Philos Stud (2008) 137:149–158
DOI 10.1007/s11098-007-9173-1

# What you don't know can't hurt you: realism and the unconceived

**Anjan Chakravartty**

Published online: 1 November 2007
© Springer Science+Business Media B.V. 2007

**Abstract** Two of the most potent challenges faced by scientific realism are the underdetermination of theories by data, and the pessimistic induction based on theories previously held to be true, but subsequently acknowledged as false. Recently, Stanford (2006, *Exceeding our grasp: Science, history, and the problem of unconceived alternatives*. Oxford: Oxford University Press) has formulated what he calls the problem of unconceived alternatives: a version of the underdetermination thesis combined with a historical argument of the same form as the pessimistic induction. In this paper, I contend that while Stanford does present a novel anti-realist argument, a successful response to the pessimistic induction would likewise defuse the problem of unconceived alternatives, and that a more selective and sophisticated realism than that which he allows is arguably immune to both concerns.

**Keywords** Realism · Antirealism · Underdetermination · Pessimistic induction · Entity realism · Structural realism · Properties

## 1 Introduction

The title metaphor of Kyle Stanford's *Exceeding Our Grasp* is wonderfully provocative. Anyone with even the slightest familiarity with how scientific research is conducted knows that scientists routinely exceed their grasp, for this is part of the

Presented at a symposium on P. Kyle Stanford's *Exceeding Our Grasp: Science, History, and the Problem of Unconceived Alternatives*, at the Pacific Division APA meeting in San Francisco, 2007.

A. Chakravartty (✉)
Institute for the History and Philosophy of Science and Technology, Victoria College,
University of Toronto, 91 Charles Street West, Toronto, ON, Canada M5S 1K7
e-mail: anjan.chakravartty@utoronto.ca

very process of theorizing and experimentation. We are constantly reaching beyond what we think we know, into realms about which we may know little or nothing—that is part of the heuristic of scientific investigation, and indeed, we could not proceed otherwise. Since we are sadly incapable of magically knowing much if anything about scientifically interesting phenomena in the absence of careful and systematic work, it seems the sciences *must* proceed by means of incorrect and incomplete characterizations of their domains of interest, at least to begin with, and probably a lot longer. This necessity may suggest that we are not always (or perhaps never) in a position to believe such characterizations, and this is where debates concerning the epistemic grasp of the sciences begin.

Stanford's primary goal is to issue a novel sceptical challenge regarding scientific knowledge, and more specifically, to furnish a new argument against scientific realism (roughly, the view that our best theories yield approximately true descriptions of both observable and unobservable aspects of the world), on behalf of antirealism (broadly, the denial of realism regarding the unobservable).[1] The debate between realists and antirealists is longstanding, making novelty something to cherish. It is Stanford's contention that his sceptical challenge—the problem of unconceived alternatives (PUA)—is ultimately telling against realism. It is not merely that in the sciences, we exceed our grasp, but rather that in the philosophy of science, realists exceed their grasp, and it is this latter claim that he hopes to establish.

In this paper I will suggest, however, that while PUA is indeed a new argument, it does not in fact generate any new consequences for the debate between realists and antirealists. For while there is a clear sense in which it is a novel formulation of antirealist scepticism, this turns out not to have any novel import for the central issues contested in the debate itself. As a result, I worry that focusing on PUA serves to distract from precisely these issues. And while Stanford is aware of these issues and has important things to say about them, he does not take us forward in the way that he should, given his careful attention to historical detail. My main contention in this regard will be that we should give realism its due: the view is made into a straw position if one does not take seriously the extent to which it is consistent with the idea that the sciences exceed their grasp. In closing, I will briefly sketch an outline of what a more sophisticated realism might look like, and what further considerations might genuinely take this discussion forward, based in no small measure on what I see as potential awaiting realization in Stanford's incisive approach to assessing scientific knowledge.

