An Antirealist Explanation of the Success of Science

Author(s): P. Kyle Stanford
Source: *Philosophy of Science*, Vol. 67, No. 2 (Jun., 2000), pp. 266-284
Published by: The University of Chicago Press on behalf of the Philosophy of Science Association
Stable URL: http://www.jstor.org/stable/188724
Accessed: 26-09-2016 19:14 UTC

# An Antirealist Explanation of the Success of Science*

P. Kyle Stanford†‡
Department of Logic and Philosophy of Science, Department of Philosophy
University of California, Irvine

I develop an account of *predictive similarity* that allows even Antirealists who accept a correspondence conception of truth to answer the Realist demand (recently given sophisticated reformulations by Musgrave and Leplin) to explain the success of particular scientific theories by appeal to some intrinsic feature of those theories (notwithstanding the failure of past efforts by van Fraassen, Fine, and Laudan). I conclude by arguing that we have no reason to find truth a better (i.e., more plausible) explanation of a theory's success than predictive similarity, even of its success in making *novel* predictions.

**1. Introduction.** The strongest plank in the support for Scientific Realism is an argument classically articulated by Popper (1963), Smart (1968), Putnam (1975, 1978), and Boyd (1984) and which has received a powerful new formulation in the hands of such authors as Alan Musgrave (1988) and Jarrett Leplin (1997). The argument is that the only satisfactory explanation for the success of our scientific theories is that they are true (or approximately true, or true in those respects which are actually responsible for their success) in something very like the classical correspondence senses of these terms. This is sometimes called the Ultimate Argument for Scientific Realism (first by van Fraassen (1980)), or the Miracle argument (because, paraphrasing Putnam (1975, 73), Realism 'is the only philosophy

*Received September 1999; revised November 1999.

†Send requests for reprints to the author, Department of Logic and Philosophy of Science, Department of Philosophy, University of California, Irvine, CA 92697-5100.

‡Thanks are owed to the members of my Spring 1999 graduate seminar for their input on the argument I present below and to Pen Maddy, Philip Kitcher, two anonymous referees for *Philosophy of Science,* and especially to Jeff Barrett for their encouragement and helpful comments on earlier drafts of this paper.

Philosophy of Science, 67 (June 2000) pp. 266–284. 0031-8248/2000/6702-0008$2.00

that doesn't make the success of science a miracle'), and it is often thought of as strong enough alone to settle the case in favor of Realism: no matter what Realist nits Antirealists can find to pick, no alternative account of the epistemic status of our scientific theories can be taken seriously unless it provides some compelling alternative explanation of the success of those theories.

I begin by explicitly laying aside any number of legitimate and important concerns about the very coherence or intelligibility of the Realist's correspondence notion of truth and proceed to consider the position of a character I will call the 'Epistemic Antirealist', who grants the Realist the significance and defensibility of a correspondence conception of truth as well as the claim that there *is* always some theory true of a given scientific domain in this correspondence sense, but who insists that we are never in a position to know whether any theory we have discovered, tested, and/or applied is in fact this theoretical truth of the matter or not.[1] In other words, I propose to make things as difficult as possible for the Antirealist who seeks to defend her position from the Ultimate Argument.

What I hope to show (past Antirealist failures notwithstanding) is, first, that there is indeed an explanation available even to such an Epistemic Antirealist for the success of our scientific theories, and, second, that it is far from clear that the Realist explanation for the success of science is even the best explanation of that phenomenon, much less the only available explanation of it.

**2. Antirealism and the Success of Our Scientific Theories.** The history of Antirealist thinkers' efforts to explain (or explain away) the success of science does little to inspire confidence in this enterprise. Consider, for example, Bas van Fraassen's suggestion (1980) that the empirical adequacy of a theory may be cited in explanation of its success. Van Fraassen is aware that his Realist opponent will find this unsatisfying, a mere 'verbal' explanation, but he argues that it is illegitimate to insist that there must be some further explanation for a theory's success: "that the observable phenomena exhibit these regularities, because of which they fit the theory, is merely a brute fact, and may or may not have an explanation in terms of unobservable facts 'behind the phenomena' " (1980, 24). If the Realist objects to relying upon brute regularities or coincidences that do not themselves have an explanation in terms of deeper structure, he points out, the explanatory demand she posits is incoherent, for our explanations must rest content with *some* such brute regularities and coincidences in any case: it "does not even make sense," he insists, to claim that we must

1. I take van Fraassen, for one, to be such an Antirealist, but nothing turns on this.

eliminate coincidences or accidental correlations in general in order to obtain satisfactory explanations.

Fair enough. Perhaps van Fraassen's constructive empiricist explanation is not to be rejected *merely* because it appeals to brute facts or unexplained correlations, but this is a far cry from showing that just *any* appeal to brute facts or unexplained regularities should be accepted as explanatory. The trouble with van Fraassen's argument is not that it appeals to brute regularities or brute facts, but that it doesn't give us any reason to think that it is at just this particular point—at the brute fact of the empirical adequacy of a theory—that we should draw the line and end our demands for explanation. Explaining the success of a theory by appeal to its empirical adequacy is, in essence, to explain why some of the observational consequences of a theory are true by pointing out that all of its observational consequences are true; as Musgrave points out (1988, 242), this "is like explaining why some crows are black by saying that they all are." This is not to deny that subsumption under a generalization can ever constitute an explanation, still less to insist that explanations must make no appeals to brute facts or unexplained correlations; nonetheless, it seems perfectly natural and appropriate in this instance to ask in turn what it is that enables the theory to be empirically adequate or accounts for its empirical adequacy. Van Fraassen gives us no reason for ending our search for explanations with empirical adequacy, and no justification for refusing to answer the question at just this point.

