Inductions, red herrings, and the best explanation for the mixed record of science

P.D. Magnus*

December 22, 2009

This is the author's final draft of a paper forthcoming *The British Journal for the Philosophy of Science.*
e-mail: `pmagnus(at)fecundity.com`
web: `http://www.fecundity.com/job`

Abstract

Kyle Stanford has recently claimed to offer a new challenge to scientific realism. Taking his inspiration from the familiar Pessimistic Induction (PI), Stanford proposes a New Induction (NI). Contra Anjan Chakravartty's suggestion that the NI is a "red herring", I argue that it reveals something deep and important about science. The Problem of Unconceived Alternatives, which lies at the heart of the NI, yields a richer anti-realism than the PI. It explains why science falls short when it falls short, and so it might figure in the most coherent account of scientific practice. However, this best account will be antirealist in some respects and about some theories. It will not be a sweeping antirealism about all or most of science.

Kyle Stanford [Sta01][Sta06] claims to mount a new challenge to scientific realism. Taking his inspiration from the familiar Pessimistic Induction (PI), Stanford proposes a New Induction (NI). Anjan Chakravartty [Cha08, p. 153] has recently called the NI a "red herring", arguing that it poses no new problem for realism. Patrick Enfield suggests similarly that "Stanford might deny it, but… what really does the work in this argument are 'historical' claims associated with the pessimistic induction" [Enf08, p. 883]. I argue that this charge is unjustified. Moreover, the Problem of Unconceived Alternatives — which lies at the heart of the NI — yields a richer anti-realism than the PI. It explains why science falls short when it falls short, and so it might figure in the most coherent

*This paper grew out of a seminar at the University at Albany, SUNY. I want to thank all the participants, especially John Milanese and Elizabeth Gray. Thanks to Anjan Chakravartty, Gregory Frost-Arnold, Kyle Stanford, and anonymous referees for feedback on earlier drafts.

account of scientific practice. This antirealist element in our best explanatory account of science may ultimately be galling to realists, but it won't be all wine and roses for antirealists either.

In §1, I review the familiar Pessimistic Induction. In §2–§3, I discuss how best to understand Stanford's New Induction. In §4, I consider the objection that the NI is a red herring. I argue that it is not. In §5, I suggest that the challenge Stanford mounts to realism leads to a richer account of science than previous arguments like the PI.

1 The Pessimistic Induction

The Pessimistic Induction (PI) relies on the history of science in an attempt to establish anti-realism in this way:

PI-1 The historical record reveals that past theories which were successful typically turned out to have been false.

PI-2 So, our present scientific theories, although successful, will probably turn out to be false.

PI-3 Therefore, we should not believe our best present theories.

In the voluminous literature replying to the PI, realists have tried to block the argument in a number of ways. Several kinds of objection are worth distinguishing.

First, even admitting that there is a pattern in the historical record, generalizing from the pattern might be fallacious. Various statistical considerations have been used to try and block the move from history to PI-1. Lewis [Lew01] has argued that it commits a base rate fallacy. Lange [Lan02] has argued that it commits what he calls the turnover fallacy. To be sure, generalizing from history is a tricky business.

Second, realists try to block the inference from PI-1 to PI-2 by insisting that present theories are better in some respect than abandoned past theories. If this difference can be made explicit, so much the better — but even in the absence of an explicit description, a realist can still claim that there is *some* difference. Stanford acknowledges the possibility of this reply and accepts it *arguendo* [Sta06, p. 10–11], but he does not seem to give it much weight; after answering the third sort of reply (discussed below), he considers the PI established. Nevertheless, this realist reply need not be hopelessly *ad hoc*. One might consider this to be part of a *realist package*, in the sense of Boyd [Boy90]. The idea behind the package reply is this: The anti-realist may match the realist on each individual point, with each individual contention coming to a stalemate. Nevertheless, realism does better as an overall view — it provides a more coherent explanatory account of science. Part of the realist account is that present theories are better than past theories, and so we have some reason to believe that they are.

Third, realists say that the parts of the theories that got the world right were the same parts that were responsible for the theories' success. These parts have not been rejected, but instead continue as parts of present theories. Just as we still believe in the working parts of past theories, we should now believe in the working parts of our present theories. Arguments of this kind are given by Kitcher [Kit93] and Psillos [Psi99], among others. The approach admits that past theories were literally false, while insisting that they were nevertheless partly correct. Following Chang [Cha03], I'll call this approach *preservative realism*.

