80) admits that Rutherford used both of these assumptions 'for purposes of explaining the scattering results', which constituted the main empirical success of his model compared to Thomson's. But surely we would not want to be realists about assumption (ii), so we need to claim that only assumption (i) was responsible for the model's success. The challenge then is to distinguish in a principled way between those theoretical postulates which were responsible for success and those which were not. Unfortunately, Harker's observation that 'by Rutherford's own admission, [assumption (ii)] was an idealization and in no sense confirmed by the data' (91) does not go beyond Psillos's (1999, chap. 6) earlier attempts to base such a distinction on the scientists' own judgments, which Stanford (2006, sec. 7.3) has shown to be unreliable (see also Chang 2003).

A final proposal on how to identify trustworthy theory parts uses the notion of 'minimally interpreted mathematical parts' of theories (Votsis 2011). The guiding idea is that mathematical structure is often preserved in theory change, and an appropriate interpretation of that structure might yield a realism capable of withstanding the PUA. However, much hinges on what exactly one means by 'minimal interpretation' here. Votsis (2011, 1229) characterizes it as 'the bare minimum that needs to be assumed in order for there to be an appropriate inferential relation' from the equations to the empirical predictions. Without further qualification, this seems compatible with a strict instrumentalism, according to which theories are nothing more than useful tools for deriving empirical predictions. Hence, it is not immediately clear how the appeal to minimally interpreted mathematical parts of theories is supposed to support any kind of realism. At least, more needs to be said about what is to count as an appropriate minimal interpretation, a question to which I will return in subsection 3.2.

### **2.3 The induction from past to present**

Having discussed several arguments against the premise (RTU), let us now consider the inductive step performed within (NI). The validity of such an inference from the

history of science to our present situation depends on the assumption that the present is relevantly similar to the past. This assumption can be challenged in various ways.

First, if the evidence supporting our present best theories is qualitatively better than the evidence that supported our best theories in the past, then an induction from past to present theories might be problematic. In other words, while an induction of the (NI) type may count as some kind of second-order evidence against the truth of present theories, such evidence needs to be balanced against the first-order evidence we have in support of these theories (Psillos 2009, sec. 4.4), and this evidence is, in some cases at least, massively more impressive than the evidence for the theories on which Stanford bases his NI (Votsis 2007). However, as a general strategy against the NI, this idea faces the dilemma familiar from the previous subsection: If the standard of first-order evidence required for realism is set too high, then much of present science will fail to meet that standard; if it is too low, then many theories now considered false will meet the standard, and the NI will not be blocked.

Furthermore, pointing to differences between past and present theories is not a sufficient response to the NI, because the NI is, as Stanford (2006, sec. 2.3) emphasizes, an induction over theorists rather than theories. The second way of criticizing the NI takes this fact into account and argues that present scientists differ relevantly from past scientists, due to improvements in scientific methodology (Roush 2010, Devitt 2011). It seems clear that there is indeed such a difference between the past and the present. What is less clear is whether the difference in scientific methodology bears on how well scientists are equipped to deal with the PUA.<sup>4</sup> If improvements in methodology just lead to more empirical success, but not to a better grasp of the space of alternative theories, then present theorists do not differ relevantly from past theorists, as far as the PUA is concerned.

Finally, one might argue that Stanford's induction from past to present can be blocked even if it is agreed that present theorists are, individually, not relevantly different from past theorists. This is because the scientific community is simply much larger now than it was in the past (Fahrbach 2011), and because later scientists have had more time for their increasingly robust scientific traditions to uncover alternatives (Ruhmkorff 2011). But notice that these developments have been accompanied by other developments that might well have the opposite effect: It is at least plausible that the contemporary mechanisms of research funding lead to a decrease in the investigation of fundamentally distinct alternatives to dominant theories (Stanford 2006, 132). Therefore, although there is an undeniable difference between the scientific community in the present and the past, it is not self-evident that this difference undermines the NI.

### **3 Causal Knowledge as a Criterion for the Realist**

As the discussion in subsection 2.2 has shown, the realist's appeal to continuity across theory change as an argument against the PUA needs to be made precise in terms of a criterion that permits the identification of parts of theories which are likely to be retained in the future. Furthermore, this criterion must be applicable

---

<sup>4</sup> While Devitt (2011) does not address this issue, because he thinks that the onus is on the antirealist to show that the past and the present are relevantly similar (290n11), Roush (2010, 55) explicitly connects improvements in methodology with the PUA. In support of the view that there is such a connection, she refers to her earlier treatment of Perrin's work on the atomic hypothesis (Roush 2005, 218-223). I will discuss this example in section 4, where we will see that her approach is, by itself, not sufficient to establish the immunity of the atomic hypothesis with respect to the PUA.

prospectively, that is, it must not depend on knowing which parts were in fact retained in later theories. Chakravartty (2008, 155) proposes ‘detailed causal knowledge’ as such a criterion, arguing that ‘if one has a detailed enough causal knowledge of something, knowledge that allows one to manipulate it in highly systematic ways, then there is no better warrant for knowledge’. This suggestion alone would not go a long way towards rebutting (NI), since Stanford has already dealt with realist attempts to capitalize on causal considerations. But as I will now try to show, Chakravartty’s proposal has the resources to withstand Stanford’s criticisms. However, his account is rather sketchy, so it is necessary to look at the relevant passages of Stanford (2006) in some detail, in order to see why Stanford’s criticism of earlier causal strategies invoked by realists does not apply to Chakravartty’s strategy.

### 3.1 How Chakravartty’s proposal differs from earlier causal strategies

Stanford (2006, sec. 6.3) discusses several realist attempts to argue for successful reference of theoretical terms despite changing theoretical descriptions associated with these terms. Of interest for our purpose is his critique of Psillos’s ‘causal-descriptivist’ account of reference, which argues for referential continuity across theory change by allowing for changes in theoretical descriptions as long as they leave the ‘core causal description’ of a given entity intact. For example, Psillos (1999, 286) claims that there is referential stability between nineteenth and twentieth century electrodynamics in that ‘the term “luminiferous ether” may be seen as referring to the electromagnetic field’, because ‘the core causal description associated with the term “electromagnetic field” takes up the core causal description associated with the term “ether”’. Stanford’s (2006, 151) reply is as follows:

Of course, this account of the matter invites the realist to choose the core causal descriptions she associates with the central terms of past theories rather carefully, with one eye on current theories’ claims about nature, so there is more than a whiff of ad hoc-ery about the proposal.