Similar problems afflict a strategy characterized (although not endorsed) by Arthur Fine for explaining the success of science without any commitment to Realism, a strategy dubbed 'surrealism' by Jarrett Leplin (1988). The surrealist strategy makes central use of the 'as if' operator (or some equivalent), and Fine claims (1986, 154) that to every Realist explanation of some set of phenomena, there corresponds a better instrumentalist explanation: the claim that the world is 'as if' the theory were true suffices to explain the phenomena but takes less of an epistemic risk and commits us to less. One worry about this surrealist strategy, articulated by Musgrave (1988, 243–244), is whether it is really any more than verbally distinct from the constructive empiricist's appeal to empirical adequacy. After all, if in saying that the world is 'as if' the theory is true we mean simply that the world is *observationally* as if the theory is true, this is simply to assert the empirical adequacy of the theory, and, as an explanation of the theory's success, faces the problem noted above. On the other hand, it is hard to see what *more* about the world the surrealist means to suggest is 'as if' the theory is true, short of simply asserting the truth of the theory. Thus, the surrealist has some work left to do if she is to convince us that she is actually in possession of some explanation for the success of science distinct from both the Realist appeal to truth that she is trying to avoid and the constructive empiricist appeal to empirical adequacy that we cannot accept.

But even if this problem can be solved and the surrealist can carve out some appropriate middle ground between these possibilities, this explanation for the success of our scientific theories will still face the same charge of abortiveness that was leveled against van Fraassen's appeal to empirical adequacy. Consider the surrealist claim that the world is (behaves?) just as if a particular theory were true—is the Realist within her rights to regard this fact, too, as an unexplained coincidence or an extraordinary miracle crying out for explanation and inviting the actual truth of the theory as its explanans? It might seem not, for we might read the surrealist's claim that the world is 'as if' the theory were true to imply that the theory is not, in fact, true. But recall that the point of embracing Fine's surrealist explanation is supposed to be that it is less epistemically risky and commits us to less than the Realist alternative—thus, surrealism must be understood to imply agnosticism about the truth of the theory rather than the theory's falsity. Thus, the surrealist explanation must be understood as not taking any position on the truth of the theory, and yet asserting something like 'the world is (for all we can tell) just as if the theory were true', and this fact does indeed both cry out for some further explanation and invite the truth of the theory as just the thing that would do the trick. Like van Fraassen's constructive empiricist, the surrealist can neither satisfy the further explanatory demand she creates, nor give us some reason to think that the explanatory demand can be legitimately refused at just this point.

Van Fraassen proposes a further strategy for explaining the success of science, the so-called 'Darwinian' strategy, introduced by way of an analogy. While his Realist opponents would (à la St. Augustine) explain the fact that the mouse runs from the cat by claiming that the mouse *perceives that* the cat is its enemy, that its thought is adequate to the structure of nature, van Fraassen counsels a Darwinian approach instead: "the Darwinist says: Do not ask why the *mouse* runs from its enemy. Species which did not cope with their natural enemies no longer exist. That is why there are only ones who do" (1980, 39). Similarly, van Fraassen suggests that there is no mystery about why science is successful, for "any scientific theory is born into a life of fierce competition, a jungle red in tooth and claw. Only the successful theories survive—the ones which *in fact* latched on to actual regularities in nature" (1980, 40).[2]

2. This Darwinian proposal is actually offered (in van Fraassen 1980) in reply to Putnam's Ultimate argument *explicitly considered* (probably mistakenly) as something *besides* a version of J. J. C. Smart's 'cosmic coincidence' argument (that the success of our scientific theories would be a 'cosmic coincidence' if they were not true), an argument van Fraassen takes himself to have satisfactorily answered with his earlier appeal

This Darwinian account has been rightly attacked, perhaps in the most convincing detail by Philip Kitcher (1993). Kitcher argues that van Fraassen appeals to Darwinian explanation in the relatively shallow manner that has traditionally drawn charges of tautology against Darwinism: it offers at best a bare appeal to the high fitness of some organisms (or theories) in the relevant environments to explain why they have persisted.[3] Genuinely substantive Darwinian explanations, Kitcher points out, go much further, offering an analysis of precisely how the possession of particular traits by organisms in their present environment (or by an ancestral population in some actual past environment) contributed to the increased survivorship and/or reproductive success of those organisms. Indeed, the Darwinian analogy plays into the hands of van Fraassen's Realist opponents, who have an analysis of the 'organism-environment' relationship that accounts for the success of our theories close at hand—namely, the truth of those theories—pointing up the fact that van Fraassen himself has nothing in the way of genuine explanation to offer on this score. His Realist opponents can embrace the Darwinian analysis and insist that we must appeal to the truth of our successful theories in order to complete any real Darwinian explanation of their persistence and success.[4]

to the empirical adequacy of our theories and his insistence that explanations must end at some point with an appeal to brute facts and accidental regularities. It is therefore curious that a number of commentators have construed this Darwinian proposal as all that van Fraassen has to say in explaining the success of our scientific theories; Musgrave (1988, 242), for example, characterizes empirical adequacy as an explanation van Fraassen might, but does not, give for the success of scientific theories, and the Darwinian explanation as what van Fraassen actually has to say on the matter. It is also noteworthy that the Darwinian proposal is explicitly offered as an answer to the question "why we have successful scientific theories at all," rather than as a way to explain the success of particular scientific theories (1980, 39), although in fairness to his commentators, van Fraassen immediately goes on (1980, 40) to treat the Darwinian proposal as an answer to the latter explanatory demand rather than the former. In any case, it is the latter explanatory demand which concerns us at the moment (see my discussion of Laudan 1984, below), and I am arguing that neither of van Fraassen's strategies constitutes a satisfying way to address it.