Stanford calls this an appeal to "partial confirmation" and replies to it at length [Sta06, chs. 5–6]. We might be able to tell such a story retrospectively, he argues, but it should come as no surprise that the parts of the successful past theories *that we now think were the core of them* were the parts *that we now think are true*. What is judged to have been useful is what is judged to have been true, but the standards by which we judge what was efficacious are the same standards by which we judge what was true. Moreover, preservative realism only allows the realist to say after the fact which parts of a theory were doing the work. That is, the realist can only retrospectively identify which parts of a superseded theory were likely to be true and thus were worthy of belief. It is not enough for scientific realism as a general doctrine that we can, in some special instances, divide the confirmed from the unconfirmed parts of present theories. A general doctrine requires a general criterion. Stanford does not claim *a priori* that such a forward-looking criterion is impossible, but notes that we do not have such a criterion and do not have any prospects for finding one. If Stanford is right, then the realist cannot justify belief in our present theories — we will only be able to say at some point in the future what it is that we would now have been justified in believing today. Stanford calls this a "pyrrhic victory" for realism. They win the preservative battle at the cost of their realist souls.

2 Stanford's New Induction

Stanford poses a problem which he means to be a hybrid of the PI and the problem of underdetermination. He dubs it the Problem of Unconceived Alternatives, and he means to establish it by a New Induction (NI). In this section, I briefly summarize the NI.

Scientists typically cannot address all the relevant theoretical alternatives. There are theories which never explicitly occur to them as live possibilities, but which later scientists take seriously. Yet scientists often rely on eliminative inference; that is, they rule out all but one of the possibilities and conclude that whatever remains must be right. (On the importance of eliminative inference to science, see *inter alia* Kitcher [Kit93, ch. 7] and Bird [Bir06].)

Of course, eliminative induction is sometimes justified. It is an "old axiom," as Sherlock Holmes says, "that when all other contingencies fail, whatever remains, however improbable, must be the truth" [Doy05, p. 926]. To take a mystery-story example, suppose we know that one of five suspects in the se-

cluded manor must have committed the murder. If evidence exonerates four of them, then the fifth must be the killer. It is always logically possible to question the exhaustion premise in such a case, but there might not be any alternatives worth taking seriously in a mundane case like this. The requisite premise is contingent and known only fallibly, but we can know it in some cases. Stanford will admit as much, and he is not concerned to challenge *every* use of eliminative inference [Sta06, p. 29 ff.].

In scientific cases concerning distant regions or exotic features of the world, however, Stanford argues that we cannot be confident that we have considered all the possibilities. The idea is illustrated in a fanciful way by Kurt Vonnegut, who imagines meeting the late Isaac Newton during a near-death experience. Vonnegut writes:

> It isn't enough for Newton that during his eighty-five years on Earth he invented calculus, codified and quantified the laws of gravity, motion, and optics, and designed the first reflecting telescope. He can't forgive himself for having left it to Darwin to come up with the theory of evolution, to Pasteur to come up with the germ theory, and to Albert Einstein to come up with relativity. "I must have been deaf, dumb, and blind not to have come up with those myself," he said to me. "What could have been more obvious?" [Von99, pp. 51–2]

Whatever Vonnegut might imagine, later theories were not obvious to earlier scientists — and future theories are not obvious to us. As such, we would be unjustified in believing theories about esoteric domains — even theories that are superior to any other theory we can presently think up. There might and indeed *are likely to be* significant unconceived alternatives in science. Stanford claims that these worries undercut our "fundamental theories", by which he means our theories about "the constitution of matter... the remote history of the Earth... the most minute working of our bodies... and the structure of the farthest reaches of the universe..." [Sta06, p. 32]. Note that this use of 'fundamental theory' is pretty broad, including parts of biology and geohistory along with microphysics.

The argument is something like van Fraassen's complaint against inference to the best explanation that the best explanation we can think of might still be pretty bad, if it is the best of a bad lot. Whereas van Fraassen only appeals to this as a possibility, Stanford tries to provide a positive reason to think that scientists actually have overlooked serious possibilities; cf. [Sta06, pp. 47–8, fn. 8]. In order to show that the Problem of Unconceived Alternatives plagues actual science, Stanford appeals to the historical record: "[T]he historical record suggests that in science we are typically unable to exhaust the space of likely, plausible, or reasonable candidate theoretical explanations for a given set of phenomena before proceeding to eliminate all but a single contender, but this is just what would be required for such eliminative inferences to be reliable" [Sta06, p. 29]. This New Induction is aimed to establish antirealism in this way:

NI-1 The historical record reveals that past scientists typically failed to conceive of alternatives to their favorite, then-successful theories.

NI-2 So, present scientists fail to conceive of alternatives to their favorite, now-successful theories.

NI-3 Therefore, we should not believe our best present theories insofar as they are the result of eliminative inference.

