Predictive Similarity and the Success of Science: A Reply to Stanford

Stathis Psillos

Philosophy of Science, Vol. 68, No. 3 (Sep., 2001), 346-355.

Stable URL:
http://links.jstor.org/sici?sici=0031-8248%28200109%2968%3A3%3C346%3APSATSO%3E2.0.CO%3B2-X

Philosophy of Science is currently published by The University of Chicago Press.

Predictive Similarity and the Success of Science: A Reply to Stanford*

Stathis Psillos†‡
Department of Philosophy and History of Science
University of Athens

P. Kyle Stanford (2000) attempts to offer a truth-linked explanation of the success of science which, he thinks, can be welcome to antirealists. He proposes an explanation of the success of a theory T_1 in terms of its predictive similarity to the true theory T of the relevant domain. After raising some qualms about the supposed antirealist credentials of Stanford's account, I examine his explanatory story in some detail and show that it fails to offer a satisfactory explanation of the success of science.

1. Introduction. Typical antirealist explanations of the success of science have refrained from appealing to truth-linked explanatory concepts. This is quite natural. The target of the antirealist is the realist 'no miracle argument', viz., (roughly) that the best explanation of the success of science is that scientific theories are (approximately) true. If the alternative antirealist explanans made use of some truth-linked notion to explain success, then it seems that antirealists would concede the basic point of the 'no miracle argument'.

In his (2000), P. Kyle Stanford attempts to offer a truth-linked explanation of the success of science, which, he thinks, can be welcome to antirealists. While dismissing standard antirealist explanations—such as van Fraassen's Darwinian account or Fine's "surrealist" one (cf. 267–270[1])—

*Received October 2000; revised February 2001.

†Send reprint requests to the author, Department of Philosophy and History of Science, University of Athens, 37 John Kennedy Street, 16121 Athens, Greece; email: psillos@phs.uoa.gr.

‡Many thanks to Kyle Stanford and two anonymous referees of *Philosophy of Science* for useful comments on an earlier version of this paper.

1. Unadorned page references are to Stanford 2000.

Philosophy of Science, 68 (September 2001) pp. 346–355. 0031-8248/2001/6803-0006$2.00

as inadequate, he proposes an explanation of the success of a theory T_1 in terms of its *predictive similarity* to the *true* theory T of the relevant domain. This is clearly a truth-linked explanation. It is the appeal to the *true* theory T that grounds predictive similarity between T_1 and T as a candidate for the explanation of the success of T_1. After raising some qualms about the supposed antirealist credentials of Stanford's account, I examine his explanatory story in some detail and show that it fails to offer a satisfactory explanation of the success of science.

2. Is Stanford's Account Antirealist? I take the dialectic of the debate around the explanation of the success of science to be the following: where realists go—very roughly—from success to truth, antirealists should try to show that there are alternative explanations of this success which do not involve the truth of the successful theory. Stanford's truth-linked story seems to fall between two stools. His explanation of the success of each and every individual theory T_i is not couched in terms of T_i's *own* truth, and hence it seems to fall within the antirealist camp. Yet, the success of each and every individual theory T_i is explained by means of a relational property of T_i, viz., its predictive similarity to the "true theoretical account of the relevant domain." (275) The presence of this relational property in the explanans of the success of a theory T_i commits Stanford to the existence of a true theory of a certain domain and hence makes theoretical *truth* relevant to the explanation of T_i's success—even if T_i's own theoretical truth is not, according to him, relevant. So, it seems that Stanford's account could also fall within a—suitably refined—realist camp.

Indeed, there are further reasons to worry about the suitability of Stanford's account for antirealists. He takes an "Epistemic Antirealist" to be, in effect, a skeptical realist. (267) This philosophical character claims that "there *is* always some true theory of a given scientific domain in [the] correspondence sense, but . . . 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." (ibid.) Note that even if Stanford's explanation of the success of science was adequate, it would be fit only for the foregoing skeptical realist. So, all it could possibly show is that if one was an antirealist of this specific sort, one could have an alternative explanation of the success of science.