3. At worst it offers a bare appeal to the fact of their survivorship to 'explain' why these theories have survived.

4. In a curious (and often overlooked) footnote (1980, 40n), van Fraassen seems to recognize this problem for the Antirealist, deferring discussion of it until he addresses pragmatic virtues and explanation proper (Ch. 4, §4, and Ch. 5), but I do not see how the latter discussions will ultimately help. The footnote suggests that van Fraassen thinks there are two distinct kinds of Darwinian explanation, one aimed at general questions (like why organisms run from their predators or why science is successful) and another aimed at more particular questions (like why *mice* run from their enemies or why Balmer's formula for the line spectrum of hydrogen survives as a successful hypothesis). If so, this is simply a mistake: there is only a single program of Darwinian explanation and it does not support van Fraassen's case in the way he imagines.

Let us consider finally, then, an alternative Antirealist reply which rejects the demand for an explanation of this kind at all. Laudan (1984, 92) points out a crucial ambiguity in the demand that we 'explain the success of science': on the one hand, we might want to be told the features possessed by theories in virtue of which they are able to make such impressive predictions, but we might be asking instead for some account of why the theory-selection procedures characteristic of scientific activity are so reliable in identifying theories with this sort of predictive power. These are demands to explain quite different phenomena; Laudan suggests first that it is usually the latter that we are "driving at" when we ask for an explanation of the success of science, and second, that this explanatory demand can be answered quite convincingly without any appeal to the truth (or indeed, to any semantic feature) of our theories whatsoever. Further, we can answer this latter explanatory demand in a piecemeal and local fashion, in that we can appeal to quite particular practices of theory-selection to explain how we are able to avoid equally particular methodological pitfalls: we isolate variables to avoid the fallacy of post hoc ergo propter hoc, we use control groups in our experiments to distinguish baseline occurrences of phenomena from experimentally induced ones, we perform experiments in double-blind fashion to prevent expectations from influencing either experimental subjects or experimenters, and so on.[5] In each case we can show how the methodological stricture in question contributes to our ability to identify predictively reliable scientific theories without appealing to the truth (or any intrinsic feature) of the theories so selected.

Laudan does us an important service in disentangling the demand to explain the ability of a theory to make successful prediction from the demand to explain how science is able to identify theories which are able to do so, and perhaps even makes a promising start on addressing the latter explanatory demand. What Laudan does not do is give us any reason to think that the former explanatory demand is somehow illegitimate or misguided, or provide us any resources for addressing it.

Leplin (1997) uses a revealing analogy to illustrate why both of these explanatory demands are legitimate ones. If someone observes the two Wimbledon finalists on television and asks why they are such great tennis players, there are two ways to take this question: we may either take it as a request to explain why *Wimbledon finalists* are great players, in which case it is appropriate to note the difficult hurdles which must be surmounted in order to reach the finals of Wimbledon, or we may take it as a request to explain why these *particular individuals* are such great players, in which case our explanation must cite relevant features *of those players*

5. These are Laudan's examples.

(like their training and their native athletic abilities) which enable them to surmount those hurdles where less gifted players failed. Leplin continues:

> Analogously, to explain why the theories *that we select* are successful, it is appropriate to cite the stringency of our criteria for selection. But to explain why *particular theories*, those we happen to select, are successful, we must cite properties *of them* that have enabled them to satisfy our criteria. Laudan argues as though the fact that our explanatory question can be given a reading on which stringency of criteria for selection is the right answer obviates the need to consider attributes of successful theories themselves. (1997, 9; emphasis in original)

The charge is not quite fair, for what Laudan actually says is that he suspects that it is generally the former explanatory demand rather than the latter that we are "driving at" (1984, 92) when we demand that the success of science be explained. Still, Laudan nowhere offers any support for this suspicion, nor does he give us any reason to think that the quite distinct demand to explain the success of our particular scientific theories can or should be dismissed as unimportant, illegitimate or otherwise not in need of an answer.

The state of the playing field is not very promising, then, for the Antirealist: it seems that she can explain the success of science only by changing the rules of the game and explaining quite a different phenomenon than the one for which the Realist accused her of being unable to provide any explanation at all. A serious and apparently legitimate explanatory demand survives these Antirealist maneuvers, one that Leplin characterizes in the following way:

> if we ask of successful theories *why* they are successful, we need an answer that goes beyond an explanation of why science in general produces successful theories; we need an answer that appeals to attributes that discriminate among theories. Why does *this* theory work, while others equally the products of diligence and preferred methods fail? (1997, 8; emphasis in original)

The next section will argue that there is indeed an explanation for the success of our scientific theories that is available even to the Epistemic Antirealist which neither appeals to the truth of those theories nor ignores the legitimate insistence that we be able to explain the success of particular scientific theories by appeal to some intrinsic feature of those successful theories themselves.