I have deliberately presented the PI and the NI so as to highlight their parallel structure. Each begins with an historical basis: a pattern of failures. An inductive inference carries us to the present case, from which an antirealist conclusion follows.

Yet the two differ importantly in their details. They rely on different historical patterns, and the basis for each induction is different; the PI is an induction over *scientific theories*, while the NI is an induction over *scientists*. Even though lack of historical evidence might make it hard to suss out facts about past scientists, there is some fact of the matter as to how many people there were between (for example) 1800 and 1850. It is unclear whether there is any comparable fact of the matter about the number of theories that were considered in the same period. How do we individuate theories so as to count them? and what level of attention counts as consideration? Scientists are at least a well-defined sample space in a way that theories are not.

Also, the PI yields antirealism about all our present theories, while the NI only yields antirealism about present theories that are established in a specific way. The PI is usually used to argue for antirealism which marks the limits of what should be believed at the boundary between the observable and the unobservable. The NI instead suggests an antirealism which marks the limits at the boundary between conclusions drawn without eliminative inference and conclusions relying on eliminative inference. This is still antirealism — a view which argues that some of our best scientific theories are not even approximately true — but the focus has shifted. As Stanford explains:

> A number of recently influential challenges to scientific realism have embodied an important turn away from the idea that the limits of observability mark a crucial epistemic boundary distinguishing justifiable from unjustifiable scientific beliefs. In its place, it has been suggested that the scope of justifiable scientific belief is limited instead by a concern for our ability to effectively consider the space of alternatives to a hypothesis we seek to evaluate and to consider their empirical consequences. [Sta09, pp. 253–4]

3 Reaching and overreaching

The NI and the Problem of Unconceived Alternatives, as I have presented them, pose *retail arguments* for antirealism (arguments that raise doubts about scientific claims of a specific, targeted kind) rather than *wholesale arguments* (ar-

guments yielding a generic conclusion about the theoretical claims of science). The distinction is admittedly a matter of degree, but the NI as I have so far presented it is toward the retail end of the continuum.[1] It applies only to scientific claims that are the result of eliminative inference, and not even to all of those. It applies only if scientists do not have good reasons for thinking that all serious possibilities are among the ones that have been considered. Although this is more general than arguments just about a single theory, it is a far cry from familiar realist or antirealist arguments that are meant to apply in a sweeping way to our mature sciences. The two labels are helpful for constrasting two readings of the New Induction, because it is possible to reconstruct it as a wholesale argument.[2]

Stanford sometimes seems to have aspirations of a wholesale argument. In his original presentation of the NI [Sta01], he presents it without any special connection to eliminative inference. It goes roughly like this: *In the past, scientists failed to consider theoretical possibilities that were just as well-confirmed as their favorite, conceived theories. So present scientists are probably in the same situation, and so our best present theories are no better confirmed than other possibilities that we have not yet imagined. Therefore, we ought not believe our best present theories.*

The unrestricted conclusion about our best present theories makes this a wholesale argument. If successful, it delivers antirealism about all science in one shipment.

Yet, in refining the NI, Stanford explicitly cautions that "the limited skepticism... motivated [by the NI] should certainly not extend to every scientific claim or hypothesis and may even have different force as applied to scientific exploration in different domains" [Sta06, p. 37]. This caution makes the NI a retail argument, depending on the details to such a degree that it will not yield antirealism about all science, most science, most mature science, or any such sweeping category. The crucial matters are whether the claims are obtained by eliminative inference and (if they are) whether scientists actually exhausted the relevant possibilities. However, just a few sentences after explaining this "limited skepticism," Stanford claims that the NI poses "a distinctive general challenge to *virtually all* those fundamental theories concerning remote domains of nature that lie at the heart of the contemporary scientific conception of the natural world" [Sta06, p. 37, my emphasis]. Note again the broad use of 'fundamental.' Further passages could be provided to support the retail or wholesale reading, and I think there is a genuine tension. In a recent paper, he argues for an "expansive conception of the scope of the problem" [Sta09, p. 256]. Although Stanford explains the NI as a retail argument when he is being careful, he has wholesale aspirations.

[1] 'Scientific realism' is sometimes defined modestly as being the claim that there are *some* unobservable entities in which one could be justified in believing. Given this definition, realist success with even a single retail argument would be enough to establish realism. Once we are distinguishing retail arguments from wholesale arguments, however, we should also distinguish possible positions between monolithic ANTIREALISM and REALISM.

[2] I take the distinction from work co-authored with Craig Callender [MC04].

Regardless, I argue that any wholesale version of the NI overreaches. After this section, the remainder of the paper will be addressed to the retail version with which I began.