Let us, for the time being, assume that Stanford has offered an explanation of the success of science. Typical arguments from the success of science—be they realist or antirealist—take the weight of explanation of the success of a theory to be carried by some relation (objective: truth, empirical adequacy; or quasi-objective: the worlds behaves 'as if' the theory were true) between the specific theory and the world. Stanford is quite explicit that he wants the weight of explanation to be carried by a "relationship of predictive similarity *between two theories,*" one being the the-

ory whose success needs explanation, the other being "the true theoretical account of the relevant domain"—what I have called T. (276; his emphasis) Given a correspondence theory of truth, which he amply grants to his antirealist, there is traffic between the true theoretical account of the relevant domain and the world (the particular domain, to be specific). So, to say the least, it is an overstatement that Stanford's own explanation depends on the relationship between two theories. The "true theoretical account of the relevant domain" automatically licenses access to the world—on a correspondence account of truth—and hence Stanford's explanation reduces to a relationship between a theory T_1 and the world (as this is captured by "the true theoretical account of the relevant domain"). The only complication he adds is that the predictive success of T_1 is not directly explained by its own relation to the world, but by the (correspondence) relation to the world of a theory T that is predictively similar to T_1. All the same, the predictive success of T_1 is indirectly explained by T_1's own relation to the world, since it is a fact about T_1's own relation to the world that T_1's predictions are also predictions of the theory that corresponds to the world. The world is not lost, and this is the realist's world. So, on Stanford's account, it is the world—via the true theory T—that carries the weight of the explanation of the successes of each and every theory predictively similar to T. All this might make us feel uneasy as to how much service to antirealism Stanford has done.

Alternatively, it may be that Stanford's account is distinctively antirealist because predictive similarity to truth collapses to some standard antirealist story: a theory predictively similar to the true theory is an empirically adequate theory, an 'as-if true' theory etc. Stanford considers this possibility and says that even if this collapse is inevitable "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." (276 note 11; his emphasis) I'll leave it to the antirealists to engage in domestic disputes as to whether Stanford's account does (or should) indeed collapse to more standard ones. My own worry is twofold: first, as I will argue shortly, I doubt that Stanford's story is genuinely explanatory. Second, even if it was, we would still want to know why it "constitutes . . . a demonstrably appropriate terminus for the chain of explanatory demands." (276)[2] But more on this in a moment.

2. Stanford takes the view that the predictive similarity of T_1 to T is an "intrinsic" property (feature) of T_1 (cf. 276–277). This is quite puzzling, since predictive similarity is a relational property. If, as I suspect, predictive similarity to T is an *extrinsic* property of T_1, then there is a further legitimate task to look for an intrinsic property that explains the success of T_1 (be it truth or empirical adequacy or what have you). So, Stanford's account—even if genuinely explanatory—is incomplete.

3. The Detour via T. Antirealists should not just assume that successful theories have been false. In order to claim their outright falsity they need an independent argument.[3] So, they should leave open the issue of whether they are true or false, and try to argue that truth is not required for the explanation of their success. I am stressing this tedious point because Stanford seems to waver a bit. The overall spirit of his antirealist argument is, along the lines I mentioned, viz., it tries to offer a general explanation of the success of scientific theories, leaving open the issue of their truth or falsity. When, however, it comes to the specifics, he restricts his attention to false-but-successful theories. So, he says that "there is an explanation available . . . to . . . an Epistemic Antirealist for the success of our scientific theories." (267) But he formulates his story with a reference to *false* theories: "*the success of a given false theory in a particular domain is explained by the fact that its predictions are (sufficiently) close to those made by the true theoretical account of the relevant domain.*" (275; his emphasis) This shift might be explained by his general view that if a theory were true, its truth would explain its success. So, it seems that the pressure for an explanation is acute only for false theories. But, given that he tries to make a case for an "Epistemic Antirealist"—who is, in fact, a skeptic about the truth of scientific theories (cf. 267)—his account must be unrestricted: it must apply to any theory, either false-but-successful or not-known-to-be-true-but-successful. In this section, I shall examine his alternative explanation as applied to an arbitrary successful theory T_1 and then—in the next section—I shall restrict my attention to false theories.

My prime point is that, far from being an adequate explanation of predictive success, Stanford's account just *is* an assertion (or re-statement) of the skeptical attitude encapsulated in his "Epistemic Antirealist." Take a theory T_1 which is predictively successful, but not known to be either true or false. Pressed to explain this success, Stanford's antirealist doesn't want to appeal to T_1's own (approximate) truth; on the other hand, she doesn't want to make this success a fluke. So, for each predictively successful T_1, she posits a true theory T of the same domain and claims that what explains the success of T_1 is the predictive similarity between T_1 and T. What, then, is really remarkable in Stanford's account is that we should not appeal to the truth of any specific theory T_1 to explain T_1's own success—the explanation has to go via a predictively similar true theory T.