And further:

This case for the referential status of central terms in successful past theories simply invites from the historical record a renewed form of the pessimistic induction itself, this time concerning our ability to distinguish (at the time a theory is a going concern) which of our beliefs about an entity are *actually* part of its core causal description. (152)

What Stanford has in mind here is the fact that past scientists have repeatedly included claims we would now consider utterly false in the core causal descriptions of their theoretical entities. As an example, he mentions Maxwell’s claim that the ether must be a material substance, because the energy transmitted by electromagnetic waves ‘cannot be contained in any vessel except the inmost substance of material things’ (Maxwell [1873] 1955, 493; quoted in Stanford 2006, 152). Generalizing from such examples in the spirit of the pessimistic induction, Stanford concludes that ‘we cannot rely on our *own* judgments about which of the descriptions we associate with our own terms are *genuinely* part of their own core causal descriptions’ (153).

Returning now to Chakravartty’s proposal, it is tempting to identify what he calls ‘detailed causal knowledge of something’ with Psillos’s ‘core causal description’ (of something) and, consequently, to suspect that Stanford’s critique is equally damaging to both of these proposals. But this impression is misleading, as can be seen by attending to another important aspect of Chakravartty’s position: his insistence that scientific realism is ‘first and foremost a realism about well-confirmed *properties*’ (2008, 155). On this view, continuity between successive theories is ensured not by their referring to the same entities, but the same properties. The difference between

the two accounts is quite subtle, because Psillos (1999, 293), too, emphasizes continuity at the level of properties. However, he identifies them as ‘properties by virtue of which [the posited entities] play their ascribed causal role’. This invites Stanford’s complaint that we may only be able to single out which properties these are with the benefit of hindsight. Taking again the above example, we nowadays know that the electromagnetic field can play its causal role without being a material substance, but as we have seen, Maxwell explicitly denied that this could be true of the ether, and Psillos does not give us a criterion on the basis of which Maxwell and his contemporaries could have thought otherwise.<sup>5</sup> By contrast, Chakravartty’s identification of the properties to which realists should be committed makes no reference to the entities supposed to possess these properties. Realism, Chakravartty (2007, 47) argues, is primarily concerned with ‘causal properties one has managed to detect’ (‘detection properties’). On the other hand, realists should remain agnostic with regard to ‘any other putative properties attributed to particulars by theories’, which he calls ‘auxiliary properties’. The question of how to demarcate detection properties from auxiliary properties will be discussed in the next subsection. For the moment, I just note that one does not need to rely on retrospective judgement to perform that task. For example, I will argue that even at the time of Maxwell, physicists could have classified the amplitudes or frequencies of (what they took to be) ether waves as detection properties, while recognizing the ether’s materiality as an auxiliary property. If this is so, then Stanford’s argument against our ability to reliably fix the core causal descriptions of our theoretical entities does not translate into an argument against our ability to tell detection properties from auxiliary properties.

This refinement of realism is also effective against a second objection raised by Stanford to the realist’s use of causal reasoning. Having criticized Philip Kitcher’s (1993, sec. 5.4; 2001, sec. 5) distinction between ‘working posits’ and ‘presuppositional posits’, Stanford (2006, sec. 7.2) considers a possible improvement of this distinction on behalf of the realist, to the effect that realists might want to commit themselves to the existence of only those posits to which theories ascribe direct causal roles, treating other posits as merely presuppositional (or ‘idle’). He takes a dim view of this idea:

Although perfectly natural, this suggestion seems to run afoul of any number of discarded theoretical posits that *were* ascribed direct causal roles in the production of phenomena by the successful explanatory practices of their respective theories. To such familiar examples as phlogiston and caloric fluid we’ve seen that we might fairly add Darwin’s gemmules, Galton’s stirp, and Weismann’s biophors. (Stanford 2006, 172)

A realist who has taken Chakravartty’s lesson to heart need not be embarrassed by Stanford’s list of theoretical entities which were once believed to play a causal role but were subsequently abandoned. As we have seen above, there may well be discontinuity on the level of entities, but continuity on the level of properties. Thus, a sophisticated realist does not need to (implausibly) deny that, for example, past theories indeed ascribed a direct causal role to caloric. And he can even admit that Hasok Chang (2003) has refuted Psillos’s (1999, 130) claim that ‘caloric’ was not a central term (see also Stanford 2006, 175-179). But he can maintain that the properties which were ascribed to caloric and which we now regard as mistaken were

---

<sup>5</sup> When Psillos (1999, 139) discusses Maxwell’s ‘differentiated attitude towards the several parts of a theory in view of what evidence supports them’, he may have in mind something similar to the criterion I will discuss in the following. But his failure to state it explicitly (or to provide any other criterion) invites Stanford’s charge of rationalization *post hoc*.

just auxiliary properties, while a case can be made for the retention of some detection properties which appeared in the caloric theory, such as the specific heat of air ( $c_p$  and  $c_v$ ).

Chakravartty's focus on properties instead of entities might make him vulnerable to another one of Stanford's arguments. This new strategy, the objection goes, secures nothing more than a pyrrhic victory for the realist, since a realism which forfeits its commitment to theoretical entities, settling for the mere commitment to (some) properties, does no longer deserve to be called realism. However, the fact that properties are the primary focus of the realist's response to (NI) does not imply that his commitment is limited to properties alone. Once the reality of certain properties is established, the realist can argue for the existence of entities, based on the fact that properties often cohere to form interesting units (Chakravartty 2007, 63-66). In some cases, the reality of some entities follows even more directly, as will be seen for the atomic hypothesis in section 4.

### 3.2 Causal realism and the detection/auxiliary distinction

The analysis so far has shown that Stanford's arguments against the realist's appeal to causal knowledge do not, by themselves, threaten Chakravartty's proposal. On the other hand, it has also shown that the viability of this new causal strategy depends crucially on our ability to tell detection properties from auxiliary properties 'at the time a theory is a going concern', as Stanford puts it. This necessitates a closer look at Chakravartty's (2007, 47-54) account of the demarcation between detection and auxiliary properties. Here is the gist of his suggestion:

Detection properties are connected via causal processes to our instruments and other means of detection. One generally describes these processes in terms of mathematical equations that are or can be interpreted as describing the relations of properties. As I will attempt to show, one can thus identify detection properties as those that are required to give a *minimal interpretation* of these sorts of equations. (48)

But, as discussed at the end of subsection 2.2, the problem with the notion of 'minimal interpretation' is that the corresponding commitment depends crucially on which reading of 'minimal' is applied. Without further qualification, Chakravartty's (2007, 48) advice to treat as auxiliary anything that 'goes beyond what is minimally required to do the work of science: to make predictions, retrodictions, and so on' just results in instrumentalism. To be of any use for the realist, 'minimal interpretation' thus needs to be read as 'minimal realist interpretation'. Chakravartty makes this reading explicit in his discussion of Fresnel's equations: 'The existence of certain properties is minimally required to give a realist interpretation of these equations' (49). Unfortunately, Chakravartty does not offer an explication of the appropriate (realist) understanding of minimality that does not itself depend on having already identified some properties as indispensable. This makes his claim that 'one can [. . .] identify detection properties as those that are required to give a minimal interpretation' somewhat misleading: It is not the case that we can approach the relevant equations equipped with a well-understood notion of 'minimal interpretation' in order to learn which of the properties described by them are detection properties. Rather, it seems to be the other way round: Only once we know which properties are detection properties are we in a position to give an appropriate minimal (realist) interpretation of the equations in question.