The intermediate conclusion of the wholesale argument is that there are unconceived rivals to our best present theories. Even if future evidence will eventually favor the unconceived, that does not show that present evidence does not favor the conceived. Our best present theories might still be better confirmed by present evidence than the unconceived rivals, and then we would be justified in believing them. In order to undercut realism, the argument requires that the conceived and the unconceived are now *equally well-confirmed*.[3] Since this is an induction, this further requires that (in past scientific episodes) unconceived alternatives were not merely *compatible* with then-present evidence, but were equally well-confirmed by it.

Consider the case of classical mechanics and special relativity in 1785.[4] All of the observations made at that time were of objects moving slowly enough that relativistic effects were negligible, and it would have been impossible for a physicist then to show that the Galilean transformations rather than the Lorentz transformations applied. However, this does not show that the two theories were equally well-confirmed. Given scientific standards of plausibility at that time, the Lorentz transformations would just have been a sceptical scenario. It does not matter that observations a century later would make relativity a serious possibility.

Stanford may reply in this way: The fact that physicists did ultimately accept special relativity shows that it was a scientifically serious possibility, so scientists in 1785 would have been wrong to just set it aside. Indeed, he writes, a theory must "be regarded as a scientifically serious alternative possibility rather than a mere skeptical fantasy [if] it is ultimately accepted by some actual scientific community" [Sta06, pp. 20–21]. Yet this argument only makes sense if being a scientifically serious alternative is a timeless property of a theory. To see that it is not, consider possibilities that we think we can set aside. For example, it is common in science fiction stories for characters to discover that they have been living in a computer simulation. They unplug from the machine, sign up with a crew in the real world, and learn to reprogram the simulation. If we were to have experiences like these, then we would come to accept that we had heretofore lived in a simulation. Without such experiences, the possibility of living in a simulation is just a sceptical fantasy. As Stanford himself insists, any reasonable method will allow us to set it aside. As such, the fact that brain-in-a-vat scenarios are not serious possibilities is not a timeless *a priori* fact; it depends on whether we have had experiences like those in the sci-fi stories or not. Just as we can set them aside, then, scientists in 1785 could have set aside effects depending on the ratio between an object's velocity and the speed of light (had special relativity been proposed).

One might still insist that emerging from a simulation would show that

[3] One might think that this would not even be enough. If two well-confirmed theories were equally well-confirmed (one might say) then one would be justified in believing either one.

[4] The example roughly follows my [Mag06].

brain-in-a-vat scenarios had always been serious — even though we could not have known it. If we must now take seriously what we *might* ultimately have to take seriously in the future — even allowing for futures that would subvert a great deal of what we now think we know — then we have to take all sceptical possibilities seriously. This is just to give up the distinction (which Stanford accepts) between serious and sceptical possibilities, and that would just be to give up all knowledge knowledge claims whatsoever.

So the standards for scientific plausibility are relative to background knowledge and so (effectively) time dependent. *Prima facie* this is no help for the realist. If our present best theories only appear warranted relative to today's standards but those standards may change tomorrow, then we would be hasty to actually believe our present best theories. However — as numerous realists have argued — realism is compatible with the historical-variability of standards. Standards depend on the background assumptions of scientists at a time. As theories improve, so do the assumptions and the methods. Improving methods support further improvement in theories. According to realists, theories that are closer to the truth allow for methods that are more robust and reliable; more reliable methods allow for more accurate theories. Assessments of plausibility and seriousness are part of the standards at a time. This dialectic between theory and method is part of the realist package.[5]

An antirealist might still insist: The time-dependence of plausibility judgments makes science at any particular time unreliable. It is not merely that certain options are dismissed as sceptical scenarios, but further that some dismissed possibilities are ultimately preferable to the theories that are taken seriously.

Yet the realist will say that past methods nonetheless yielded *approximately* true theories. These were not as close to the truth as later theories, but the methods were not as refined as later methods.

Perhaps the antirealist can push back, but we cannot press too far without giving up the NI as a distinct argument. The complaint cannot merely be that theories believed in the past were ultimately repudiated in favor of theories that had not been believed; that's just the PI. At the same time, the complaint needs to allow for present scientists to rationally reject sceptical scenarios like Cartesian demons.

The situation here is similar to Lipton's discussion of what he calls *the problem of underconsideration* [Lip96]. The problem arises from granting that scientists can reliably judge the relative merits of different theories but insisting that scientists might not have even imagined a true theory. This alleged to yield an intermediate scepticism: The comparative judgement that one theory is the best of those considered is insufficient to justify belief in that theory. The eliminative inferences that worry Stanford fit just this pattern: We reliably conclude that several possibilities are false but fail to even consider the true account.

[5] See Boyd [Boy85][Boy90] and (although he does not use the 'package' metaphor) Kitcher [Kit02]. For the treatment of plausibility judgments as part of background knowledge, see also Salmon [Sal90].

