

Author's Response

By P. Kyle Stanford

I first want to thank my commentators for the constructive spirit in which they have approached my work and the care with which they have considered it. In response I'll begin with some small but crucial points of contention about what the central argument of *Exceeding Our Grasp* really requires. The New Induction (NI) claims that at the time we have embraced any given scientific theory on the strength of a given body of evidence, there typically have been fundamentally distinct theoretical alternatives also well-confirmed by that evidence (often including those accepted by later scientific communities) that simply remained unconceived by theorists at the time.¹ Because this predicament has recurred systematically throughout the history of virtually every scientific field, we have every reason to believe and no reason to doubt that it is probably our current predicament as well.

But following Magnus (2006), Saatsi worries about how we can know that “the Newtonians, for example, operating in their cultural and scientific context, would have accepted the relativistic framework as a plausible alternative to theirs, given the data they had” (xx). To insist that any genuinely worrying unconceived alternative meet this demand, however, surely gets things the wrong way round. The NI suggests, after all,

¹ Note that although a serious version of the NI strictly requires only that past unconceived alternatives were *not ruled out* by the available evidence (cf. EOG p. 19n.), I sought to defend the stronger claim of equal confirmation for past unconceived alternatives in part to help deflect the suggestion that such alternatives remained unconceived only because they were poorly confirmed at the time (see below).

that we should expect to find presently unconceived theoretical alternatives that stand in just the relationship to present evidence as, say, General Relativity did to the evidence for Newtonian Mechanics or Mendelian genetics did to the evidence for Darwinian Pangenesis, and these previously unconceived alternatives would *actually come to be accepted* by later scientific communities. Thus, if features of the “cultural and scientific context” would have prevented such alternatives from being regarded as “plausible” at the time that they remained unconceived, this should lead us to worry about the stability of whatever features of the “cultural and scientific context” inform our *own* judgments of plausibility, not to conclude (quixotically) that genuinely threatening unconceived alternatives never really existed and therefore probably do not now!

To put the matter another way, insisting that Newtonians would have rejected General Relativity as implausible only helps to undermine the force of the NI if we *also* assume both (i) that this is what prevented Newtonians from conceiving of General Relativity in the first place *and* (ii) that there are no comparable changes in store for scientific communities of the present day in the background assumptions, range of evidence, or whatever other features of the “cultural and scientific context” are supposed to have grounded both the (hypothetical) judgment that General Relativity was not a plausible competitor and the (consequent) failure to conceive of it. After all, features of scientific and cultural context have *actually* varied historically in ways that have undermined such previously entrenched judgments of implausibility, and we would seem to have little reason to believe that we are now at the end of this process or that our grasp of the relevant evidence is now substantially complete. So if indeed Newtonians would

have dismissed General Relativity as “implausible” had they considered it, this more forcefully challenges the idea that we can straightforwardly rely on our own standards of scientific plausibility than the idea that the epistemic predicament of earlier scientific communities is also our own. And the fact that previously unconceived alternatives like General Relativity have been ultimately accepted is important to the NI not because it implies that earlier theories have been overturned, but instead because it shows that these theoretical possibilities were, in addition to being supported by the available data, scientifically serious and “plausible” in the only sense that really matters here.

Of course, in *Exceeding Our Grasp*, I tried to evade the need to debate any claim of relative fixity or privilege for relevant features of our own cultural and scientific context. My detailed historical investigation of 19th Century theories of inheritance and generation sought to illustrate that we are not good at exhausting the space of well-confirmed theoretical alternatives *even within a single shared “cultural and scientific context”*. Thus I do suggest that Darwin, for example, would have regarded the fundamental mechanisms of inheritance later proposed by Galton, Weismann, and even Mendel as perfectly serious and plausible competitors to his own account had he managed to conceive of or consider these possibilities (which is not to say he would ultimately have accepted any of them). And if this pattern is indeed general, it gives us good reason to doubt that we are conceiving of all the well-confirmed theoretical alternatives that would count as “plausible” even by just the lights of our *own* “cultural and scientific context”.