This is not to say that Chakravartty's characterization of detection properties is unilluminating, but its success does not depend on identifying them by means of a minimal interpretation. In fact, Chakravartty seems to elucidate both concepts ('detection property' and 'minimal interpretation') with the help of a prior notion of



‘causal connection’ or ‘causal contact’. Not only does he (as seen in the above quotation), characterize detection properties as the ones which are causally connected to our instruments, but he also summarizes his ‘recipe of the minimal interpretation’ as a commitment ‘only to structures with which one has forged some significant causal contact’ (54). Of course, this only pushes the problem one step back. In order to demarcate detection properties from auxiliary properties, we now have to know what counts as significant causal contact.

At this point, Chakravartty’s account can profit from some other recent work in the development of scientific realism. What I have in mind is the distinction between causal and theoretical warrant, introduced by Mauricio Suárez (2008) and further developed by myself (2012) in the context of what I call ‘causal realism’. More precisely, I suggest that this distinction adequately captures and explicates the intuitions behind Chakravartty’s detection/auxiliary distinction. By providing such an explication, causal realism turns Chakravartty’s (2008) sketch into a robust strategy against Stanford’s NI.

On my account of the distinction between causal and theoretical warrant, the former is generated by an inference to the best explanation if and only if the corresponding explanation fulfils three criteria, denoted by ‘non-redundancy’, ‘material inference’, and ‘empirical adequacy’ (Egg 2012, 261). It is the second of these criteria which gives us a first hint at how causal warrant connects to the concept of a detection property: Material (as opposed to formal) inference is defined as the kind of inference ‘that results in ascribing to a concrete entity a property for which there is a well-defined notion of what it means to modify it’ (266). Having such a well-defined notion is crucial for establishing causal contact with a hypothesized property  $P$ , because in order to claim that some observed phenomenon causally depends on the presence of  $P$ , one must be able to specify what would have happened if  $P$  had not been present. The properties mentioned in the definition of material inference can therefore legitimately be called ‘detectable properties’ (265n), since they are the properties with which we can in principle establish some causal contact. But not every detectable property is a detection property. In other words, the possibility of establishing causal contact is not always actualized, since the criterion of material inference is not alone sufficient for causal warrant, but only in conjunction with the other two criteria, empirical adequacy and non-redundancy. If a hypothesis postulating a (detectable) property  $P$  turns out to be empirically inadequate, that is, if it implies incorrect statements about observable things and events, then there is obviously something wrong with the claim that  $P$  is connected via causal processes to our instruments and other means of detection. In that case,  $P$  is not a detection property and the corresponding hypothesis lacks causal warrant (267). Similarly, a hypothesis lacks causal warrant if it does not fulfil the requirement of non-redundancy, that is, if there are other (scientifically respectable) hypotheses capable of explaining the relevant phenomena (261-263).<sup>6</sup> A property  $P$  postulated by such a hypothesis is not a detection property, because if the phenomena can also be

---

<sup>6</sup> The foregoing considerations concerning the ‘minimal interpretation’ may have raised the worry that the criterion of non-redundancy can never actually be fulfilled, because there is always the alternative to choose an instrumentalist interpretation of equations instead of a realist one, and both options allow us to make the same predictions, retrodictions etc. However, this is a kind of redundancy that is not a proper subject for the debate on scientific realism, because it is just as damaging to common sense realism as to scientific realism (see Stanford’s (2006, 13) comments on the ‘Craigian reduction’ of a theory). The restriction to ‘scientifically respectable’ hypotheses in the definition of non-redundancy aims to exclude this kind of cases.

explained by an alternative hypothesis which does not make reference to  $P$ , then observing these phenomena does not amount to making causal contact with  $P$ .

In sum, relying on the characterization of causal warrant in terms of the three above-mentioned criteria provides us with a straightforward way to distinguish detection properties from auxiliary properties: The former are simply those for which we have causal warrant. This matches very well with Chakravartty's account: Material inference spells out a crucial precondition for what he calls 'forging significant causal contact', empirical adequacy captures his idea that detection properties allow us to make predictions, retrodictions and so on, and non-redundancy expresses the belief that these properties are indispensable to performing these tasks.

However, Stanford's challenge creates a problem for this characterization, because it threatens the applicability of the criterion of non-redundancy. To take the PUA seriously is to acknowledge that even if a certain hypothesis seems to be the only one which can account for some phenomena, we may not be justified in asserting that this is really so. As a consequence, if non-redundancy were understood in the strong sense of there being no other empirically adequate hypothesis at all (whether known or unknown), then we could never claim any causal warrant in cases in which the PUA is a matter of concern, because we could never justifiably claim that the requirement of non-redundancy has been met. Therefore, if causal realism is to do any work in solving the PUA, non-redundancy has to be understood in the weaker sense of there being no other known hypothesis that accounts for the phenomena. Distinguishing these two senses of non-redundancy allows us to reformulate Chakravartty's claim that our knowledge of detection properties is immune to the PUA in the following way: Hypotheses that are causally warranted, which is to say that they conform to the criteria of non-redundancy (in the weak sense), material inference and empirical adequacy, are likely to be non-redundant in the strong sense.