Lipton argues that this scepticism is untenable. Principles of inductive inference are necessarily contingent and empirical, which means both that they must be learned and that they may be improved as we learn more about the world. As a result, the admission that scientists can make even comparative judgements between theories requires that their background theories be approximately true. One might give up the premise that scientists can reliably compare theories, but then the worry just becomes familiar Humean despair about ampliative inference. The problem of underconsideration fails to provide a distinct, successful argument for antirealism. As Lipton puts it, "The intermixture of methodological principle and substantive belief, in part a consequence of the essential role of background belief in theory evaluation, makes it unclear how even to formulate the global argument..." [Lip96, p. 104]. The further point here, in relation to Stanford's NI, is just that plausibility judgements are one component of scientists' inductive practice. Since many realists embrace the interdependence of theory and method, there is no reason for them to be worried or emabarrassed by the interdependence of theory and judgements of plausibility.

Understood as a wholesale argument, the NI distills down to some other familiar sceptical problem. So it is better to understand the NI as a retail argument, applying only to scientific theories of a limited kind.

4 The red herring objection

This section begins by considering how a preservative realist might answer the NI. Stanford offers a counter reply, but it is equally well a defense of the PI. Since the retail NI's conclusion is weaker than the PI's, this threatens to make the NI redundant. I argue that this worry can be answered.

The preservative realist answer to the PI, were it successful, could just as easily be given to the NI: Past theories were not the full truth, but they did get some things right. Past scientists failed to conceive of theories which would later supersede the theories they could imagine, but the better theories agree with the working core of the past theories. So scientists were justified in believing in the approximate truth of their best theories. And we are justified in believing in the approximate truth of ours.

As we have seen, Stanford has a reply to this realist maneuver: We can only determine which parts of their theories past scientists ought to have believed from our vantage point in their future. They would have been well-advised to be cautious antirealists. So, he concludes, "the most... sophisticated recent defenses of realism from the historical record do not constitute an effective reply to the original *pessimistic induction*..." [Sta06, p. 184, my emphasis]. Further down the same page, he insists the upshot of this discussion is "that the challenge to our scientific beliefs posed by the *problem of unconceived alternatives* is quite genuine, and that the most promising proposals concerning how we might circumscribe the problem... simply will not work" [Sta06, p. 184, my emphasis]. So Stanford's pyrrhic victory argument would show that the PI and the NI are both serious threats to realism. Yet PI-3, the conclusion of the PI, is

stronger than NI-3, the conclusion of the NI. The realist will need a response to the PI regardless of the NI. And perhaps a refutation of the PI, like preservative realism, will refute the NI as well. This makes it natural to say, as Chakravarty does, that the NI is "a novel red herring" [Cha08, p. 153]. (See also Enfield [Enf08].) It is a new argument for an antirealist conclusion, but it does not make any new trouble for the realist. A realist needs to answer the PI anyway, and (the objection goes) answering the PI will yield an answer to the NI without any further work. Call this the *red herring objection* to the NI. There are two significant ways that it might be answered. I will state them briefly and then consider each in more detail.

First, preservative realism is not the only realist answer to the PI. There are other ways to block the PI that, even if they were successful, would not undo the NI. The NI is places new demands on the realist and so is not a red herring.

Second, PI-1 enters into the PI as a brute datum; that is, the PI relies on but provides no explanation for the dismal historical record. If we had independent reasons to think that scientists would not exhaust the relevant alternatives, however, then NI-1 would express just the kind of pattern that we would expect. Moreover, NI-1 explains why PI-1 is true (insofar as it is true). Thus, the Problem of Unconceived Alternatives provides an explanation for the bases of both the NI and the PI. And in this explanation, the NI is more fundamental. (So it's not a red herring.)[6]

4.1 The NI is more robust than the PI

Recall the first two responses to the PI, from §1.

First, statistical considerations might undercut the inductive generalization in the PI. There is no way to treat all past theories as a population or pull a representative sample of them. For example, there is no sensible way to enumerate the population of theories that were considered by physicists between 1800 and 1850 — and the problem gets worse when we try to consider theories from history altogether. As Stanford notes [Sta06, p. 44], however, the NI is not a generalization over past theories. Instead, it is a generalization over past *scientists*. It would not be trivial to sample the population of past scientists, but it is at least coherent to treat past scientists as a population. And so the NI does not have the same weaknesses *qua* statistical inference that the PI has got.[7]

Bird [Bir07] resists the inductive step of the PI by denying that the past failures described in PI-1 are a representative sample. There are, he suggests, "a raft of important and long-standing discoveries that have never been falsified" [Bir07, p. 80]. If he is right, then the inference from PI-1 to PI-2 fails. Note, however, that a litany of discoveries which have not been falsified would not provide an answer to the NI. Rather, one would need a raft of cases in which

[6] I want to thank Elizabeth Gray for helping me to appreciate this line of response.