**3. Predictive Similarity and the Success of Science.** Let us begin by asking how we would go about explaining the success of a theory that we already took to be false. Ironically, we can get a start by considering the answer offered to this question in the course of J. J. C. Smart's classic defense of Scientific Realism (1968). There he says,

> Consider a man (in the sixteenth century) who is a realist about the Copernican hypothesis but instrumentalist about the Ptolemaic one. He can explain the instrumental usefulness of the Ptolemaic system of epicycles because he can prove that the Ptolemaic system can produce almost the same predictions about the apparent motions of the planets as does the Copernican hypothesis. Hence the assumption of the realist truth of the Copernican hypothesis explains the instrumental usefulness of the Ptolemaic one. Such an explanation of the instrumental usefulness of certain theories would not be possible if *all* theories were regarded as merely instrumental. (151; emphasis in original)

This seems quite a plausible analysis of the particular case: we do indeed explain the success of the (revised) Ptolemaic system of epicycles[6] by pointing out how closely its predictions approximate those of the true Copernican hypothesis. Let us call this relationship the *predictive similarity* of the Ptolemaic system to the Copernican. Furthermore, it seems that this appeal to predictive similarity is indeed the natural place to end our demand for explanation of the success of the Ptolemaic system: asked further, *why* it is that the Ptolemaic system approximates the predictions of the true Copernican one or *how* it, in particular, is able to accomplish this magnificent feat, we would appropriately (and could only) either direct the questioner to the details of the Ptolemaic system itself, to see how its specific predictions arise from the mechanics of the theory, or greet her with a puzzled look and a shrug. No further explanation of what *intrinsic feature of the theory* enables it to be successful is appropriate or possible.[7]

It might seem that there is indeed more to say about the success of Ptolemaic astronomy. In particular, it might seem tempting to suppose that we can achieve a somehow deeper explanation of the success of Ptolemaic astronomy by pointing out a kind of 'structural similarity' or 'isomorphism' between the Copernican system and the particular (successful) version of the Ptolemaic system in question. The problem with this suggestion is not that there is no such structural similarity, but rather that such structural similarity is so easy to come by as to be explanatorily

6. Note that we are not here concerned with the success of the Ptolemaic *strategy* of spinning out epicycles (which was *historically* successful in part because it is capable of reconstructing virtually *any* set of predictions about changes in relative celestial position), but rather with the success of one particular theory of celestial motions generated by means of this strategy.

7. Here I am disputing (by example) Leplin's (1997, 14) contention that an appeal to the fact that a theory makes the same predictions as the true theoretical account of the matter is simply not itself explanatory.

vacuous: between any two theories that make similar predictions over a domain of any significant extent, there is sure to be *something* we could fasten onto as a structural similarity or isomorphism between them. (Indeed, this can typically be achieved by working backwards from their systematic range of similar predictions: something we *could* characterize as a structural similarity or isomorphism is bound to show up.) Thus, it explains nothing more *about* the success of a theory that is predictively similar to the true account of the matter to be told further that it bears some 'structural similarity' to that true account.[8]

We should, then, accept Smart's contention that explaining the success of the false Ptolemaic hypothesis simply requires pointing out both the truth of the Copernican hypothesis and the fact that the Ptolemaic hypothesis is able to generate sufficiently similar predictions to those of the Copernican hypothesis (in the relevant domain). But the general moral Smart draws from the case should give us serious pause. Notice that the actual *content* of the Copernican hypothesis plays *no role whatsoever* in the explanation we get of the success of the Ptolemaic system: what matters is simply that there *is* some true theoretical account of the domain in question and that the predictions of the Ptolemaic system are sufficiently close to the predictions made by that true theoretical account. To see this, consider the explanatory impact of dropping from Smart's favored explanation of the success of the Ptolemaic system ('because it is predictively similar to the true, Copernican system') either the information that it is the *Copernican* system to which the Ptolemaic is predictively similar, or that it is the *true* system to which it bears this relation. If we drop out just

8. The illusion that the mere fact of structural similarity or isomorphism carries further explanatory significance is created by the fact that the *details* of the particular structural similarity at issue in some particular case *can* contribute to explaining why two theories make much the same predictions (over particular domains); indeed, articulating the relevant structural similarity in a particular case will often simply *amount to* something very like showing how the mechanics of the false theory produce predictions that are, in fact, sufficiently similar to those generated by the true account. But this cannot be a source of comfort to the Realist's explanationist defense of Realism, for the *particular* structural similarity at issue will be quite different in different cases, while the Realist's defense of her position requires that the only plausible explanation of the success of our theories appeal to some *general* feature that obtains in all or most cases and that ensures the truth (or substantive identity to the truth) of those theories. But as we have seen, the only sense in which structural similarity or isomorphism to our present theories could qualify as a *general* feature of past false-but-successful theories must be characterized in a way that is so vague and easily satisfied that it cannot do any real work in explaining the success of those theories: while their predictive similarity to the true theoretical accounts of their respective domains is genuinely explanatory, the fact that this predictive similarity will invariably be sufficient to generate *something or other* that could be characterized as a structural similarity or isomorphism between the two theories is itself of no further explanatory significance.

the information that it is the *Copernican* system to which the Ptolemaic is predictively similar, we are still left with an explanation (to wit, 'because its predictions are nearly identical to those made by the true theoretical account of the matter') that renders perfectly understandable *why* Ptolemaic astronomy is so successful.[9] We could not, however, drop out just the information that the Ptolemaic system is predictively similar to the *true* account without undermining our answer's explanatory value: it is perfectly *un*helpful to inform the questioner that the Ptolemaic system makes nearly identical predictions to those of some other particular theory (the Copernican hypothesis) unless the questioner already knows that this other theory is itself the truth of the matter. Thus, Smart notwithstanding, it is the fact that the Ptolemaic system is predictively similar to the *true* theoretical account of the relevant domain that explains its usefulness, not that it is predictively similar to the *Copernican* hypothesis as such.