But note that an appeal to the truth of T can (and does) explain T's own success. For the true theory T of a domain is trivially predictively

3. This might well be the role of the argument from the Pessimistic Induction. But a) it needs independent support; and b) its credentials have been recently challenged by many realists (cf. Kitcher 1993, Psillos 1999).

similar to itself and hence, on Stanford's account, T's own success is explained by T's own truth in a fully realist way. What is the relevant difference between a specific theory T_1 and the true theory T in virtue of which the truth of T_1 cannot be used as an explanation of T_1's success, while the truth of T can be used as an explanation of T's own success? It can't be, at this stage at least, that T_1 is false. The skeptic is not entitled to the assertion that an arbitrary T_1 *is* false. Given that it is an open issue whether T_1 is true or false, the relevant difference can only be that T_1 is doubted (not known) to be true, whereas T is not (it is by default *the* true theory). If we take this difference seriously in our attempt to ground the asymmetry between T_1 and T, we merely reiterate the skeptical position associated with Stanford's Epistemic Antirealist: what forbids us to explain T_1's success by means of its own (approximate) truth is that T_1 is not known to be true. (Alternatively, that T_1 may be false.)

Could Stanford reply that the predictive similarity between T_1 and T does, nonetheless, explain T_1's success? I think there is no explanatory advance here. There are three reasons why an abstract appeal to an unfathomable true theory T cannot carry the weight of the explanation of the success of T_1. *First*, the role of T is epistemic. As we have seen, T_1 would carry the weight of the explanation of its own success, had it been (known to be) true. T simply covers for the skeptic's unwillingness to make a direct link between explanatory success and truth, unless the truth is certified. *Second*, for an antirealist, it would be just enough to note that T_1 has hit upon relevant universal regularities in order to explain its success, without going via T. Stanford thinks that his own story explains why T_1 has hit upon universal regularities, while a standard antirealist story just assumes that. Yet—from the antirealist point of view—he also just assumes that there is a true theory and that it is such that it entails correct descriptions of the relevant regularities. *Third*, from a realist point of view, the detour via T is simply an expression of excessive caution. Unless there are specific reasons to doubt T_1 (or to consider it false), the weight of the explanation of its success should be carried by the (approximate) truth of T_1 itself. Perhaps, all is not lost for Stanford. Perhaps, the detour via T gives some flimsy ground to his antirealist to claim that since T_1 was successful, its success must not be a fluke. But admitting that T_1's success was not a fluke amounts to nothing like an adequate (let alone sufficient) explanation of its success.

While we are at it, let me also deliver on the second of the promises made in the end of the previous section. Does the appeal to predictive similarity to T constitute "a demonstrably appropriate terminus for the chain of explanatory demands"? (276) I think there is always another question to ask, to which Stanford owes us an answer: is the predictive similarity between T_1 and T a brute fact? If it isn't there is further need

for explanation of this fact. If it is, then we still need to be told a story as to why this brute fact should be posited and taken seriously.

4. In Defence of Smart. Let's now concentrate on Stanford's more specific thesis, viz., that his account explains why *false* theories have been successful. In his argument, Stanford (273) appeals to the following point by Smart:

> 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. (Smart 1968, 151)

Stanford says that "we should . . . 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)." (274) But he disagrees with the "general moral" that Smart draws from this case. For Stanford

> the actual *content* of the Copernican hypothesis plays *no role whatsoever* in the explanation we get of the successes of the Ptolemaic system: what matters is 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. (274; his emphasis)

So, he insists, in explaining the success of Ptolemy's theory we can drop out any reference to the actual content of the Copernican theory, but we cannot "drop out just the information that the Ptolemaic system is predictively similar to the true account [whatever this may turn out to be] without undermining our answer's explanatory value." (275) Hence, he concludes, it is enough to posit the existence of a predictively similar *true* theory of the relevant domain instead of talking about the truth of any specific rival theory in order to explain the success of a false theory. Is he right?