[7] Of course, the NI is also about theories *qua* the favourite theories of past scientists. Yet, for each scientist, we might record their favourite theory. So we don't need to treat the theories themselves as the population which we are attempting to sample.

the range of alternative possibilities had been constant; that is, areas where scientists had not imagined any new possibilities. Perhaps such a raft could be lashed together, but it suffices for my point here that it would be distinct from the raft of cases that Bird relies on in answering the PI. As such, the NI is a novel problem for the realist.

Second, it is possible for the realist to block the PI by insisting that present theories are somehow different than past theories; for example, that present theories are mature or well-tested in a way that past theories were not. To resist the NI, though, the realist must instead insist that present *theorists* are somehow different from past theorists; they must insist that present scientists somehow consider all serious alternatives to their preferred theories even if past scientists did not. Perhaps a realist would argue for such a claim, but in any case it is distinct from the premise they need for their argument against the PI.

Lipton [Lip00] provides a version of this response to the PI. He insists that failure of past theories is *exactly what we would expect* if our theories were getting better, and so PI-1 is entirely compatible with realism. He writes:

> What undermines the inference from past falsity to general incompetence in science is what we know about how science develops. We know that this process is in some ways akin to natural selection, with theories generated and selected. [Lip00, p. 202]

The analogy with natural selection is apt for answering the PI. If theories are developed and refined by a process of selection, then the demise of unfit theories in the past is a reason to expect fitness among our present theories. Yet natural selection can only pick organism types that have arisen as mutations, and theory selection can only pick theories that scientists have imagined. Unless there is a reason to think that theorists have gotten better at exhausting the relevant possibilities — and the analogy gives us no such reason — the NI goes through. So the NI should still worry a realist.

For the realist who answers the PI in these ways, the NI is not a red herring. It is a distinct, novel problem. And it demands an answer.

4.2 The NI is more fundamental

Stanford offers the NI in order to show that the Problem of Unconceived Alternatives actually arises in scientific practice. The Problem of Unconceived Alternatives can be posed as a separate worry; it is distinct from any particular argument that attempts to exploit it and from any antirealist conclusion drawn from it. The Problem of Unconceived Alternatives, formulated independently, amounts to this claim:

UA Scientists typically cannot consider or address all the relevant, serious theoretical alternatives.

UA grounds NI-2, without an appeal to history. As such, if there were an argument for UA that did not rely on an an induction from the history of science,

then that it would underwrite a distinct antirealist argument. Contrawise, the PI is just a straight induction from past failure to probable future failure: Past scientists ought to have been antirealist, so we should be too. There is no underlying thesis that might be motivated or exploited in different ways.

It is plausible to think that humans are not the kind of cognizers who can consider all the relevant alternatives. We can consider a handful of possibilities at the most. As Stanford says, "we are not good at conceptually exhausting spaces of serious alternative possibilities with the sort of amorphous and indefinite contours characteristic of those from which we draw our fundamental theories of nature" [Sta06, p. 45]. He suggests that cognitive science might provide evidence for this claim. If it could, then we would have an independent reason to believe UA.

I want to suggest, further, that *scientists* are especially unlikely to exhaust the space of alternatives. As Kuhn [Kuh70][Kuh77] argues, scientific progress typically occurs in the context of a well-developed theoretical framework — what he famously but ambiguously calls a 'paradigm.' Scientists accept the basic assumptions and methods of the framework, working to extend it in the details. During periods of crisis, when the dominant framework is weighed down by anomalies, some scientists attempt to resolve the problem by developing a novel framework. If the novel framework is successful, the crisis ends with a scientific revolution. The new framework replaces the old one, and ordinary scientists doing normal science work within it until the next crisis.

The Kuhnian picture as I've just summarized it is surely too simple, because parts of the framework can change incrementally over time and different framework commitments can make complicated constellations across the scientific community. However, this basic insight is correct: Scientific development relies on scientists having a shared set of commitments which they are not questioning all at once. When the dominant commitments are no longer adequate for scientific development, scientists go on when they find *some* other framework which is sufficient to support further development. It would be counterproductive for them to wait until *every* possibility had been surveyed, even if that were somehow possible.[8]

My suggestion, then, is that the basic structure of scientific methodolodgy — the way that a successful scientific community works — makes UA true. If this is right, then NI-1 and PI-1 are just the historical patterns we should expect. That is, we should expect past scientists to have relied on workable theories but not to have exhausted the space of possibilities. And we should further expect past scientists to have found theories that were successful enough at the time but ultimately proved to be false. PI-1 is true because of NI-1, which is true because of UA.