## 2 Unconceived alternatives

PUA is intriguing for several reasons. One is that it effectively recasts antirealist concerns about the epistemic status of scientific theories as concerns about the

[1] Reflecting the empiricist bent of influential antirealisms, 'unobservable' here is traditionally applied to anything that cannot be detected by the unaided senses. Stanford (2006, p. 12) prefers to speak of an 'inaccessible domain of nature,' which includes items too small to be seen unaided, but also items that are spatially or temporally distant, rare, or otherwise hard to investigate. The difference is substantive, but will play no material role in the discussion here.

cognitive abilities of scientific theorizers. This raises genuinely interesting questions about the limits of our cognitive capacities and how they are exercised in scientific contexts, which might be helpfully engaged by employing the idea of unconceived alternatives as an analytic probe. While such questions are clearly important to cognitive scientists and anyone interested in scientific or human reasoning more generally, however, I do not believe that worrying about PUA will teach us anything new about the reasonableness (or not) of scientific realism. In order to see why this is so, let me begin by describing the nature of PUA, by considering its relations to prior antirealist arguments.

There are three main, widely discussed foci of antirealist unease concerning realism: scepticism about abduction, or inference to the best explanation (IBE); the underdetermination of theory by data or evidence (UTE); and the pessimistic (meta-)induction (PI). I will leave IBE aside for present purposes, for Stanford has no objection to it as a form of inference per se. The realist faces problems in applying IBE in certain scientific contexts, he maintains, but the source of these problems is not an inherent flaw in IBE as a form of inference. Rather, these difficulties stem from PUA, which has important features in common with UTE and PI. Thus, with the goal of clarifying what sort of argument PUA is in mind, let me focus on these latter two antirealist concerns.

In essence, UTE can be described as a two-step worry. First, there is an assertion to the effect that any given scientific theory has empirically equivalent rivals; and then, there is an assertion to the effect that, given the first assertion, there is reason to doubt that any given theory is true. To coin a phrase, theories are underdetermined by the empirical evidence. The considerations adduced to demonstrate this claim vary across different presentations, and the assertions themselves can be made more or less strongly. For example, one might claim that theories are empirically equivalent and thus underdetermined *in principle*; that is, given all possible evidence. Conversely, one might claim that theories are empirically equivalent and thus underdetermined *in practice*; that is, given the evidence one has at any given time.[2] Many antirealists discuss UTE in principle, no doubt assuming that this would constitute a more powerful reason for scepticism than mere UTE in practice, but Stanford dismisses the former as a form of radical scepticism. His concern is with UTE in practice, or what he (2006, p. 17) calls 'recurrent, transient' UTE: the idea that given any actual theory, $T$, there are other theories that are consistent with or as equally well-confirmed as $T$, on the basis of the evidence available, and thus that we have reason to doubt that $T$ is true. The reason that Stanford is so much more interested in the weaker form of UTE than the stronger, is that in the case of the weaker, he believes that he can adduce historical evidence to demonstrate that its component assertions are very likely true.

[2] Like some in the literature, Stanford reserves the term 'empirical equivalence' for theories that are equivalent with respect to all possible evidence, whereas I (and others) also apply it to theories that are equivalent merely with respect to the available evidence, noting the difference. Following Sklar (1975), Stanford uses the term 'transient underdetermination' in connection with the latter case, but I dislike this usage, for when theories are empirically equivalent with respect to the available but not all possible evidence, there is no guarantee that (the putative problem of) underdetermination is transient, since there is no guarantee that differentiating evidence will always be obtainable, and if it is, that it will be obtained.

Let me now turn to PI, for it is the historical nature of this argument that serves as a model for PUA. In essence, PI can also be described as a two-step worry. First, there is an assertion to the effect that the history of science contains an impressive graveyard of theories that were previously believed, but subsequently judged to be false (for example, because their central terms do not refer, or because their central theoretical descriptions are incorrect). Second, there is an induction on the basis of this assertion, whose conclusion is that current theories are likely future occupants of the same graveyard. PUA mirrors this argument structure closely: first, Stanford asserts that the history of science is typified, at any given time, by a conspicuous absence of theories that are consistent with or as equally well-confirmed (on the basis of the available evidence) as those believed, but that are subsequently conceived and adopted. Second, there is an induction on the basis of this assertion, whose conclusion is that current theories are likely in the same boat as previous ones, in that future science will adopt theories that we have simply not conceived, but that are consistent with or equally well-confirmed by the evidence we currently have. Both PI and PUA are historical inductions, but emphasizing different things: the falsehood of theories, versus an inability on the part of scientists to conceive of alternatives to false theories.