Notice that (contra one natural way of reading Smart's final remark above) this explanatory appeal to predictive similarity works perfectly well even if all actual theories we possess are considered to be false: what matters is that there are facts of the matter about the underlying mechanisms in the domains about which we theorize and that our theories (sometimes) make predictions that are (sufficiently) close to those made by the true accounts of the relevant mechanisms. This suggests a natural strategy for generalizing the results of this case into an explanation of the success of any theory whose basic claims are substantially false: *the success of a given false theory in a particular domain is explained by the fact that its predictions*[10] *are (sufficiently) close to those made by the true theoretical account of the relevant domain.* Perhaps most importantly, it *is* appropriate to end our demand for explanation there: just as we saw in the case of explaining the success of the Ptolemaic system, above, it is inappropriate to ask what further characteristic *of the theory* accounts for or explains its predictive similarity to the truth.

Perhaps the constructive empiricist's or the surrealist's explanation of the success of our scientific theories can be charitably reinterpreted along

9. Of course, our questioner will then want to know *how* we were able to develop a theory with this impressive characteristic, but this is to ask a different question (see below).

10. Here and below I use 'prediction' in an extremely broad sense, encompassing not only predictions of future observational outcomes, but also the conditional or counterfactual predictions that make our theories successful tools for intervention in the world, and the retrodictions that confer upon false theories whatever sort of explanatory value (if any) they are capable of having. In other words, I will use 'predictive success' and the like as shorthand for *whatever* kinds of success the Realist suggests can only be explained by appeal to the truth of a theory.

the lines of this proposal, although this will take some doing. Constructive empiricists and surrealists each appeal to a relation *between a theory and the world* to explain the success of that theory: either to the accuracy of a theory's predictions about the world (i.e., to the theory's success itself), in the one case, or to the surprising fact that the world is 'as if' the theory is true, in the other. But the proposal offered here does not appeal to a relationship between a theory and the world at all; instead it appeals to a relationship of predictive similarity *between two theories*. This is also why predictive similarity is not simply a redescription of predictive success (i.e., 'making true predictions'): it constitutes a relationship between the false, successful theory and the true theoretical *account* of the relevant domain that makes clear why it is no mystery or miracle that the successful theory enjoys the success that it does, without requiring that this theory itself be true.[11] Most important of all, this account invokes an explanation of the success of our scientific theories which constitutes, as we have seen, a demonstrably appropriate terminus for the chain of explanatory demands, and this is precisely what existing appeals to empirical adequacy, surrealism, and the Darwinian strategy simply fail to do.

Of course, the bare explanatory appeal to predictive similarity leaves us profoundly disappointed, for it leaves us wanting to know how it is that we are able to generate and identify false theories which have such marvelous predictive abilities. This is a perfectly legitimate question, but it is crucial to note that it raises an explanatory demand quite distinct from the demand to explain the success of any particular theory; indeed, what it raises is simply the further explanatory demand that Laudan has identified and tried to address—the demand that we explain how the characteristic methods of scientific inquiry are able to reliably generate and select such powerful theories. By surmounting particular methodological challenges of the sorts Laudan identifies, we dramatically increase our chances of generating and picking out theories that are either true or that make the same predictions as the true theoretical account of the relevant domain; of course, addressing *this* explanatory demand does not call for an appeal to any intrinsic feature(s) of our successful theories at all. What this illustrates, of course, is that once we understand the *intrinsic* feature that explains the success of false scientific theories, we see that by far the most important and interesting work of explaining the success of our sci-

11. Moreover, even if one insists that predictive similarity must collapse into empirical adequacy or surrealism, we will still have made philosophical progress, for we will then have seen why an explanation in terms of *that* intrinsic feature of our theories does indeed (appearances to the contrary) constitute a sufficient explanation of their success.

entific theories arises in addressing the completely distinct demand to explain how science is able to arrive at such theories in the first place.[12]

The overarching point for our present purposes, however, is that the perfectly legitimate demand that we explain the success of our scientific theories by appeal to some *intrinsic* feature of those theories is not, in fact, one for which the Antirealist has no answer, carrying the day for Scientific Realism. It turns out that the appeal to predictive similarity answers this demand and does so in a way that we recognize as exhausting the legitimate *scope* of the demand in parallel cases: asked how or why a theory which makes (sufficiently) similar predictions to the true theoretical account of some domain is able to do so, we can only greet the questioner, just as we did in the case of Ptolemaic astronomy, with the detailed mechanics of the theory or with a puzzled look.

**4. Novel Predictive Success.** In a recent book (1997; see also Musgrave 1988), Jarrett Leplin acknowledges that Antirealists can account for many forms of scientific success, but insists that the ability to make *novel* predictions is a form of success that can only be explained by imputing some measure of truth to the theories which enjoy it. It is well worth noting, then, that the sort of explanation I have proposed is fully adequate to explain the ability of theories to make successful, novel predictions (in either the more traditional or in Leplin's somewhat idiosyncratic sense of 'novel'[13]). That is, the fact that a theory makes predictions that are (sufficiently) close to those made by the true theoretical account of some domain suffices to explain why its *novel* predictions in that domain are successful, and answers this explanatory demand in a way that terminates the demand for explanations in parallel cases.[14]

12. Thus, we may have an argument after all for Laudan's claim that the latter explanatory demand is what we are typically 'driving at' in trying to explain the success of science.

13. Novelty, for Leplin, amounts to the satisfaction of two conditions that are not entirely intuitive: for a result predicted by a theory to count as novel for that theory, there must be an adequate reconstruction of the reasoning leading to the theory which does not appeal to even a qualitatively generalized description of the result, and second, no other extant theory may predict even a qualitatively generalized description of the result (for details, see Leplin 1997, Ch. 3). Leplin ultimately eschews our intuitions about novelty in favor of attempting to pick out just those forms of predictive success for which truth is the only available explanation. But these differences in the construal of 'novelty' do not affect the ability of a theory's predictive similarity to account for its ability to predict novel results (see below).