For a start, I think that Stanford misreads the central message of Smart's point. Ptolemy's theory was successful (instrumentally useful) as

opposed to being unsuccessful. Given that its falsity is not enough to explain this difference (false theories can be both successful and unsuccessful), there must be some appeal to some other feature to explain this difference and hence to explain why a false theory can be instrumentally useful. Being a realist, Smart goes for the truth of the *Copernican* theory to explain this difference. How is this done? By its own theoretical arsenal, the Copernican theory explains and predicts the relevant regularities (in particular those that lent credence to Ptolemy's theory). So, the truth of the Copernican theory explains the theory's *own* instrumental usefulness. But then any other theory—in particular any other false theory—that generates the same predictions is bound to be instrumentally useful (since any such theory is bound to be predictively similar to the Copernican).[4]

Now, were we to drop "the actual *content* of the Copernican hypothesis," as Stanford suggests, we would be left with no (first-order) explanation of the regularities that led credence to the Copernican (and the Ptolemaic) system. Nor would we explain why the *Copernican* theory was instrumentally useful. We would just state that it was instrumentally useful. Consequently, we wouldn't explain why a theory which was predictively similar to the Copernican was also instrumentally useful. I think this point needs to be stressed. The content of the Copernican theory is employed to explain the success of the Copernican theory itself and hence to explain why any false theory, in particular Ptolemy's, which is predictively similar to the Copernican had to be successful and instrumentally useful.[5]

In any case, as noted in the previous section, an abstract appeal to the existence of a true theory T of the same domain does little to genuinely explain why a *false* theory was successful. A recurring concern is that if we are just told that the false theory is predictively similar to the true theoretical account of the relevant domain, we are merely assured of the fact that its success should not be taken to be a fluke. We don't thereby explain it. Positing predictive similarity to the true theory T is no less

4. In support of my suggested interpretation of Smart's main message consider the following. Smart ends his point by saying: "Such an explanation of the instrumental usefulness of certain theories would not be possible if *all* theories were regarded as merely instrumental." I take this to imply that unless *some* instrumentally useful theories were taken to be true (and hence, I add, they were taken to explain their own instrumental usefulness), there would be no explanation of instrumental usefulness. In fact, a few lines later Smart adds: "If we have nothing but instrumental laws, they may explain in the sense of enabling us to predict, but they do not explain in the sense of reducing the brutishness of the facts." (1968, 152)

5. Of course, the Copernican theory is, strictly speaking, false. But to say that it is false is not to say that its actual content can be dropped out in the explanation of its own success. After all, the Copernican theory is a) approximately true and b) truer than Ptolemy's.

mysterious than just taking the success of a false theory to be a brute fact. Naturally, Stanford disagrees. He notes that "it is inappropriate to ask what further characteristic *of the theory* accounts for or explains its predictive similarity to the truth." (275; his emphasis)[6] But calling this question "inappropriate" is premature. We can certainly ask this question for the true theory. It would be totally pointless to say that what explains the success of the true theory is its predictive similarity to itself. It would be more acceptable to say that the further characteristic of its own truth is what explains its success. In particular, it would only help to cite the content of the true theory. But then why not look for a further characteristic in false-but-successful theories? That they are false does not imply that they are not, say, approximately true in ways in which their predictive similarity to truth is to be expected.

To be sure, Stanford (273–274) insists that an appeal to some sort of "structural similarity" between the false theory and the truth (or a truer theory, anyway) adds nothing further to the explanation of the false theory's success in terms of predictive similarity. He bases this contention on the claim that "such structural similarity is so easy to come by as to be explanatorily vacuous." (273–274) I doubt that such structural similarity is so pervasive among different theories, especially if we do not just restrict our attention to similarity of mathematical content. But suppose it is. Just because something is easy to come by, it does not mean that it is explanatorily vacuous. It might be easy to find (structural) similarities in the behavior of people in a certain group (e.g., football fans or professional philosophers). But this does not mean that it's explanatorily vacuous to cite these similarities as an explanation of their shared actions. Even if there is more to be said by way of explanation, what is already said is not vacuous. Similarly, it might be easy to find structural similarities between theories, but this does not show that it is explanatorily vacuous to cite these structural similarities in order to explain their predictive similarity.

There is a general reason, I think, why Stanford's appeal to predictive similarity to truth fails. Because of the properties of predictive similarity, it leads to odd (and counter-intuitive) results. Here is an example. Suppose there are two false theories T_1 and T_2 (say Ptolemy's and Copernicus's) which, before we read Stanford, we take to be predictively similar to each other. Suppose also that we follow Stanford in explaining the success of T_1 in terms of a true theory T and the success of T_2 in terms of the very