I hope the nature of the challenge presented by PUA is now clear. It is a version of UTE plus a historical induction, leading to a sceptical conclusion about scientific knowledge. That is, it marries a version of UTE to the model of historical demonstration exemplified by PI. And given that this precise formula is not identical to either UTE or PI as traditionally presented, Stanford does succeed in furnishing a novel sceptical argument with which to challenge scientific realism. But given that it is offered in the context of an on-going discussion between realists and antirealists, the immediate question then becomes: what work does this argument do?

## 3 An old dispute re-described

One way to tackle the question of what contribution PUA makes to this debate is to consider what sort of response it requires of realism. What might realists say, faced with this challenge? They might take issue with the historical induction component, and claim that the history of the sciences does not in fact suggest that scientists typically fail to conceive of later-accepted theories, but that would be foolish, since clearly, they do. Alternatively, realists might take issue with the UTE component, and claim that theories are not typically underdetermined by the available evidence. This would be less obviously false, but not (I submit) terribly promising either, for given that generally speaking, later-accepted theories must account for the evidence supporting their predecessors, there is at least a *prima facie* case for the idea that later theories are generally consistent with this evidence, and at least as well confirmed by it as their predecessors. I suspect that neither of these responses will do, and though there may be some things to be said in their defense, let me move on to consider a response that I think is more promising.

The best response to PUA, in my view, is simply to grant PUA. Or less glibly and more correctly, to grant that the phenomenon of unconceived alternatives is a fact of scientific life, but to dispute Stanford's contention that this ultimately spells the death of realism. *Of course* scientists do not typically conceive of all promising alternatives to their own theories, some of which their descendents will propose and accept, and today's scientists are no exceptions. Anyone for whom this is news has clearly not been paying attention. The real question of interest here is whether there is anything like a principled continuity across scientific theories over time, which would allow realists to latch on to certain aspects of theoretical descriptions as likely being approximately true. (By 'principled,' here, I mean that there should be some epistemic principle or principles according to which realists may identify what these retained elements are; I will say more about this later.) Indeed, this is what realists generally think: our best theories today are our best attempts to describe the world thus far, to be replaced with better attempts in due course; in the meantime, we have good reason to believe that certain aspects of today's theories, including those describing the unobservable, are on the right track, and significantly so.

This response to PUA requires spelling out, but before turning to that task, let me first make the point I promised earlier regarding the import of PUA for the debate between realists and antirealists. The response just mentioned on behalf of realism is not, of course, something I have just now invented. It is the sort of response that many and perhaps most realists give *already*, not to PUA, but to PI. If tenable, the notion of principled continuity constitutes a response to PI, because if there is some epistemic principle or principles that would allow the identification of aspects of theories that are likely to be retained, it hardly matters that past theories are generally considered false. The sophisticated realist may *grant* PI—indeed, one might think it silly not to—but nonetheless maintain that the falsity of past and present theories does not preclude knowing that some aspects of these theories pertaining to unobservables are approximately true. Note how this response, which would serve to answer antirealist scepticism in connection with PI, would *also* serve to answer antirealist scepticism in connection with PUA. For again, the fact that we have not conceived of theories that we may adopt in future does not preclude believing that some aspects of current theories pertaining to unobservables are approximately true. So while PUA is a novel argument, the natural response to it is something that realists were already enjoined to provide. It requires nothing new by way of rebuttal. In this sense, PUA is something of a novel red herring.