14. Admittedly, in the Ptolemaic case it was stipulated that the successful theory is false, while this is an open question regarding an arbitrary theory with a record of successful novel prediction. Nevertheless, it could hardly be the case that predictive similarity is insufficient to explain the success of a theory whose truth status is unknown, but adding

Leplin is wrong, then, to think that the ability to make successful, novel predictions is a form of scientific success for which *only* the truth of a theory can account, in either his sense of 'novel' or the more traditional one. But it is natural to ask whether the truth of a theory might not be a more *plausible* explanation of a theory's ability to make accurate novel predictions than the fact that its predictions are sufficiently close to those made by the true theoretical account of the matter in the relevant domain.[15] That is, given the evidence (successful novel prediction), might it not simply be more likely that a theory is true than that its predictions are sufficiently close to those of the true theoretical account of the relevant domain?

To address this question, we must first remind ourselves of the *sort* of truth that the history of scientific inquiry makes it reasonable for the Realist to assert on behalf of even our most successful scientific theories. The long record of theories with distinguished empirical successes that are nonetheless now judged to be false prevents any straightforward inference from the success of a theory (even its success in generating novel predictions) to its truth, simpliciter.[16] Realists typically respond to this skeptical induction by insisting that we have a historical warrant for believing that our most successful theories are 'approximately true'[17] or by attributing some other kind of attenuated truth to such theories.[18] Leplin, for example,

the information that the theory is false somehow *enables* predictive similarity to *become* a sufficient explanation of its predictive success.

15. Indeed, Musgrave's (1988) claim is simply that the truth of a theory provides the best available (satisfactory) explanation of its ability to make accurate novel predictions and that this makes it reasonable to tentatively suppose that the theory is true.

16. The locus classicus for this 'skeptical induction' over the history of science is Laudan 1981.

17. The other main option for the Realist is to find some relevant feature that has *not* been reliably exhibited by past theories ultimately judged false, but that is reliably exhibited by some other group of theories (e.g., those of current science) which are therefore protected from the skeptical induction. Making successful novel predictions would not, as Leplin realizes, seem a promising candidate for such a feature: after all, the paradigm case of successful novel prediction, the Poisson 'bright spot', was made by Fresnel's (false) wave theory of light (see Leplin 1997, 83–85).

18. This includes thinkers who argue (e.g., Kitcher, 1993, Ch. 5) that the *parts* of past theories that are actually responsible for their success are typically true. The problem with this strategy, of course, is that at the time of our commitment to a theory it is not usually possible to separate its operative elements from the extra baggage (assuming that this separation is coherent at all), as Maxwell's famous remark (paraphrased in Laudan 1981, 114) that "the aether was better confirmed than any other theoretical entity in natural philosophy" reminds us. Thus, what the strategy really offers us is the extremely abstract Realist commitment that we will be able to *retrospectively interpret* our present theories *in light of our later theories* as having had true features that were responsible for their success. This abstract commitment does not lay the Antirealist's

argues that we should regard theories enjoying novel predictive success as 'partially true', but holding a theory to be partially true "is only to suppose that its explanatory mechanisms capture some of the features of natural processes well enough not be misleading as to how the effects these mechanisms explain are actually produced" (1997, 104), where accuracy of representation is relativized to particular respects in which and the particular purposes for which an entity is being represented (103); it is to believe that "the theory is not on the wrong track, is not basically misleading as to what entities or processes are actually producing our observations, and pursuing it will be progressive" (133) and that "the areas in which it proved wrong are less important, in a certain respect, than the areas in which it was right"(133–134). In short, a theory is to be judged partially true just in case its development can be retrospectively interpreted as (on balance) a step forward on the road to our present theory, something which brought us closer to what we now believe, rather than pushing us farther away from it. Thus, the sort of partial, approximate or attenuated truth that Leplin is able to assert (in light of the history of science[19]) on behalf of theories that enjoy novel predictive success is extraordinarily weak.[20]

With this reminder of the sort of truth that Leplin, at least, finds it reasonable for the Realist to attribute, in light of the history of science, to theories enjoying novel predictive success, let us return to the question of the comparative plausibility of our candidate explanations for the novel predictive success of a theory: given the evidence that a theory has made

concern to rest, of course, for it does not help us to identify which (if any) of the actual claims of our present theories are those on whose truth we may safely rely.

19. Leplin does ultimately reject the skeptical induction, claiming (1997, 145, 175) that past theories which have enjoyed sustained records of novel predictive success have indeed turned out to be partially true, but in order to make this claim he must water down the assertion of partial truth to amount, once again, simply to the claim that "their eventual failure is not total failure" and that "those of their theoretical mechanisms implicated in achieving [their] warrant are recoverable from current theory" (145).

20. It is curious to note that Musgrave's otherwise subtle development of the form of the Ultimate Argument does not involve an inference to the merely approximate or partial truth of our theories; instead it offers a highly qualified ('It is reasonable to (tentatively, of course) accept . . . ') inference to their truth, simpliciter. But this would seem simply to be an oversight on Musgrave's part. He recognizes that the form of the argument he describes relies on the major premise that "It is reasonable to accept a *satisfactory* explanation of any fact, which is also the *best* available explanation of that fact, as true" (1988, 239; emphasis in original), but this is just the premise that seems perfectly *un*reasonable in light of the history of science. It seems that Musgrave should instead claim at most that it is reasonable to accept a satisfactory explanation of any fact, which is also the best available explanation of that fact, as *partially* or *approximately* true.

a successful novel prediction (NP), we want to know how likely it is that the theory is true in the Realist's partial, approximate, or attenuated sense (AT) and to compare this with the likelihood that the theory is simply predictively similar to the literal truth in the relevant domain (PS). A straightforward application of Bayes's Theorem reveals that in order for p(AT/NP) to be greater than p(PS/NP), p(NP/AT) $\times$ p(AT) must be greater than p(NP/PS) $\times$ p(PS). It would seem that the most promising strategy for the Realist to pursue in seeking to make the case for the latter claim would be for her to argue that p(NP/AT) is significantly higher than p(NP/PS), which would establish the inequality so long as p(PS) is not correspondingly higher than p(AT).[21] Let us consider, then, the relationship between p(NP/AT) and p(NP/PS).