To see how this explication improves on Chakravartty's proposal, let us return to the example of ether waves introduced in the previous subsection. The challenge is to show how one can recognize the amplitude of such a wave as a detection property and the substantiality of the corresponding medium as an auxiliary property, without yet knowing about the former's retention and the latter's rejection in the subsequent theory. Working with Chakravartty's definition, it is not easy to see how this could be done, because there is a sense in which the ether's substantiality is just as much 'connected via causal processes to our instruments' as the amplitudes of ether waves. More precisely, if one presupposes Maxwell's belief (quoted above) that no energy can be transferred without a substantial medium, then we actually detect the ether's substantiality whenever we detect the energy of an electromagnetic wave. The reason why the ether's substantiality should (unlike the amplitude of an ether wave) still not count as a detection property becomes clear once we turn to the concepts of causal realism, in particular the criterion of material inference. The criterion is certainly fulfilled in the case of a light wave's amplitude, because there is a well-defined notion of what it means to modify that property. There are well-known experimental procedures for this, and even in cases where experimental intervention is not possible in practice, it is perfectly clear what it would mean to intervene on the amplitude of an electromagnetic wave. By contrast, no such procedures or well-defined notions exist in the case of the ether's substantiality. Even if one accepts Maxwell's presupposition connecting substantiality to the capacity for energy transfer, it is not clear what it would mean to intervene on that property, and hence, the corresponding hypothesis does not meet the requirement of material inference. If, on the other hand, one insists that there is a well-defined notion of modifying substantiality independently of the capacity for energy transfer (after all, we now know that energy transfer can occur with or without a substantial medium), then the possibility of an electromagnetic field without a substantial ether is no longer an unconceived

alternative but a known one, and hence, non-redundancy fails even in the weak sense mentioned above. In either case, it follows that there is no causal warrant for the ether's substantiality, which therefore fails to be a detection property.

This shows that, by drawing on the conceptual resources of causal realism, the distinction between detection properties and auxiliary properties can be made sufficiently precise to underlie a realism with the prospect of withstanding Stanford's argument from unconceived alternatives. The extent to which this prospect is realized will depend on how well this account fits the historical record of scientific reasoning, hence the need for detailed case studies as the one presented in the next section. What has been shown so far is that none of the arguments by which Stanford attacked earlier versions of realism are effective against the causal strategy described in this section, such that there is (until further notice) no reason to worry about unconceived alternatives to causally warranted claims. Furthermore, it seems unlikely that Stanford's arguments can be modified so as to demonstrate causal realism's vulnerability to the PUA. The reason for this is that causal warrant is closely associated with causal explanations of experimental phenomena, rather than with fundamental theorizing. Now admittedly, it would not be correct to say that unconceived alternatives are only a problem for fundamental theorizing, not for explaining experimental results. Indeed, experimenters regularly grapple with questions such as the following: Have we taken into account all external influences on the apparatus? Did nothing go wrong in the calibration process? Is the signal more than just an artefact of our sophisticated data analysis? In dealing with these questions, it sometimes happens that relevant alternative explanations of the phenomena are overlooked. But the historical record of these unconceived alternatives does not possess the right kind of structure to support Stanford's NI. Unconceived alternatives appear here and there in the history of scientific experimentation, but there is nothing like the systematic pattern described in (RTU), and accordingly, there is no analogue to (NI) which could call into question the scientists' ability to reliably identify the causes of experimental phenomena.

## **4 Causal Realism, Unconceived Alternatives, and the Atomic Hypothesis**

The previous section ended with a claim about the history of experimental science, and such a claim can of course be disputed on historical grounds. In fact, Stanford (2009b) can be interpreted as doing just that. Discussing Jean Perrin's experimental work on Brownian motion and the associated case for the existence of atoms and molecules, Stanford argues that Perrin's experimental genius did not prevent him from being subject to the PUA. The present section will address this objection. Analyzing the case of Perrin will also allow me to connect the rather abstract considerations of the previous section with actual scientific practice. In particular, this historical example will illustrate three central elements of the causal realist's response to the PUA: the distinction between detection and auxiliary properties, the difference between fundamental theorizing and the causal explanation of phenomena, and the inferential step from detection properties to the reality of unobservable entities.

### **4.1 Perrin and the philosophers: some initial observations**

Brownian motion is an irregular movement of microscopic particles suspended in a fluid, discovered by the botanist Robert Brown in 1827. Perrin was by no means the first to interpret this phenomenon as an effect of the constant movement of the molecules which constitute the fluid, but there is a wide consensus that the unprecedented accuracy of his experiments on Brownian motion, carried out between 1908 and 1911, played a crucial role in establishing that atoms and molecules actually

exist.<sup>7</sup> However, there is considerable (and ongoing) disagreement about why Perrin's work played such a crucial role. The proposed answer which concerns us here is the one given by Roush (2005), but before discussing this approach, some remarks about other philosophical treatments of this historical episode will help to set the stage. The philosophical literature on Perrin's work is so comprehensive that I will not try to give a complete summary here. I merely wish to discuss the part of the debate which is most closely related to causal realism, namely the arguments surrounding Nancy Cartwright's (1983, 82–85) and Wesley Salmon's (1984, 213–227) reconstruction of Perrin's reasoning in terms of a causal inference. The two accounts are not identical (Cartwright sees Perrin's argument as an instance of a more general inference scheme, namely 'inference to the most likely cause', while Salmon focuses on the 'common cause' structure of Perrin's argument), but they are closely enough related to be treated together in the present context. The central idea in both treatments is that the atomic hypothesis causally explains a wide variety of very different physical phenomena and that it would be an incredible coincidence if it managed to do so without being true. Even though this approach has been widely criticized, it still seems essentially correct to me. I will therefore briefly address the different lines of criticism that have been directed against it.

Deborah Mayo (1986) argues that Cartwright misses an important element when she takes Perrin to infer the reality of molecules from the remarkable agreement in the estimates for Avogadro's number based on entirely different physical phenomena. Before making such an 'argument from coincidence', Perrin had to check what Mayo (1986, 49) calls 'the experiment's *internal validity*', that is, the adequacy of the model underlying the empirical predictions tested in one particular experiment, regardless of its applicability to other phenomena. For the study of Brownian motion, this meant to demonstrate that the movement of a Brownian particle was indeed a random process, not depending on the movement of neighbouring particles. This is the same element of Perrin's argumentative chain that Roush (2005) focuses on, and my discussion of Roush's position will show that causal realism can readily accommodate this element. Indeed, it seems that Mayo herself does not actually consider Cartwright's analysis as mistaken, but merely as incomplete (see also Mayo 1996, 216–217).

A second line of criticism is that the causal inference which Cartwright and Salmon take to be the essence of Perrin's reasoning can only assure *that* the phenomena in question have a (common) cause, but not what that cause is (Achinstein 2001, 250;<sup>8</sup> Psillos 2011, 358n14). Stanford makes a similar remark in the context of Roush's proposal, and my discussion of it in subsection 4.4 will show that causal realism provides a substantial (though not exhaustive) characterization of what the cause of Brownian motion is.

Finally, it has been argued that Cartwright and Salmon cannot explain why Perrin's work was such a decisive step in convincing the scientific community of the reality of atoms, since the results they take to be central seem to have been available well before Perrin took on the issue (Maddy 2007, 404; Psillos 2011, 358n14). As mentioned above, I accept that the argument from coincidence, which Cartwright and Salmon emphasize, needs to be complemented by Mayo's insight about Perrin's demonstration of the randomness of Brownian motion. Therefore, I concede Maddy's and Psillos's point to the extent that the Cartwright-Salmon account is not quite sufficient to fully appreciate the importance of Perrin's work. On the other hand, the

---

<sup>7</sup> It is not immediately clear whether Stanford (2009b) departs from this consensus or merely criticizes one particular way of characterizing Perrin's achievement. I will return to this question in subsection 4.4. For explicit criticism of the consensus, see van Fraassen (2009), but also the response by Chalmers (2011).