As Winther is centrally concerned with what is being left *out* of this story, I should perhaps say explicitly in this connection that I fully share the interest of his Constructivist in how the details of human pragmatic agency and cognitive machinery affect the constitution of evidence and the processes of scientific change more generally, but all of this seems to me to complement my aims in *Exceeding Our Grasp*, rather than compete with them. Indeed, what his Constructivist proposes is largely an empirical exploration of the various dynamical processes that help explain how and why particular unconceived alternatives remain unconceived by particular (human!) scientists and scientific communities. Of course, in *Exceeding Our Grasp* I also tried to avoid tying my exploration of the reality and consequences of our repeated failures to conceive of the full range of well-confirmed theoretical alternatives to any particular account of the (presumably heterogeneous and untidy) *sources* of those failures.

In any case, whatever the reasons for our repeated failures to conceive of the full range of well-confirmed theoretical alternatives even within a single historical and scientific context, it remains true that realists have sometimes sought to insulate present theories from invidious comparison with their predecessors in ways that, if successful, might help to undermine the NI as well as the PI. In this connection we should consider Saatsi's claim that I ignore the recent realist focus on predicting novel or surprising phenomena as the kind of success that genuinely requires the truth of a theory that enjoys it. This emphasis on novel predictive success is itself, of course, nothing new: it has been a staple of realist argument at least since Herschel and Whewell used it to defend what they saw as legitimate uses of the method of hypothesis against inductivist critics

(see Laudan, 1981, 127ff). But notice that such an appeal would have to work quite differently against the NI than the PI. Against the PI, realists can plausibly suggest that it is simply unwarranted to draw conclusions about the ultimate fortunes of theories that did not enjoy any novel predictive success from the ultimate fortunes of those that do – after all, this is a difference between the two sets of theories that might well make it inappropriate to project inductively from one to the other. But the same appeal will not work in the case of the NI, for here we are instead projecting from the repeated failure of past theorists and scientific communities to conceive of the full range of well-confirmed alternative theoretical possibilities to the likely failure of present scientists and scientific communities to do so. The fact that some of the *theories* we have discovered along the way manage to successfully predict the existence of novel phenomena does nothing to show that the attempts of past *theorists* to exhaust the space of theoretical alternatives well-confirmed by the evidence are relevantly unlike the attempts of present *theorists* to do so or that there cannot be well-confirmed unconceived alternatives to a theory that enjoys success of this kind. So the appeal must be different. The claim must be that novel predictive success is a reliable sign that we have found a true theory *even if there are* well-confirmed alternatives to it that remain presently unconceived: we can safely ignore any such alternatives because only the (approximate) truth can (probably) have novel predictive success, so the novel predictive success of our present theory tells us that we have (probably) already found the (approximate) truth.

Even if this claim were true, of course, it would simply give up realism concerning theories in the many domains of science for which confirmation seems to

come by way of broad explanatory scope and unifying power (important parts of the biological sciences, for example) rather than predictive success of any kind, much less novel predictive success. But in any case, *we already know that the claim is false*. Many predictions of novel phenomena have been made by theories that have turned out to be fundamentally mistaken, including the paradigmatic example of Fresnel’s formulation of the wave or “undulatory” theory of light, which predicted that there should be a bright spot at the center of the shadow of a perfectly circular disc – a prediction famously treated as a *reductio* of the theory (by Poisson) until Arago bothered to actually perform the relevant experiment! Notice that despite Saatsi’s suggestion that the PI “seems to play a major role in NI as well” (xx), and Psillos’ suggestion that the NI places realism “in jeopardy only if the pessimistic induction were sound” (Psillos, xx; see also Chakravartty 2008), the difference between the PI and NI is crucial here: even if (as Saatsi suggests) the class of theories with novel predictive success that have subsequently been overturned is indeed too small to form a convincing basis for an inductive projection, it is nonetheless large enough to undermine the view that novel predictive success is a clear sign of (approximate) truth that therefore allows us to simply dismiss our independently motivated worries about the possibility and significance of unconceived alternatives.