The issue to which PUA points, as PI did before it, concerns the notion of principled continuity, and whether this notion can be made tenable. Many realists and antirealists have recognized this, and Stanford has important things to say on the subject. Indeed, I take his remarks about continuity—regarding versions of this idea offered by John Worrall (1989), Philip Kitcher (1993), and Stathis Psillos (1999)—to be rather convincing. That is not to say that the conclusions he draws from these remarks are convincing, however. The specific accounts he cites are unsatisfactory, perhaps, but only because they are not fully developed as they stand. There are important ways in which some of these accounts are instructive, I believe, and as it turns out, on the right track. Supplemented with further details, they may well turn out to be rather compelling after all. The details required are there in the realist position, awaiting articulation. Let me give the briefest sketch of a proposed articulation now.[3]

## 4 Selective scepticism/optimism

Stanford identifies two major concerns that arise immediately in response to the realist appeal to continuity. The first is that, by restricting realism to only parts of scientific theories, this strategy represents what he calls a 'pyrrhic victory' for realism. The worry is that by sacrificing other parts of theories to PI or PUA, realists give up too much, and thus find themselves unable to trust the pronouncements of our best theories as they might otherwise hope to do. The second concern about continuity is that there are, he suggests, no good prospective criteria on the basis of which to selectively confirm just those parts of theories that realists should believe. The worry here is one of rationalization post hoc. If one is unable to furnish reasonable grounds for thinking that certain parts of theories concerned with the unobservable are likely to be retained, the identification of retained elements can only be done in retrospect, on the basis of what *has* been retained de facto, rather than by employing an epistemic principle.

I suspect that these problems are potentially even more worrying for the realist than Stanford suggests, so let me aid his cause for a moment before demurring. First, the concern regarding pyrrhic victories is raised in the context of entity realism, whose advocates invoke causal theories of reference to safeguard claims about the existence of certain unobservable entities despite changing theories about them. Thomson, Millikan, Rutherford, Bohr, and so on throughout the twentieth century all referred to the electron, it is maintained, despite having very different theories about the electron. But surely it is at least somewhat misleading to claim that they all believed in the same thing, given that they believed such different things *about* electrons. The worry here is not merely that one's realism is now pyrrhic, but that furthermore, it is trivial—on this view, one is never wrong! Second, the concern regarding prospective criteria is raised in the context of Kitcher's (1993, pp. 140–149) suggestion that some parts of theories are merely 'pre-suppositional,' while others are 'working posits,' and that realists can believe in the latter. Kitcher does not say much about what general criteria might identify working posits; Psillos (1999, chaps. 5 and 6) does, but Stanford argues, not compellingly. The worry here, I suggest, is not so much that realists can only identify things to believe in retrospect, but that more worryingly, they are not entitled to even that much. For given that we are nowhere near (and will likely never arrive at) the end of scientific inquiry, retrospect realists are in no position to know what elements of current theories will be retained there, if any, and thus in no position to endorse what has persisted until now.

Having agreed that the problems Stanford raises are serious, however, let me now contend that sophisticated realists are not so easily refuted. There are three keys, I think, to seeing this point of view. None of them is intended to stand alone, but they

[3] A sketch is all I can provide here, but for an elaboration, see Chakravartty (2007, Part I).

are mutually reinforcing. The first is that realists do in fact have at least one promising criterion for the prospective identification of parts of theories that are likely to be retained in future, *viz.* detailed causal knowledge. (I suspect that there may be other promising criteria, but I will limit myself to this candidate here.) This realization was the outstanding contribution of entity realism: Ian Hacking (1982), Nancy Cartwright (1983, chap. 5), Ronald Giere (1988, chap. 5), and others were right to argue 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. The problem with entity realism is not its proposed criterion with which to identify belief-worthy aspects of theories, but rather its misidentification of what it is, more precisely, that such causal knowledge establishes. Entity realists misidentify what is retained across theory change—a mistake that Stanford inherits, unnoticed—and this leads to the second key.