[8] A realist may object that by invoking Kuhn I have prejudiced the case against realism. What I have identified as the 'basic insight' is inescapable, but a realist might still say that the blinders of normal science lead scientists to focus on theories that are approximately true. My suggestion here does not rule out that possibility. However, if it could be shown that scientists will generate the true theory then that would constitute a direct solution to the problem of unconceived alternatives.

5 Explanationist antirealism

In recent decades, realists have often relied on inference to the best explanation (IBE); see e.g. Lipton [Lip04]. However, the previous section ended with the suggestion that the historical record described in PI-1 is explained by the historical record described in NI-1, which in turn is explained by UA. The Problem of Unconceived Alternatives thus figures as part of the best explanation for scientific practice and the history of science; it is at least part of the explanation of when and why science *fails* to get the truth. If we accept IBE, then the power of the explanation might be sufficient to establish UA even if the suggestions about cognitive science and methodology are not decisive. Here it is possible for the antirealist to do more than answer the realist in the details. UA and the antirealism it entails contribute to an antirealist package. This package is a rival to the whole realist package, because it promises to offer a coherent account of scientific practice. This is not merely a pyrrhic victory for realists; they are hoisted on their own petard.

We should be careful, however, because it is not yet clear whether the package will be thorough-goingly antirealist. As we saw in §3, Stanford sometimes seems to want a sweeping conclusion that applies to most of science. For example, he writes that "it is quite likely that even the most genuinely impressive and instrumentally accomplished theories of contemporary science will ultimately be replaced by more powerful conceptual tools offering fundamentally different conceptions of nature that have presently not yet even been conceived" [Sta06, p. 211]. Although Stanford himself says that the NI will not engender scepticism about prehistoric or microscopic creatures [Sta06, p. 33], he claims that it will undercut claims that rocks are made up of atoms, that spiders share a common ancestor with humans, that the light of distant stars can be deflected by gravity, that there are gluons, that there are genes, or that there is dark energy [Sta06, p. 210]. He says that these are all connected to 'fundamental theories' (in his sense), but such strident antirealism overreaches his argument. NI-3, the antirealist conclusion underwritten by the NI or by UA, is not unrestricted pessimism. It only calls into question scientific claims that are the conclusion of eliminative inferences, when we are unjustified in believing we have exhausted the serious alternatives.

Stanford [Sta06, chs. 3–5] argues that scientists in the 19th century relied on problematic eliminative inferences in their theorizing about the material basis of heredity. Present arguments for realism about genes might be different, however, and Stanford has shown neither that they are eliminative nor that they are as problematic as their predecessors. To consider other specimens on Stanford's list of allegedly dubious claims:

- The claim that gravity redirects the light from distant stars does not require eliminative inference. Eclipse observations confirm it for our sun, which is a typical star, and no eliminative inference is required for the inference to distant stars. Of course there is *some* inference involved, but it is parallel to the projective inferences that ground our belief in

prehistoric creatures — and Stanford accepts those.

- The claim that humans share common ancestors with spiders is trickier, but it is not readily obvious that it depends on eliminative inference. This is just a specific claim about prehistoric creatures, after all.
- For gluons and dark energy, the case still needs to be made — but surely even avowed realists should be cautious about such exotica. Realism should not try to paint all of science with a uniform coat of belief.

The kind of antirealism we are led to by Stanford's arguments is doubt about the deliverances of a particular kind of argument. Insofar as the theoretical claims of science depend on that kind of argument, we should be cautious. That they do in particular cases, however, will depend on an examination of the details. As I have argued, the NI and the Problem of Unconceived Alternatives are best seen as retail arguments for antirealism. If we accept them, we will be antirealists about present theories that are established by eliminative inference and for which we are unjustified in believing that we have surveyed the relevant alternatives. Yet we cannot say in advance exactly which theories this will sweep out. We will be antirealist about some things and realist about others. The resulting package will neither be wholly realist nor wholly antirealist.

One might hope that case studies will underwrite a generalization, allowing for a statistical wholesale argument. To put it simplistically, if 90% of cases went antirealist, then one might argue that unsurveyed cases have a 90% chance of going antirealist. This is possible in principle, but I am not sanguine about the possibility. It is unlikely that scientific episodes to which NI-3 applies will form a homogeneous reference class. The situations that force scientists to rely on eliminative inference may vary from domain to domain. The constraints that make it impossible for scientists to justifiably think they have eliminated all other serious possibilities are surely even more varied. (Espescially since, as I argued in §3, what counts as a serious possibility depends on historical context.) So case studies will probably not allow for any interesting, high-level generalization.