In any case, it is not quite fair to suggest that my discussion *completely* ignores the recent realist emphasis on the importance of novel predictive success. I explicitly note that some of the theories of inheritance and development I discuss to which well-confirmed alternatives remained unconceived enjoyed such novel predictive success,

such as Weismann's widely influential theoretical prediction of the need for a reduction division in the formation of sex cells, and the prediction made by teleomechanist thinkers that gill slits should be found in the course of human ontogenetic development. These examples further illustrate that novel predictive success is no proof of either the approximate truth of a theory that enjoys it or the absence of fundamentally distinct unconceived alternatives to that theory that are also well-confirmed by the evidence in support of it (including the ability to make the very prediction(s) in question).²

Defenders of novel predictive success might try to seek criteria of novel prediction that would exclude such problematic examples, but down this road looms an unpleasant dilemma for the realist. The stricter her criteria for genuinely novel predictive success, the more of contemporary science she excludes from realist treatment while nonetheless failing to eliminate troublesome paradigmatic instances like the Poisson bright spot. But the more permissive her criteria of genuine novel predictive success become, the more conclusive evidence history provides that such success is no clear sign of truth, the absence of unconceived alternatives, or anything else that will help her case.

Of course, Saatsi makes clear that his favored response to this sort of challenge is to weaken the epistemic entitlements upon which he thinks sensible realists should insist, and both he and Psillos think that I have unfairly evaluated the strategy of selective

² Such examples also introduce a version of what I have sometimes called the "threshold" problem for scientific realism: even if we concede that (some) contemporary theories have *more* novel predictive success than any ultimately rejected past theory, this gives us no reason to believe that we have now crossed over some kind of threshold in this respect such that these predictive powers are now finally substantial *enough* to ensure that there are no presently unconceived well-confirmed alternatives to those theories or that they must be true even if there are.

confirmation, in which realists argue that only those *parts or aspects* of theories centrally implicated in their empirical achievements (especially their achievement of novel predictive success) are to be believed. In *Exceeding Our Grasp* I argued that such selective realism was not genuinely supported by the merely retrospective judgment that the parts of past theories that were truly essential to their empirical successes are just those that have turned out to be accurate: this convergence is explained just as well by the fact that realists evaluate *both* what a past theory got right *and* what was truly necessary for its success using our present theoretical account of the relevant scientific domain. Such a strategy of analysis is virtually guaranteed to produce the convergence that the realist celebrates between the parts of past theories judged both true and essential for success whether the present theories used to make this evaluation are themselves even approximately true or not. In the absence of some *prospectively applicable* criterion for identifying the essential parts of a theory in advance of later developments, then, selective realism gets no credit for passing a test it could hardly have failed.

But Psillos, Saatsi, and even Winther's Constructivist insist that the required judgments could have been and can now be made prospectively: we can, they suggest, examine a theory in isolation to determine which of its parts, aspects, or elements are genuinely required for any (successful, novel) predictions it has made (or are involved in its successful "problem-solving strategies") and which are superfluous, and then proceed to ask independently whether those parts, aspects, or elements have been subsequently preserved in and/or ratified by the lights of present theories. If this claim is correct, of course, it introduces something of a puzzle: why did we (or the relevant scientific

communities) *ever* believe more than those parts or aspects of past theories on which their empirical successes really depended? Saatsi concedes that, at least *prima facie*, the novel predictive successes of Fresnel's wave theory depended on radically false theoretical assumptions, but insists that the empirical successes of Newton's mechanics depended only upon the fact "that a rather large pool of data about moving bodies could be captured with outstanding accuracy by his laws of motion, [etc.]" (xx). If so, why did the many reflective and methodologically scrupulous Newtonian scientists ever believe anything more than the laws of motion themselves? And why did Priestly, or Maxwell, or Lavoisier, or Darwin, or Galton, or Weismann, not to mention their contemporaries, ever believe more of what their own theories proposed about nature than those parts or aspects genuinely required for their successes?