<sup>8</sup> Interestingly, this criticism does not appear in Achinstein's (2002) paper dedicated to Salmon's argument.

alternative accounts put forth by Maddy and Psillos strike me as even less convincing explanations of Perrin's historical importance. Maddy (2007, 406) observes that 'before Perrin's successes, the case for the existence of atoms had hinged on aggregate behavior' and claims that 'Perrin's accomplishment was to establish a link to [the] behavior of individual molecules'. But this last claim is not true in any but a very loose sense, given that a Brownian particle undergoes some  $10^{21}$  collisions with molecules of the surrounding fluid per second (Chandrasekhar 1943, 23). Psillos (2011, sec. 3.3) reconstructs Perrin's reasoning by means of a probabilistic argument, but it remains unclear why Perrin's experiments should have been crucial for establishing any of the argument's three premises. The first premise states that the likelihood ratio  $P(n=N|\neg AH)/P(n=N|AH)$  (where  $AH$  stands for the atomic hypothesis and  $n=N$  for the claim that the number of molecules in a mole is equal to Avogadro's number  $N$ ) is very small. This follows from Perrin's theoretical model and does not depend on any of his experimental results. By contrast, the second premise does report an experimental result, namely that  $n=N$  is the case. But if, as Psillos claims, 'most of the ways to calculate Avogadro's number were known (and they were in agreement) before Perrin brought them together in his books' (358n14), then establishing this premise was not a new achievement. The same is true for the third premise, stating that  $P(AH)$  is 'not very low'. According to Psillos, Perrin evaluated this prior probability 'by eliminating several alternative potential explanations of Brownian movement' (357), but clearly, others had done this before him (Nye 1972, 21-28).

## 4.2 Roush and Stanford on Perrin

Building upon Mayo's ideas, Roush takes the decisive element of Perrin's research to be his demonstration that Brownian motion is completely random, exhibiting 'no systematic effects, no dependencies or correlations between the motions of one particle and another or tendencies in the motion of a single particle'. Roush goes on to claim that, by demonstrating the complete randomness of Brownian motion, Perrin confirmed what she calls 'the modest atomic hypothesis', that is, the hypothesis 'that there are atoms and molecules, understood merely as spatially discrete sub-microscopic entities moving independently of each other, i.e., at random' (Roush 2005, 219). This is the claim that Stanford attacks in his (2009b) paper.

The disagreement between Roush and Stanford is illuminating, because, from the perspective of causal realism, both of them are partly right, but also partly wrong about this case. Let us see why. Unsurprisingly, Stanford takes issue with Roush's (2005, 219) claim that 'there do not seem to be any hypotheses that could explain a random walk in the Brownian particles that are not included within this [modest] atomic hypothesis'. He cites two hypotheses as counterexamples, taken from Fine (1991, 91): Instead of being caused by molecular collisions, Brownian motion could be due to electrostatic forces among the particles themselves (in conjunction with exchange forces with the medium) or it could be uncaused, a manifestation of some fundamental randomness in nature. I do not think that the first example really supports Stanford's argument, because any causal influence of electrostatic forces among the Brownian particles on their movement would result in some correlation between the movements of nearby particles, thus contradicting Perrin's conclusion that Brownian motion is fully random.<sup>9</sup> Since Stanford does not dispute that Perrin

---

<sup>9</sup> Stanford (2009b, 260) anticipates this kind of argument and admits that 'the bare appeal to "electrostatic forces" does not immediately or straightforwardly entail random motion of the Brownian particles'. He nevertheless deems such an entailment possible, depending on 'the characteristics of the electrostatic forces and their interaction with the exchange medium'. The idea seems to be that the

indeed established this conclusion, Fine's first example is of no use to him. By contrast, Fine's second example does support Stanford's argument by accounting for complete randomness in Brownian motion without relying on the atomic hypothesis. This shows that Roush is wrong when she equates Perrin's demonstration of randomness with a confirmation of the atomic hypothesis.<sup>10</sup>

Using the terminology introduced above, we can summarize the argumentative situation up to this point by saying that the randomness of Brownian motion became, thanks to Perrin, a detection property. And while I agree with Stanford that this does not yet establish the reality of atoms, I also agree with Roush that it was a significant achievement. Drawing on the resources of causal realism, it is even possible to explain precisely how Perrin managed to turn randomness from an auxiliary into a detection property. Among the three necessary criteria for such a transition (see subsection 3.2 above), non-redundancy is the least relevant to this case, for two reasons: First, the logical space of possibilities is rather simple here, 'since the motion is either random or it is not' (Roush 2005, 219; quoted approvingly in Stanford 2009b, 257). Second, and more importantly, considerations of non-redundancy cannot even start before there is a clear notion of what it would mean for the system to have one or the other of the alternative properties. In other words, it needs to be established that the randomness hypothesis meets the criterion of material inference. Perrin did this by deriving precise quantitative expectations about the behaviour of Brownian particles from Einstein's (1905) theory of Brownian motion, which depended crucially on the supposition that the motion is completely irregular (Mayo 1996, sec. 7.4). The comparison of these predictions with the experimental results then showed that the randomness hypothesis also satisfied the criterion of empirical adequacy, whereas its negation was incompatible with experimental data. Taken together, these findings endow the hypothesis with causal warrant, which amounts to saying that randomness is a detection property.

### 4.3 From Brownian motion to the reality of atoms

Hence, the causal realist and Roush agree that detecting the randomness of Brownian motion was an important step in Perrin's argument. Furthermore, due to the simplicity of the logical space of possibilities (either random or not), not even Stanford views

---

electrostatic forces interact with the medium in such a way that they cancel out any correlation between the movements of nearby particles. However, the plausibility of such a model is undermined by Svedberg's experiments of 1907, which showed that the electrostatic properties of the medium had no influence on the Brownian movement of the particles (Nye 1972, 124-125).