6. Relatedly, he notes: "asked *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 predictions arise from the mechanics of the theory, or greet her with a puzzled look and a shrug." (273)

same true theory T. So, we state that (A) "T_1 is predictively similar to T" and (B) "T_2 is predictively similar to T." But predictive similarity is symmetric, and hence we can reverse, say, (B): "T is predictively similar to T_2"—call this (B'). Predictive similarity is *not* transitive, since no similarity is.[7] Hence, from (A) and (B') we *cannot* infer that T_1 is predictively similar to T_2. If we rely on predictive similarity to truth, we cannot recover the obvious judgment that T_1 (Ptolemy's theory) is predictively similar to T_2 (the Copernican theory). Here is another example. We would, intuitively, be willing to say that since Kepler's theory (T_1) is predictively similar to Newton's (T_2), then, if Newton's theory is predictively similar to the truth so must Kepler's be. But given the properties of predictive similarity we can't. For if T_1 is predictively similar to T_2, and T_2 is predictively similar to the truth (T), we cannot infer that T_1 is predictively similar to T. Do you want more? On Stanford's account no subsequent theory of a domain can explain the success of its predecessors—even if intuitively they do. Suppose—as we would intuitively—that Newton's theory explains the success of Kepler's and that Kepler's explains the successes of Copernicus's. Then, just because predictive similarity is *not* transitive we cannot infer that Newton's theory also explains the successes of the Copernican theory.[8] In fact, Newton's theory does not even explain the success of Kepler's, on Stanford's account. For even if Kepler's theory is predictively similar to the true theory T, and even if Newton's theory is predictively similar to T, it does not follow that Kepler's theory is predictively similar to Newton's. So, Newton's theory cannot even be a candidate for the explanation of the success of Kepler's theory.

7. Obviously, I take similarity to be different from identity. As a perceptive referee noted, similarity relations in mathematics are (as part of their definition) transitive, e.g., when we say that two sets are cardinally similar to each other.

8. These problems would, of course, not arise if we talked in terms of predictive *identity*, since predictive identity is transitive. But I think predictive identity is a too strong (and implausible) notion to rely on. It may be that in the Ptolemy-to-Copernicus case the relevant predictions are identical. But it is hard to see how this case can be generalized to cover, for instance, the transition from Newton's theory to Einstein's. As the referee mentioned in the previous note pointed out, if we define predictive similarity to require only that there be some limiting process that connects the predictions of two theories, then transitivity would be restored. For instance, since Einsteinian predictions tend to Newtonian ones as some parameter P tends to zero, and if we suppose that a new Zweisteinian theory gives Einsteinian predictions as some second parameter P' tends to zero, then we do have transitivity in that Zweistein's predictions tend to Newton's as both P and P' tend to zero. All this is correct. But note that a) Stanford does not say at all precisely what "predictive similarity" means; and b) in order to restore transitivity by some kind of limiting process we need to appeal to the *content* of the relevant theories, and in particular to some structural similarities that make possible these limiting processes—for instance that we can derive a limiting case of Newton's second law within Einstein's theory.

Stanford might be happy with an image of science according to which the success of each theory is explained afresh in terms of its own predictive similarity to the truth. So, we start with the Copernican theory and explain its success in terms of its predictive similarity to the truth, then we come up with Kepler's theory and do the same for it, then we devise Newton's theory and do the same with it, and so on. Isn't it then miraculous that all these theories are false, utterly disconnected with each other (so that we cannot talk about their being approximately true, nor about their structural similarity) and yet they all are predictively similar to the truth? Well, isn't it?[9]

5. Conclusion. I have said nothing positive about the realist explanation of the success of science.[10] But if my arguments against Stanford's antirealist alternative are sound, then yet another challenger to the full realist story fails to discredit it. Stanford himself has offered good arguments against some other standard antirealist alternatives to the realist explanation. So, leaving aside some other problems that it faces (e.g., the charge of circularity), the realist story—and the 'no miracle' argument in particular—withstands yet more pressure.

REFERENCES

Kitcher, Philip (1993), *The Advancement of Science*. Oxford: Oxford University Press.
Psillos, Stathis (1999), *Scientific Realism: How Science Tracks Truth*. London: Routledge.
Smart, J.J.C. (1968), *Between Science and Philosophy*. New York: Random House.
Stanford, P. Kyle (2000), "An Antirealist Explanation of the Success of Science", *Philosophy of Science* 67: 266–284.

9. In Section 4 of his paper, Stanford says that "it is worth noting" that his account is "fully adequate to explain the ability of theories to make successful, novel predictions." (277) But, surprisingly, he offers no argument at all. Instead of tiring the reader further, I just note that most of what I have said so far applies equally well to what he says about novel empirical success.

10. For this see Psillos 1999, Chapter 4.