Saatsi admits that there is no general recipe for prospectively identifying the aspects of a theory that ground its success, but also denies the need for one, noting in agreement with Psillos (1999) that "scientists themselves evaluate the 'working'/'idle' status of their theoretical posits every day" (xx). This claim is surely right, but the case I made against Psillos in *Exceeding Our Grasp* involved showing that scientists' own judgments about which parts of their theories are crucial to their successes and/or most likely to be preserved in later theories are routinely *mistaken* in central cases (including, it turns out, in the two central historical cases Psillos offered in support of his claim).³

³ In more recent work (2009) I have argued that scientists' confidence that particular aspects of their own theories will be preserved in their successors is often explicitly grounded in the fact that they cannot conceive of any other possible cause for a given phenomenon; if so, the problem of unconceived alternatives helps to explain why scientists' judgments of essentiality or likely persistence are themselves unreliable.

Presumably our own prospective case-by-case judgments of what is really required for the success of a theory or what is most likely to be preserved from it are not likely to be better or more reliable than those of scientists themselves, but the historical evidence suggests that theirs are not nearly reliable enough to bear the epistemic weight that selective realism would need to lay upon them.

On what, then, will Saatsi, Psillos, or any of Winther's backpackers rely in determining which parts of our theories we are entitled to believe in light of their successful prediction of novel phenomena? Even our retrospective judgments do not pick out any single feature or aspect of theories that is invariably implicated in any predictive success they might enjoy. In the case of Newton's Mechanics, Saatsi tells us, the theory's empirical successes required only the truth of the laws of motion, but the realist cannot hold that such phenomenological regularities are what we should generally expect to find preserved in theoretical transitions: this would simply be constructive empiricism by another name. In the case of Fresnel's wave theory Saatsi seems happy to allow that a correct identification of the "structure" of nature produces the theory's novel predictive success. In still other cases it seems by present lights to have been the postulation of particular entities that was crucial to a theory's novel predictive success, as is perhaps evident in the case of atoms and Brownian motion (often treated as a "novel" prediction because Brownian motion was no part of what the atomic theory was originally developed to account for). Winther's Constructivist points us towards an even broader and more heterogeneous array of features that can be preserved from the "problem-solving strategies" of a theory to its successors, but all this simply testifies to the fact that

no *one* feature or aspect of scientific theories is invariably implicated in any novel predictive success it enjoys and/or can therefore be confidently predicted to survive further theoretical upheaval or replacement by presently unconceived theoretical alternatives.

Perhaps Saatsi, Psillos, and other suitably modest realists would be content with the bare and unimprovable insistence that *something* from any sufficiently successful present theory will *somehow* be preserved *somewhere* in their successors, but this seems rather a slender reed on which to hang the realist banner. Indeed, I myself do not doubt that the empirical successes of our best scientific theories (or their predecessors, for that matter) obtain in virtue of *some* complex and interesting interconnections between those theories and the world, and I certainly believe that there will be *some* systematic relationships and important continuities between present theories and their even more empirically powerful successors. But history seems to teach us that we are in no position to say with confidence what those interconnections are or how such continuity will be realized in any particular case of fundamental scientific theorizing. And if we cannot even specify in advance which parts or aspects of our best scientific theories are true and/or will be preserved in any of their successors, we seem to have given up what the realist cared most about all along.

This, of course, brings us to Psillos' concerns about the form of instrumentalism I advocate as the appropriate epistemic attitude towards many of our fundamental scientific theories. He suggests that like many illustrious predecessors this view sets aside a special group or class of beliefs as those towards which an instrumentalist attitude simply cannot

be adopted, but this description strikes me as obscuring more than it illuminates. On the view I advance, we could in principle adopt an instrumentalist stance towards any particular belief or set of beliefs we have, treating them simply as powerful conceptual tools for mediating among other beliefs we are not treating in this same way. The point is not that any particular type of belief is automatically privileged or special, but instead that we cannot coherently adopt such an instrumentalist stance towards all of our beliefs at the same time. And as it turns out, we have good epistemological reasons for taking such an instrumentalist stance towards some of our beliefs and not others.