Stanford counsels "doubt that much interesting can be said that would apply to all and only 'the claims of science' in general" [Sta06, p. 32]. I doubt furthermore that anything can be said about all 'fundamental theories' in Stanford's loose sense of the term. If I am right, the NI does secure a place for antirealism in our best account of science— but it will be a focussed antirealism that must prove its credentials in specific cases. The question is not whether we should be realists or antirealists, but when and about which things we should be realists or antirealists.

References

[Bir06] Alexander Bird. Abductive knowledge and Holmesian inference. In Tamar Szabo Gendler and John Hawthorne, editors, *Oxford Studies in Epistemology*, volume 1, pages 1–31. Oxford University Press, 2006.

[Bir07] Alexander Bird. What is scientific progress? *Noûs*, 41(1):64–89, 2007.

[Boy85] Richard [N.] Boyd. Lex Orandi est Lex Credandi. In Paul M. Churchland and Clifford A. Hooker, editors, *Images of Science*, pages 3–34. University of Chicago Press, 1985.

[Boy90] Richard [N.] Boyd. Realism, approximate truth, and philosophical method. In Savage [Sav90], pages 355–391.

[Cha03] Hasok Chang. Preservative realism and its discontents: Revisiting caloric. *Philosophy of Science*, 70:902–912, December 2003.

[Cha08] Anjan Chakravartty. What you don't know can't hurt you: Realism and the unconceived. *Philosophical Studies*, 137:149–158, 2008.

[Doy05] Arthur Conan Doyle. The adventure of the Bruce-Partington plans. In *The Complete Sherlock Holmes*, volume 2, pages 913–931. Doubleday& Company, Garden City, New York, 1905.

[Enf08] Patrick Enfield. Review of P. Kyle Stanford's *Exceeding Our Grasp*. *The British Journal for the Philosophy of Science*, 59(4):881–895, December 2008.

[Kit93] Philip Kitcher. *The Advancement of Science*. Oxford University Press, 1993.

[Kit02] Philip Kitcher. Scientific knowledge. In Paul K. Moser, editor, *The Oxford Handbook of Epistemology*, pages 385–407. Oxford University Press, 2002.

[Kuh70] Thomas S. Kuhn. *The Structure of Scientific Revolutions*. University of Chicago Press, second edition, 1970.

[Kuh77] Thomas Kuhn. The essential tension: Tradition and innovation in scientific research? In *The Essential Tension*, pages 225–239. University of Chicago Press, 1977.

[Lan02] Marc Lange. Baseball, pessimistic inductions, and turnover fallacy. *Analysis*, 62(4):281–85, October 2002.

[Lew01] Peter Lewis. Why the pessimistic induction is a fallacy. *Synthese*, 129:371–380, 2001.

[Lip96] Peter Lipton. Is the best good enough? In David Papineau, editor, *The Philosophy of Science*, Oxford Readings in Philosophy, pages 93–106. Oxford University Press, 1996.

[Lip00] Peter Lipton. Tracking track records. *The Aristotelean Society, Supplementary Volume*, LXXIV:179–205, 2000.

[Lip04] Peter Lipton. *Inference to the best explanation*. Routledge, London, 2004. second edition.

[Mag06] P.D. Magnus. What's new about the New Induction? *Synthese*, 148(2):295–301, January 2006.

[MC04] P.D. Magnus and Craig Callender. Realist ennui and the base rate fallacy. *Philosophy of Science*, 71(3):320–338, July 2004.

[Psi99] Stathis Psillos. *Scientific Realism: How science tracks the truth*. Routledge, London, 1999.

[Sal90] Wesley C. Salmon. Rationality and objectivity in science or Tom Kuhn meets Tom Bayes. In Savage [Sav90], pages 175–204.

[Sav90] C. Wade Savage, editor. *Minnesota Studies in Philosophy of Science*, volume XIV. University of Minnesota Press, Minneapolis, 1990.

[Sta01] P. Kyle Stanford. Refusing the devil's bargain: What kind of underdetermination should we take seriously? *Philosophy of Science*, 68 (Proceedings):S1–S12, September 2001.

[Sta06] P. Kyle Stanford. *Exceeding Our Grasp: Science, History, and the Problem of Unconceived Alternatives*. Oxford University Press, 2006.

[Sta09] P. Kyle Stanford. Scientific realism, the atomic theory, and the catch-all hypothesis: Can we test fundamental theories against all serious alternatives? *The British Journal for the Philosophy of Science*, 60(2):253–269, June 2009.

[Von99] Kurt Vonnegut. *God Bless You, Dr. Kevorkian*. Washington Square Press, New York, 1999.