<sup>10</sup> To be fair, Roush does not completely equate the two claims. She considers the possibility that something other than the (modest) atomic hypothesis could explain a random distribution of the motion of Brownian particles, but dismisses this possibility on the basis of the following two-stage argument: First, she argues that even if we (generously) assign a prior probability of 0.5 to the totality of these unknown alternatives, we still end up with a posterior probability for the atomic hypothesis of  $\geq 0.6$ , 'better than more likely than not'. But surely a realist wants to claim significantly more than that. So Roush adds that the atomic hypothesis also explains several other phenomena, including those Perrin used to determine Avogadro's number. These pieces of evidence, Roush concludes, 'increase the probability that there are atoms above the minimum probability I have argued for here' (Roush 2005, 221). Hence, Roush can only infer the reality of atoms from the randomness of Brownian motion by relying on the traditional argument from coincidence. I think this is the right way to go (see below), but it undermines her claim that measurement of Avogadro's number is not needed to confirm the modest atomic hypothesis (218–219).

this part of Perrin's claim as threatened by the possibility of unconceived alternatives. But how do we get from here to a robust (PUA-proof) realism about atoms? As a first step, let us take a more detailed look at Fine's second hypothesis mentioned above, which I shall call the no-cause hypothesis. Stanford (2009b, 261) admits that it is not easy to see how the appeal to fundamental, uncaused randomness could plausibly recapture Perrin's explanation for the vertical distribution of Brownian particles at equilibrium. Perrin (1910, 530) described this distribution quantitatively by means of the following equation:

$$\frac{2}{3}W \log \frac{n_0}{n} = \varphi(\Delta - \delta)gh,$$

where  $n_0$  and  $n$  denote the concentration of Brownian particles at two levels separated by a height  $h$ ,  $W$  is the mean (translational) kinetic energy of the particles,  $\varphi$  their volume,  $\Delta$  their density and  $\delta$  the density of the fluid in which they are suspended. Now there is a sense in which Stanford here concedes more to the realist than he needs to. As Alan Chalmers (2011, 721) emphasizes, Perrin was able to derive the above formula assuming only random motion of the Brownian particles, Newtonian mechanics and some elementary statistics, but nothing about molecules or the kinetic theory. In this sense, the no-cause hypothesis can explain the (experimentally observed) exponential decrease of particle concentration as a function of height. What it cannot explain is the rate of this decrease, since it does not tell us anything about the particle energy  $W$ . Nor can  $W$  be measured, because the velocity of the Brownian particles changes much too quickly to permit a reliable measurement (Perrin 1910, 528-529). In the absence of a constraint on  $W$ , there is no constraint on the value of  $n_0/n$  either, so the no-cause hypothesis would also be compatible with a uniform vertical distribution of particles ( $n_0/n \rightarrow 1$ ) or with all the particles sinking to the bottom of the vessel ( $n_0/n \rightarrow \infty$ ).<sup>11</sup>

Assessing the shortcomings of the no-cause hypothesis in this way allows us to spell out precisely what we need to assume about the cause of Brownian motion in order to arrive at a satisfactory explanation of the experimental results. Following the analysis by Chalmers (2011, sec. 5), I suggest that not all parts of the kinetic theory are needed, but only two of its central elements. The first one is a special case of the equipartition theorem, namely the assumption that, in equilibrium, the molecules constituting the liquid in which the Brownian particles are suspended have the same mean (translational) kinetic energy  $w$  as the Brownian particles themselves:  $w=W$ . This does not yet constrain the value of  $W$ , until the energy of a molecule is linked to macroscopically accessible magnitudes. The second assumption does this by postulating a specific value for Avogadro's number  $N$ , which connects  $w$  to the temperature  $T$  and the universal gas constant  $R$  via the equation  $w=3RT/2N$  (Perrin 1910, 517). If  $N=6 \cdot 10^{23} \text{ mol}^{-1}$  is assumed, the calculated value for  $n_0/n$  agrees with what is experimentally measured. Conversely, measurements of  $n_0/n$  (and  $\varphi$ ,  $\Delta$  etc.) can now be used to determine the value of  $N$  with high accuracy.

Of course, invoking the atomic hypothesis in the form of the two assumptions described here is not the only possible way to explain the vertical distribution of Brownian particles. Experimentation is a tricky business, and there is always the possibility that what the experimenter observes is not an effect of the natural processes he intended to study, but simply an artefact of the experimental setup. This is where the *argument from coincidence*, as emphasized by Cartwright and Salmon

---

<sup>11</sup> This argument occupies a central position in Perrin (1910, 554), as Chalmers (2011, 726) points out.

(but also by Chalmers 2011, 724-726), becomes crucial. If studying the vertical distribution of Brownian particles were the only way to establish a link between Avogadro's number and observable phenomena, there would be room for scepticism about the atomic hypothesis. But this is not the case. In fact, there are several different phenomena from which we can calculate  $N$ . It would be an incredible coincidence if each of these observations were an artefact, and yet all agreed about the value of  $N$  (Cartwright 1983, 84).

The argument from coincidence thus allows us to rule out the artefact hypothesis as a reasonable explanation of the observed phenomena, just as the no-cause hypothesis was ruled out by the previously discussed argument. But this does not yet address Stanford's worry about hitherto unconceived hypotheses. In order to do this, it remains to be shown that the atomic hypothesis belongs to a class of claims which is not threatened by the PUA, because it does not provide the basis for an induction of the form (NI). According to causal realism, this requires showing that the atomic hypothesis enjoys causal warrant. The foregoing discussion has already shown that it fulfils the criteria of empirical adequacy and (weak) non-redundancy, since it accurately predicts what is experimentally observed, while the known alternative hypotheses are unable to do so. The assessment with respect to material inference is less straightforward, because it is not immediately clear whether there is a sufficiently well-defined notion of what it means to modify the properties to which the atomic hypothesis refers. In particular, while it is easy to see what it means to modify the numerical value of  $N$ , one needs to specify what this means in terms of physical properties; a possible world in which the number of molecules in one mole of substance is  $N'$  instead of  $N$  may be seen as a world which has exactly the same physics as our world, but in which people simply use another definition of 'one mole'. In order to relate a change in  $N$  to a real change of physical properties, one needs to keep the definition of the mole fixed while modifying the value of  $N$ . This change then corresponds to a change in physical properties, namely the masses of all atoms, such that the number of atoms in 12 grams of pure carbon-12 is no longer  $6 \cdot 10^{23}$ , but something else.<sup>12</sup> Of course, there is no way in which we could actually bring about such a change, but practicability is not part of what the criterion of material inference requires. It is sufficient that the relevant change is well-defined, and this is the case here. In sum, the atomic hypothesis meets all the criteria for causal warrant.