The instrumentalism I advocate is strictly modeled on the realist's *own* attitude towards a theory like Newtonian mechanics, which she thinks is fundamentally false but nonetheless serves under a wide range of conditions as a powerful and reliable instrument for mediating her engagement with a variety of mechanical phenomena (the motion of rockets, moons, cannonballs, tides, etc.) to which she takes herself to have some independent route(s) of epistemic access. For the realist herself, the instrumental utility of Newtonian mechanics *consists in* the truth of many of its implications concerning rockets, moons, cannonballs, tides, etc. And for those who are prepared to adopt the very same attitude towards a much wider range of successful fundamental scientific theories, there will still remain an extremely wide range of beliefs (including most of what we sometimes call our evolving and scientifically educated common sense about the world) for which we simply do not have the same rationale for adopting instrumentalism that we have in the case of many fundamental scientific theories. The former beliefs (e.g. if I drop this cannonball out of this window, it will fall and hit the ground after X number of

seconds...) remain available to serve as the those with respect to which we think the latter (e.g. the gravitational attraction between two massive bodies creates a force...) function merely as powerful instruments for prediction and intervention, for they are either not supported by eliminative inferences at all, or the eliminative support we have for them is of a sort for which we have no historical evidence that the prospect of radically distinct well-confirmed unconceived alternatives is anything more than a speculative possibility. It is these former beliefs that I suggested an instrumentalist could treat as “strictly and literally true”, though Psillos is surely right that this is a poor choice of words: what I meant is simply that we take them to be true in the same straightforward sense that the scientific realist thinks the claims made by Newton’s radically false theory about the behavior of the rockets, moons, cannonballs, and tides of our everyday experience are more-or-less just plain true. But this is emphatically not to say that there is some special *domain* of objects, events, or phenomena concerning which our beliefs must all be strictly and literally true: after all, a cannonball is also a matter field. So when we adopt an instrumentalist stance towards a given theory we do not believe claims about objects, events, and phenomena *when* they can be understood independently of that theory, but instead *as* they can be understood independently of that theory and any others towards which we adopt an instrumentalist stance.

Does all this involve, as Psillos suggests, a “double standard” in confirmation? If so, it is a double standard I think we can and should embrace. 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 that we routinely gather knowledge about the world around us (even if this knowledge is also ultimately “theoretical” in character). If there is a double standard here it simply recommends, as all good epistemological double standards do, that we treat beliefs differently when there are important differences in the kinds of evidence we have in support of them.

Of course this picture does allow, just as Psillos suggests, that we are accumulating more and more knowledge about the world all the time. The iterative progress of scientific inquiry on which Saatsi rightly insists allows us to rule out more and more candidate theoretical accounts of the various domains of nature, and in the process develop increasingly powerful conceptual tools for mediating our engagement with nature. Along the way we steadily add to that part of our knowledge that does not rest on suspect eliminative foundations. Thus, the question isn't whether science is great, whether its iterative methods are a marvel, or even whether we make progress – it is, they are, and we do – but whether we are epistemically entitled to believe the descriptions offered by our best scientific theories (or even some systematically and reliably identifiable part or aspect of those descriptions) of otherwise inaccessible domains of

nature – and we're not. As the judo masters know, the song says it true: there's no success like failure, and failure's no success at all.

References

Chakravartty, A. 2008, "What You Don't Know Can't Hurt You: Realism and the Unconceived," *Philosophical Studies* 137: 149-158.

Laudan, L. 1981, "The Epistemology of Light: Some Methodological Issues in the Subtle Fluids Debate," Ch. 8 of *Science and Hypothesis* (New York: Taylor and Francis): 111-140.

Magnus, P. D. 2006, "What's New About the New Induction?," *Synthese* 148: 295-301.

Stanford, P. K. 2009, "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.

Department of Logic and Philosophy of Science

University of California, Irvine

Irvine, CA

USA