It is crucial to note that p(NP/PS) represents the probability of successful novel prediction given predictive similarity *with respect to all known phenomena (in the domain of the theory) to date*. The probability of successful novel prediction given *complete* predictive similarity is, of course, 100%, but the prior probability of generating a theory that is *perfectly* predictively similar to the truth is presumably quite small (the probability of successful novel prediction given truth, simpliciter (including auxiliaries and empirical assumptions), is likewise 100%, and the prior probability of generating a theory that is true, simpliciter, is likewise extremely small). The issue is whether we should expect it to be more likely that a successful novel prediction will issue from an approximately true theory or more likely that it will issue from a false theory which is nonetheless predictively similar to the truth with respect to all known relevant phenomena.

To settle this question, we must compare the baseline proportion of successful novel predictions we should expect from a theory that is approximately true to the baseline proportion of such predictions we should expect from a false theory that nonetheless manages to save the known phenomena in its domain of application. It might be tempting to suppose that the former will be quite high: the fact that, *if* all of the theories and assumptions relevant to some context of testing were true, *simpliciter*, their chance of generating an accurate novel prediction would be 100%, might seem to suggest that from *approximately* true theories we should expect *approximately* the same baseline rate of accurate novel predictions, perhaps near 100%. But this inference is unjustifiable, first because it is quite easy to imagine relatively few or minor inaccuracies in a theory that would suffice to radically undermine the accuracy of most or even all of its predictions,[22] but more importantly because it ignores the *kind* of approxi-

21. See Leplin 1997, 129, for an importantly related, although distinct, analysis of the Realist's commitments.

22. See Laudan 1981 for an argument to this effect.

mate, partial or otherwise attenuated truth that was even a candidate for being indicated by a record of successful novel prediction in the first place. The partial truth supposedly indicated by a record of successful novel prediction amounted simply to the theory not being 'on the wrong track' or 'basically misleading' about the underlying mechanisms and the (possibly many) respects in which it proved wrong turning out to be less *important* than the respects in which it proved right (from the perspective of the development of our current theoretical beliefs). It is not obvious that we should expect much in the way of successful novel prediction from such a theory: any or even all of a theory's novel predictions could quite easily turn out to be one of the (possibly many) ways in which the theory did fail or was misleading, one of the many ways in which it steered us wrong which were, nevertheless, on balance less important than the ways in which it steered us right. Such scenarios are not at all farfetched, so long as the partial or approximate truth of the theory requires only that there be *something* basically right about the underlying mechanisms that the theory envisions (or perhaps even just that the theory be a progressive step on the road to our current conception of the relevant underlying mechanisms).

It is far from obvious, then, that we should expect a much higher baseline rate of successful novel prediction from a theory with just this kind of attenuated, approximate, or partial truth than we should expect from a theory that is false but whose postulated mechanisms have nonetheless managed to save all of the known phenomena in that theory's domain of application, and perhaps it is not reasonable to expect *any* higher rate of such prediction from the approximately true theory at all. This is not to say, of course, that there are a lot of reasons to think that a theory that is not even approximately true in this weak sense (a theory which *is* basically misleading about the underlying mechanism or whose development must be seen as a step *away* from our current theoretical beliefs) will make accurate novel predictions. Nonetheless, we must consider the fact that a false theory whose mechanisms are nonetheless able to save the existing phenomena in some domain of inquiry will be making its predictions about the appearance of novel phenomena in that same domain of inquiry: it is not unreasonable to think that the systematic relationship among phenomena within the same domain of inquiry (e.g., optical phenomena, chemical properties, descent with modification in organisms) alone suffices to give even a false theory a fighting chance of making novel predictions successfully when the machinery of that theory has already proved able to save all the known phenomena in the relevant domain. In any case, it seems that the attenuated sort of truth that the Realist is forced by the history of science to assert on behalf of even theories with successful novel predictions leaves us with no clear reason to expect a successful novel prediction to be much more likely (if more likely at all) to come from an approximately or partially true theory than to come from a false theory that has nonetheless managed to make predictions sufficiently similar to the theoretical truth of the matter concerning all of the known phenomena in the theory's domain of application thus far.

If we accept this conclusion, we cannot assume with any confidence that p(NP/AT) is significantly higher than p(NP/PS), in which case the Realist's case for the claim that the approximate truth of a theory is a more plausible or likely explanation of its ability to make novel predictions than predictive similarity must rely upon an unbalanced assignment of prior probabilities: p(AT) must be assumed to be significantly higher than p(PS). But it is hard to see how the Realist could give a convincing argument that our prior probabilities would or should be assigned by any impartial judge in this way: we would prejudge the Realism debate itself in holding ourselves more likely to generate (from a given body of evidence) a theory that will eventually be judged approximately true than to generate a false but predictively similar theory, or vice versa. If anything, it would seem that there are many more ways of generating a false theory that saves a given body of evidence than of generating even an approximately or partially true theory that does so, making p(PS) higher than p(AT).[23] In any case, there does not seem to be any convincing reason the Realist can give for assigning a higher prior probability to p(AT) than to p(PS)—at worst, an impartial judge would set these two priors equal to one another—denying the Realist her only remaining tool for arguing that an appeal to the approximate truth of a theory is more plausible than an appeal to predictive similarity in explaining its ability to make successful novel predictions.