The fact that the experiments on Brownian motion do not merely confirm some predictions of the kinetic theory in general, but lend causal warrant to some specific assumptions, is a crucial achievement of Perrin's work (for a similar assessment, though not formulated in terms of causal warrant, see Chalmers 2011, 723). More specifically, it provides us with a concrete historical example of the contrast between fundamental theorizing and explaining experimental phenomena, which I emphasized at the end of section 3; the kinetic theory as a whole is a product of the former activity, but when we manage to direct the successes of its predictions to some specific assumptions, we engage in the latter. Since, by Stanford's own admission, his NI applies only to fundamental theorizing, not to explaining experimental phenomena, there is no reason to expect any unconceived alternatives to the atomic hypothesis when it comes to explaining the results of Perrin's experiments. And since

---

<sup>12</sup> There is an additional subtlety here, because the definition of the mole, in virtue of the phrase '12 grams', implicitly refers to atomic masses, if 'gram' is defined in terms of the mass of a material body (as it still is according to the International System of Units). 'Keeping the definition of the mole fixed' therefore requires that mass be defined in an independent way, e.g. in terms of Planck's constant.



the known alternatives (notably the no-cause and the artefact hypothesis) were found wanting, the case for the reality of atoms and molecules is complete.

#### 4.4 What we know about atoms

But even if the truth of the atomic hypothesis is now secured, the above attempt to dissociate it from (parts of) the kinetic theory might raise the worry that the hypothesis thereby becomes uninformative. Stanford expresses this worry concerning Roush's modest atomic hypothesis, when she dissociates it from the idea of exact localization, 'to accommodate the possibility, later discovered, that atoms exhibit the quantum mechanical property of not being fully localized' (Roush 2005, 219). The subtleties of quantum non-locality do not matter here, what matters is just that an idea which at one time was associated with the atomic hypothesis (namely exact localization) later came to be separated from it, so we might think that the two ideas should always have been kept separate. Stanford (2009b, 262) comments:

Such retrospective retelling of the story threatens to treat the modest atomic hypothesis simply as a placeholder or a bare name for whatever further inquiry ultimately decides about the causes of the phenomena that occasioned its introduction.

But we have already seen that this criticism has no force against my account of the atomic hypothesis. Firstly, it should be clear by now that the distinction between what does and does not belong to a hypothesis, on the basis of what is and is not causally warranted, can be drawn without the benefit of hindsight, so there is nothing problematic about 'retrospective retelling of the story'. Secondly, the atomic hypothesis, as construed above, is much more than a placeholder; it makes substantial claims about molecules, most notably, how many of them there are in a certain amount of substance. Finally, since  $N$  connects macroscopic with molecular magnitudes, the atomic hypothesis directly yields further information about molecules, for example concerning their masses or their approximate dimensions. This is not to say that the atomic hypothesis gives us exhaustive knowledge about atoms and molecules, but it definitely gives us more than a bare name.

It is now time to finally address a question which I have already hinted at in footnote 7 above: Does Stanford really claim that we have insufficient evidence for the reality of atoms, or does he just argue that Roush's account is insufficient to support such a realism? His (2009b) paper can be interpreted as defending only the second claim, with which I obviously agree. However, the following passage (taken from a different paper) indicates that he also embraces the first claim:

Are atoms and amoebae really on epistemological equal footing? Although we can point to a glowing blue dot in a suitably prepared photograph and say "see, that's an atom," virtually all of what we think we know about atoms comes from the role they play in a highly elaborate fundamental theory we have adopted because its empirical accomplishments are so much more impressive than those of any competing account we know of concerning the fine structure of matter. But quite a lot of what we know about amoebae (how fast they move, what they eat, how often they reproduce, etc.) does not come to us in this way, but in a variety of other ways by means of which we routinely gather knowledge about the world around us (even if this knowledge is also ultimately "theoretical" in character). (Stanford 2009a, 389)

This suggests that, although Stanford advocates realism about amoebae (which distinguishes him from a constructive empiricist), he remains sceptical about atoms. I have argued in this section that taking into account Perrin's complete argument (rather than just the part Roush focuses on) undermines such a sceptical position. But

even if we suppose that this argument fails and that Perrin did not succeed in rendering the atomic hypothesis immune to the threat of unconceived alternatives, I doubt that it is, a century after Perrin, still possible to coherently oppose realism about atoms on the basis of the PUA. The reason for this is that we today possess evidence for atoms which is much more obviously immune to the PUA than the kind of evidence Perrin could produce. Consider, for example, the famous experiment in which Donald Eigler and Erhard Schweizer (1990) used 35 xenon atoms to write ‘IBM’ on a nickel surface. The point is not that we ‘see’ these atoms ‘in a suitably prepared photograph’, but that our ability to produce such pictures testifies to our knowledge about atoms (how they behave when placed on a metallic surface, how they can be moved from one place on the surface to another, etc.).<sup>13</sup> And this knowledge does not come from a fundamental theory, but from what is experienced in the laboratory. In fact, Eigler and Schweizer (1990, 525) admit that they have only an incomplete theoretical understanding of the interaction between the xenon atoms and the nickel surface or between the xenon atoms and the tip of their scanning tunnelling microscope, but this did not stop them from knowing how to position the atoms. To put it another way, the claim that Eigler and Schweizer really manipulated single atoms and that these really have the properties which make such manipulation possible does not depend on any eliminative inference. Rather, these claims are of the same kind as the ones Stanford cites with regard to amoebae, for which the PUA has no relevancy whatsoever.

## 5 Conclusion

Stanford’s PUA poses a serious challenge to scientific realism, a challenge which has not yet been fully answered by realists. In particular, different versions of selective realism, in their attempt to show that certain parts or aspects of scientific theories are immune to the PUA, have not succeeded in reliably characterizing these parts, or they managed to do so only with the benefit of hindsight.

This paper has sought to develop a more successful type of selective realism. The starting point of my strategy was Chakravartty’s suggestion that causal knowledge is a kind of scientific knowledge which is not threatened by the PUA. I showed that this suggestion does not fall prey to Stanford’s criticism of similar earlier realist proposals, if Chakravartty’s distinction between detection properties and auxiliary properties can be explicated in a sufficiently precise way. This I attempted to do by bringing in the distinction between causal and theoretical warrant advocated by Suárez and myself.

Since the criteria defining causal warrant are rather demanding, one might worry that not many scientific claims will meet this high epistemic standard. Consequently, the causal realism which derives from this strategy might seem so meagre as to be barely distinguishable from Stanford’s epistemic instrumentalism. The case study on Perrin furnishes at least the beginning of an answer to that worry, by displaying a clear case of how the causal realist’s commitment goes beyond what Stanford would allow. Furthermore, this case illustrates how the concepts and distinctions on which my strategy rests can be applied to concrete scientific practice. More work is needed to show that causal realism also yields substantive commitments in other cases,<sup>14</sup> but at the very least, the present paper demonstrates that the limits of warranted realist belief are not as narrow as Stanford supposes.