Realism is too coarse if one conceives it at the level of entities, as Stanford and others do. Scientific realism, I believe, is first and foremost a realism about well-confirmed *properties*, and claims about the existence of various entities must be interpreted, where appropriate, in that light. Did all those theorists about electrons believe in the same entity? That is a hard question to answer, given that each of them associated a range of properties with electrons, some of which are very different. But did they all believe in the property of negative charge? Yes they did. And their belief in that property was sustained by experimental abilities to causally manipulate things having it in highly systematic ways, in virtue of the dispositions that the property of negative charge confers. A knowledge of unobservable properties and relations, I submit, is no pyrrhic victory for the realist; it is substantial. Stanford often writes as though theories and entities are the units of scientific knowledge to which realists must commit, and thus fails to note how well-confirmed properties persist through changes in theories and the entities they describe. A sophisticated realism is finer grained than he allows, and this is why his worries about pyrrhic victories and prospective criteria miss their target.

Once these key points regarding causal knowledge and causally efficacious properties, suitably elaborated, are taken more seriously, it would seem that several other realist insights can be rescued from Stanford's scepticism. Structural realists, for instance, point out that in some fields of scientific research (including many branches of physics), there appears to be a great deal of preservation of mathematical structure across theories over time. Stanford complains that the definition of structure here is vague, and no doubt, in the earliest literature that gave rise to the recent revival of structural realism, that was so. But it is possible to define the term 'structure' very precisely—in terms of relations between properties of the sort I have indicated, for example. Where such relations, described by the mathematical equations of a theory, are susceptible to the causal criterion suggested, the structures they comprise are good bets for surviving historical inductions such as PI and PUA.

Granted, this is somewhat hand waving (see footnote 3), but allow me to wave a bit more. Worrall's flagship case for structural realism is the transition in nineteenth-century theories of light, from Fresnel's wave optics to Maxwell's electromagnetism. Fresnel believed in a luminiferous aether, but Maxwell's theory was ultimately accepted in the context of a non-aetherial physics. As Worrall notes, however, certain mathematical equations (concerning the intensities of incident, reflected, and refracted light at the interface of two media) are endorsed by both. Here it is precisely because properties of light such as intensities and directions of propagation can be systematically manipulated by employing the relevant equations that the relations between properties described by them are likely to survive. And if one were to think that this advice for identifying belief-worthy parts of theories reduces realism to triviality, one would be mistaken. Some realists do incur potential embarrassment here by suggesting that we simply identify the referents of terms like 'dephlogisticated air' and 'oxygen' (thus courting the worry of triviality[4]), but a more sophisticated realism need not. For even if some of the causal roles associated with these putative gases are the same, the properties of dephlogisticated air are clearly not co-extensive with those of oxygen. Oxygen, for instance, has a fixed chemical composition, but *ex hypothesi*, dephlogisticated air does not. Scientific realism, I stress again, is first and foremost a realism about well-confirmed properties.

This leads to a third and final key to a more realistic portrayal of scientific realism. Many realists and their critics write as though a realist commitment to the approximate truth of aspects of theories is an all-or-nothing affair. That is a caricature, however, and while in many circumstances it serves as a harmless idealization, it is seriously misleading here. In actuality, realist commitment to theoretical claims is determined contextually and is a matter of degree; it is not a blanket subscription to glorious truths. In contexts where we possess extraordinarily impressive abilities to systematically manipulate causally efficacious properties, degrees of belief are concomitantly high. In cases where we have less impressive abilities, claims regarding scientific knowledge are inevitably less certain. It is striking, I think, that Stanford's central case studies focus on theories concerning the mechanism of genetic inheritance by Darwin, Galton, and Weismann, that fare quite badly in this respect. In such cases, no sophisticated realist would be especially committed. That is not to say, of course, that explanatory power, the main virtue claimed for these theories, is irrelevant. No doubt there are circumstances in which certain kinds of explanatory power are rightly given more confirmatory weight than others—a thorough investigation of this thesis remains to be undertaken. It is simply a mistake, however, to think that realists are not entitled, like other epistemic agents, to apportion their degrees of belief according to the strength of the evidence.