None of this establishes, of course, that predictive similarity is a *more* plausible explanation for a theory's sustained record of novel predictive success than approximate truth. Nonetheless, I have argued not only that the probability ratios on which the Realist's case that predictive similarity is a *less* likely explanation might be most plausibly founded—p(AT) to p(PS) and p(NP/AT) to p(NP/PS)—are extremely difficult to estimate or compare in any evidentially responsible or non-question-begging way, but also that we do not have any clear reason to expect either of these ratios to be very far unbalanced, as the Realist's argument would require. The right course of action, then, might seem to be to admit that the best we can do by way of explaining the success of a theory is to claim that it is *either* approximately true *or* predictively similar to the theoretical truth of

23. Even this claim, however, depends upon a controversial application of the thesis that theories are underdetermined by the evidence for them, so I will not rely upon it here.

the matter in the relevant domain, but admit that we don't know which: the kind of evidence available to us enables us to responsibly infer the disjunction, but not to infer either disjunct. We might even tidy up the terminology a bit by giving predictive similarity what we can call an 'inclusionary' reading—a reading on which 'the theory's predictions are (sufficiently) close to the theoretical truth of the matter in the relevant domain' is not taken to rule out the possibility that the successful theory is itself the truth of the matter—and invoking the inclusionary reading of predictive similarity (really a disjunction which includes the possibility that the successful theory is true and the possibility that it is false but predictively similar nonetheless) as our favored explanation of success.[24] In an important sense this is, of course, mere wordplay: the inclusionary reading of predictive similarity includes two distinct possible explanations for the success of a theory by appeal to intrinsic features of that theory and simply recognizes that we lack the evidential resources to responsibly choose between them.

**5. Conclusion.** I have suggested that Realists are right to insist that there is a legitimate demand to explain the success of our scientific theories by appeal to some intrinsic characteristic(s) of those (successful) theories themselves. I have tried to show, however, that there is such an explanation available even to Epistemic Antirealists which does not invoke the truth of the successful theory and which terminates the chain of demands for explanation. Thus, the legitimacy of this demand for explanation in the form that the Realist insists it must take does not settle the dispute in favor of imputing truth to our successful scientific theories in the way that the Realist supposes.

I have not denied, of course, that the truth (even the approximate or partial truth) of a scientific theory provides an explanation of its success in making predictions, novel and otherwise. I have simply argued that truth is not the only explanation of the success of our scientific theories and that, in light of the sort of (approximate or partial) truth that the history of science makes it reasonable to attribute to successful theories,

24. This 'inclusionary' reading of predictive similarity might have looked the best bet from the very beginning, as either a theory's being true or its being false but predictively similar suffices to explain its success and the disjunction is surely more probable than either disjunct. But this is just the sort of epistemic modesty that the Realist eschews in other contexts and there is a perfectly reasonable motivation for resisting it here as well: we surely want to have the most *specific* and *precise* explanation of a phenomenon to which we are entitled by the evidence. What we must recognize here is that we are not entitled by the evidence to any explanation more precise and specific than the disjunction of the truth of the theory with its falsity and predictive similarity, that is, than the inclusionary reading of predictive similarity.

it is far from clear that the truth of a theory is even the best explanation of its success, even of its success in novel prediction: in light of the sorts of evidence that are available to us, the reasonable course would seem to be to endorse the inclusionary reading of predictive success as our explanation for the success of our scientific theories. Accordingly, I suggest that the strongest plank in the case for Scientific Realism, the argument that the success of our scientific theories would be a miracle if they were not true and that only Realism can provide an explanation (or that Realism provides the best explanation) for the success of our scientific theories, will simply not bear the argumentative weight that Realist philosophers of science have tried to place upon it.

## REFERENCES

Boyd, Richard (1984), "The Current Status of Scientific Realism", in Jarrett Leplin, (ed.), *Scientific Realism.* Berkeley: University of California Press, 41–82.

Fine, Arthur (1986), "Unnatural Attitudes: Realist and Instrumentalist Attachments to Science", *Mind* 95: 149–179.

Kitcher, Philip (1993), *The Advancement of Science.* New York: Oxford University Press.

Laudan, Larry (1981), "A Confutation of Convergent Realism", *Philosophy of Science* 48: 19–48.

———. (1984), "Explaining the Success of Science: Beyond Epistemic Realism and Relativism", in James Cushing, C. F. Delaney, and Gary Gutting (eds.), *Science and Reality.* Notre Dame: University of Notre Dame Press, 83–105.

Leplin, Jarrett (1993), "Surrealism" *Mind* 97: 519–524.

———. (1997), *A Novel Defense of Scientific Realism.* New York: Oxford University Press.

Musgrave, Alan (1988), "The Ultimate Argument for Scientific Realism", in Robert Nola (ed.), *Relativism and Realism in Science.* Dordrecht: Kluwer Academic Publishers, 229–252.

Popper, Karl R. (1963), *Conjectures and Refutations.* London: Routledge and Kegan Paul.

Putnam, Hilary (1975), *Mathematics, Matter and Method (Philosophical Papers, Vol. I).* London: Cambridge University Press.

———. (1978), *Meaning and the Moral Sciences.* London: Routledge and Kegan Paul.

Smart, J. J. C. (1968), *Between Science and Philosophy.* New York: Random House.

Van Fraassen, Bas C. (1980), *The Scientific Image.* Oxford: Clarendon Press.