---

<sup>13</sup> This is obviously an instance of Hacking’s (1983) argument from manipulability. Several authors have pointed out shortcomings of Hacking’s general position, but these do not prevent the argument from working perfectly well in the case discussed here.

<sup>14</sup> For one more example, see my (2012, sec. 4) case study on the neutrino.

**Acknowledgments** I am thankful to the audience of the 14<sup>th</sup> Congress of Logic, Methodology and Philosophy of Science (Nancy, July 2011), where I presented a very early version of this paper. Later versions profited greatly from discussions with Anjan Chakravartty and Kyle Stanford. These took place during my visit to the University of Notre Dame, for which I acknowledge support by the ‘fondation du 450<sup>ème</sup> anniversaire de l’université de Lausanne’. I am also indebted to three anonymous referees for many helpful comments.

## References

- Achinstein, P. (2001). *The Book of Evidence*. New York: Oxford University Press.
- Achinstein, P. (2002). Is there a valid experimental argument for scientific realism? *The Journal of Philosophy* 99, 470–495.
- Cartwright, N. (1983). *How The Laws of Physics Lie*. Oxford: Clarendon Press.
- Chakravartty, A. (2007). *A Metaphysics for Scientific Realism: Knowing the Unobservable*. New York: Cambridge University Press.
- Chakravartty, A. (2008). What you don’t know can’t hurt you: Realism and the unconceived. *Philosophical Studies* 137, 149–158.
- Chalmers, A. (2011). Drawing philosophical lessons from Perrin’s experiments on Brownian motion: A response to van Fraassen. *British Journal for the Philosophy of Science* 62, 711–732.
- Chandrasekhar, S. (1943). Stochastic problems in physics and astronomy. *Reviews of Modern Physics* 15, 1–89.
- Chang, H. (2003). Preservative realism and its discontents: Revisiting caloric. *Philosophy of Science* 70, 902–912.
- Devitt, M. (2011). Are unconceived alternatives a problem for scientific realism? *Journal for General Philosophy of Science* 42, 285–293.
- Egg, M. (2012). Causal warrant for realism about particle physics. *Journal for General Philosophy of Science* 43, 259–280.
- Eigler, D. M. and E. K. Schweizer (1990). Positioning single atoms with a scanning tunnelling microscope. *Nature* 344, 524–526.
- Einstein, A. (1905). Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Annalen der Physik* 17, 549–560.
- Enfield, P. (2008). Review of *Exceeding Our Grasp*. *British Journal for the Philosophy of Science* 59, 881–895.
- Fahrbach, L. (2011). How the growth of science ends theory change. *Synthese* 180, 139–155.
- Fine, A. (1991). Piecemeal realism. *Philosophical Studies* 61, 79–96.
- French, S. (2011). Metaphysical underdetermination: Why worry? *Synthese* 180, 205–221.
- Hacking, I. (1983). *Representing and Intervening: Introductory Topics in the Philosophy of Natural Science*. Cambridge: Cambridge University Press.
- Harker, D. (2013). How to split a theory: Defending selective realism and convergence without proximity. *British Journal for the Philosophy of Science* 64, 79–106.
- Kitcher, P. (1993). *The Advancement of Science*. New York: Oxford University Press.
- Kitcher, P. (2001). Real realism: The Galilean strategy. *The Philosophical Review* 110, 151–197.
- Laudan, L. (1981). A confutation of convergent realism. *Philosophy of Science* 48, 19–49.
- Maddy, P. (2007). *Second Philosophy*. New York: Oxford University Press.
- Magnus, P. D. (2006). What’s new about the new induction? *Synthese* 148, 295–301.

- Magnus, P. D. (2010). Inductions, red herrings, and the best explanation for the mixed record of science. *British Journal for the Philosophy of Science* 61, 803–819.
- Maxwell, J. C. ([1873] 1955). *A Treatise on Electricity and Magnetism*, Volume II. London: Oxford University Press.
- Mayo, D. G. (1986). Cartwright, causality, and coincidence. In *PSA 1986: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, Volume 1, pp. 42–58.
- Mayo, D. G. (1996). *Error and the Growth of Experimental Knowledge*. Chicago: University of Chicago Press.
- Nye, M. J. (1972). *Molecular Reality: A Perspective on the Scientific Work of Jean Perrin*. London: Macdonald.
- Perrin, J. (1910). *Brownian Movement and Molecular Reality*. London: Taylor and Francis. Reprinted in M. J. Nye (1986): *The Question of the Atom*, Los Angeles: Tomash Publishers, pp. 507–601.
- Psillos, S. (1999). *Scientific Realism: How Science Tracks Truth*. London: Routledge.
- Psillos, S. (2009). *Knowing the Structure of Nature: Essays on Realism and Explanation*. Basingstoke: Palgrave Macmillan.
- Psillos, S. (2011). Moving molecules above the scientific horizon: On Perrin’s case for realism. *Journal for General Philosophy of Science* 42, 339–363.
- Roush, S. (2005). *Tracking Truth: Knowledge, Evidence, and Science*. New York: Oxford University Press.
- Roush, S. (2010). Optimism about the pessimistic induction. In P. D. Magnus and J. Busch (Eds.), *New Waves in Philosophy of Science*, pp. 29–58. Basingstoke: Palgrave Macmillan.
- Ruhmkorff, S. (2011). Difficulties for the problem of unconceived alternatives. *Philosophy of Science* 78, 875–886.
- Saatsi, J. (2009). Review of *Exceeding Our Grasp*. *Metascience* 18, 355–363.
- Salmon, W. C. (1984). *Scientific Explanation and the Causal Structure of the World*. Princeton: Princeton University Press.
- Stanford, P. K. (2006). *Exceeding Our Grasp: Science, History, and the Problem of Unconceived Alternatives*. New York: Oxford University Press.
- Stanford, P. K. (2009a). Author’s response. *Metascience* 18, 379–390.
- Stanford, P. K. (2009b). Scientific realism, the atomic theory, and the catch-all hypothesis: Can we test fundamental theories against all serious alternatives? *British Journal for the Philosophy of Science* 60, 253–269.
- Suárez, M. (2008). Experimental realism reconsidered: How inference to the most likely cause might be sound. In S. Hartmann, C. Hofer, and L. Bovens (Eds.), *Nancy Cartwright’s Philosophy of Science*, pp. 137–163. New York: Routledge.
- van Fraassen, B. C. (2009). The perils of Perrin, in the hands of philosophers. *Philosophical Studies* 143, 5–24.
- Votsis, I. (2007). Review of *Exceeding Our Grasp*. *International Studies in the Philosophy of Science* 21, 103–117.
- Votsis, I. (2011). The prospective stance in realism. *Philosophy of Science* 78, 1223–1234.