## 5 Moving forward with (anti-)realism

I have argued that in the debate between realists and antirealists, the probative force of PUA does not extend beyond that of PI, and that in confronting the challenge these arguments present, realism is in a much better position than Stanford contends. Let me now conclude with the surprising suggestion that there may be rather more in common between Stanford's position and the one I have been sketching than one

[4] See Hardin and Rosenberg (1982), for example.

might think. I have claimed that realism has the resources to respond to various criticisms of the sort spawned by PI and PUA, and it is no surprise, perhaps, that I find these resources lying nascent within realism, for as someone interested in seeing whether this position can be made a coherent account of the sciences, I am disposed to being charitable here. The spirit of charity is something that I also see incorporated into Stanford's positive proposal for antirealism, and it is something I believe the participants in this discussion must embrace more generally if it is to move forward.

So then, how best to move forward? Arthur Fine (1986) suggested one way during the first wave of the contemporary debate between realists and antirealists two decades ago. His friend NOA (the 'natural ontological attitude') might be described as a tantalizing combination of quietism and pragmatism. It is quietistic in insisting that realists and antirealists alike should refrain from adding anything further, such as epistemic diagnoses, to the scientific descriptions of the world they both endorse. It is pragmatic in dismissing substantive theories of truth, and in serving as a neutral position for those who see no utility in the realist-antirealist debate. Stanford's approach and my own, however, represent a rather different strategy for moving forward. Stanford ultimately advances a form of instrumentalist antirealism, and my earlier hand waving was offered in support of realism, but interestingly, I believe, our stands on these issues have in common the gemmules (stirps, germ-plasms, what have you) of a general approach: the rejection of realism and antirealism as global epistemic stances towards theories and entities, and the promotion of a more contextual consideration of claims regarding scientific knowledge.

At some points in his discussion, Stanford appears to acknowledge that there are some circumstances in which our epistemic situation is such that even his instrumentalist should believe scientific claims regarding unobservables, and this seems a promising sign. On my view, realism is an approach to scientific knowledge that is selective and graded, reflecting the state of the evidence at any given time, and the systematic uses to which one is capable of putting finely grained descriptions of the relations of specific properties (or not, as the case may be). That is consistent with both very high and very low degrees of belief in different theories, and different aspects of one and the same theory, and there are many circumstances, no doubt, in which the sophisticated realist will have just the same epistemic attitudes towards theoretical descriptions as Stanford's instrumentalist. But this is realism nonetheless, for it recognizes that there are conditions under which it is reasonable to extend belief so far as to embrace the unobservable.

If it turns out that these nuances bring Stanford's instrumentalism and a defensible realism closer together, this would represent a major step forward, not least because it would signal that the labels 'realism' and 'antirealism' no longer pick out the simple-minded opposition that most assume. If there is a lesson to be learned from PUA, it is not that realism faces new challenges per se, but that realists still have work to do in articulating precisely what forms realism takes, and under what circumstances it is a reasonable epistemic attitude to adopt towards the outputs of the sciences. What you don't know can't hurt you; what matters is how we assess what we think we know now. To that end, I am grateful for Stanford's arguments, for reminding us to think carefully about what that could mean.

**Acknowledgements** I would like to thank Arthur Fine, Peter Godfrey-Smith, Sherri Roush, and an excellent audience for a great session, and especially Kyle for the invitation, and his thoughtful and spirited replies.

## References

Cartwright, N. (1983). *How the laws of physics lie*. Oxford: Clarendon Press.

Chakravartty, A. (2007). *A metaphysics for scientific realism: Knowing the unobservable*. Cambridge: Cambridge University Press.

Fine, A. (1986). Unnatural attitudes: Realist and instrumentalist attachments to science. *Mind, 95*, 149–179.

Giere, R. N. (1988). *Explaining science: A cognitive approach*. Chicago: University of Chicago Press.

Hacking, I. (1982). Experimentation and scientific realism. *Philosophical Topics, 13*, 71–87.

Hardin, C. L., & Rosenberg, A. (1982). In defence of convergent realism. *Philosophy of Science, 49*, 604–615.

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

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

Sklar, L. (1975). Methodological conservatism. *Philosophical Review, 84*, 384–400.

Stanford, P. K. (2006). *Exceeding our grasp: Science, history, and the problem of unconceived alternatives*. Oxford: Oxford University Press.

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