CAN A THEORY'S PREDICTIVE SUCCESS WARRANT BELIEF IN THE UNOBSERVABLE ENTITIES IT POSTULATES?

Many theories posit entities that cannot be directly observed: atoms, quarks, magnetic fields, genes, mental representations, and so on. By their very nature, it is never possible to confirm the existence of such entities by direct observation. (There is, however, a question about what counts as a "direct" observation – is looking through a microscope allowed?) Nonetheless, theories that posit such entities often do make predictions that can be tested by observation, and some theories are highly successful in this enterprise. Does this sort of predictive success give us reason to believe that the posited entities really exist? One argument, often dubbed the "miracle argument," claims that the empirical success of a theory would be a *miracle*, or at least a coincidence of cosmic proportions, if the theory were wrong about the basic entities underlying the phenomena in question. Jarrett Leplin and also André Kukla and Joel Walmsley, reject the miracle argument in its broadest form. Nonetheless, Leplin argues for the realist position that we are sometimes warranted in believing in the existence of the unobservable entities postulated by science. First, there is a certain presumption in favor of belief in these entities, and none of the standard anti-realist arguments are successful in dislodging this belief, once admitted. Secondly, one specific kind of predictive success, the prediction of *novel* phenomena, *does* provide particularly strong warrant for the belief in unobservables. Kukla and Walmsley criticize this argument, and other refinements of the miracle argument, on the grounds that they presuppose a certain conception of what it would be to *explain* the predictive success of a scientific theory. They arrive at the strong conclusion that no version of the miracle argument could possibly give us grounds for believing in unobservable entities.

CHAPTER
F I V E

A Theory's Predictive Success can Warrant Belief in the Unobservable Entities it Postulates

Jarrett Leplin

5.1 The Burden of Argument

Theoretical entities are the unobservable entities that scientific theories posit to explain or predict empirical results. Inaccessible to experience, their claim to conviction derives from the acceptability of the theories in which they figure. Electrons, fields, and genes are examples normally thought to be real entities. Phlogiston and the electromagnetic ether, though once confidently embraced, have turned out not to be real. The fate of the strings and gravitons advanced by theories at the frontiers of physics is unresolved. The ontological status of theoretical entities is frequently uncertain, disputed, revised. The existence of such an entity may be denied because, with changes in theory, something different, with a related explanatory role, is thought to exist instead. Thus oxygen replaces phlogiston in the chemical theory of combustion. Or an entity may be rejected because a new theory denies it any continuing explanatory utility. The crystalline spheres of ancient and early modern astronomy are simply obviated by Newtonian gravity. In general, the question within science is not *whether* theoretical entities exist but *which* theoretical entities exist. Answers to this question change with the fortunes of theory.

Scientific realism is the position that this question can be answered on evidentially probative grounds. Many posited entities turn out not to exist and the status of many others remains unsettled. But in some cases science develops, through the testing and application of its theories, adequate reason to believe that certain theoretical entities are real. Further, according to realism, the success of theories warrants some beliefs about the nature – the properties and behavior – of these entities. For, as they are unobservable, the mere assertion of their existence, without an account of their nature, is insufficient to serve their explanatory and predictive purposes. All such realist beliefs are defeasible; new evidence could undermine them, as new evidence fre-

quently forces changes in theory. Nevertheless, according to scientific realism, such beliefs are epistemically justifiable.

Anti-realism claims, as a matter of philosophical principle, that there can never be adequate reason to invest credence beyond the range of the observable. Anti-realism regards all theoretical entities, regardless of the ontological status assigned to them within science – regardless of the available evidence – as conceptual tools whose roles are pragmatic, not epistemic. Anti-realism need not deny the reality of unobservable entities; it denies that their reality is ever warrantedly assertable by us. Anti-realism proclaims a sweeping agnosticism with respect to theory, independently of the evidence used to evaluate theories. Scientific evidence attests not to the truth of theories, nor to the existence of the unobservable entities they posit, but only to their explanatory and predictive utility.

If the dispute between realism and anti-realism is a dispute between science and philosophy, anti-realism loses. There is no *a priori* stance from which philosophy can presume to dictate the standards and methods for acquiring knowledge. How knowledge is best acquired depends on the nature of the objects of knowledge, on what the world is like, and is therefore itself knowledge to be acquired, in the way that any empirical knowledge is acquired, through scientific investigation of the world. Abstract reasoning, conceptual or linguistic analysis, appeal to common sense or intuition – or any distinctively philosophical mode of inquiry – are notoriously unreliable as a determinant of the nature or scope of scientific knowledge. Nor does autonomous philosophy deliver a consistent verdict to compare with the conclusions that science reaches. No settled, dependable method of appraisal, such as operates within the natural sciences, is available to adjudicate among the indefinite number of competing positions that philosophers fashion.

Some realists have contended that their dispute with anti-realism could be won on just this basis. Realism sides with science; comparing the progress of science with the state of philosophical inquiry, the attraction of siding with science is an advantage that no philosophically generated anti-realism can overcome. This is essentially what realists are saying when they claim that but for their position science is mysterious, its successfulness an unprojectable accident. To reject realism, they suggest, is to reject science itself. Science's detractors gleefully agree. They think they can infer from their rejection of realism that the success and progressiveness of science are illusory, that its epistemic status is no better than that of any other social institution or practice.

Unfortunately, the resolution of the dispute over realism does not reduce to one's attitude toward science. Essentially, this is because science and philosophy are not autonomous. Philosophical assumptions are ineliminable from the reasoning by which science fixes its ontological commitments. And substantive scientific results often support philosophical limitations on science. It is open to the anti-realist to contend that realist beliefs cannot be read off the record of what science achieves, but must be read in via a certain, optional, philosophical interpretation. Science can, or even does, operate just as well without such beliefs.

For example, science operates just as well with a system of fundamental equations that admit of no coherent physical interpretation as it does with conceptually tractable laws, for no physical commitments at the theoretical level follow in any case. From a (consistent) anti-realist perspective, the quest for an understandable interpretation

of quantum mechanics is misdirected. The anti-realist therefore disagrees with theoretical physicists as to the importance of interpreting quantum mechanics, even as he insists that his philosophy is consonant with scientific practice.

But this rejection of realism's scientific pedigree concedes that the priority of science would favor a philosophical position that science really did require. And this concession places the argumentative burden on anti-realism. Science claims to discover and to learn the nature of certain theoretical entities. "Electron," for example, is a purportedly referential term, and properties of electrons – mass, charge, and spin – are held to be well established. According to anti-realism, all claims to discover or learn such things are mistaken. Given the priority to which science is entitled, the opening move in the debate should be to argue that its characteristic theoretical hypotheses cannot be warranted.

Accordingly, I elect to begin with anti-realist arguments. In subsequent sections, I shall consider two significant challenges to realism. Although unsuccessful, they may be judged sufficiently compelling to shift the burden of argument back upon the realist. Therefore, I shall follow with development of an independent argument for realism. In summary, I contend that realism is the default position, and that the case to be made for switching to anti-realism is at best indecisive. At the same time, a compelling defense of realism is available.

5.2 Underdetermination

A major source of anti-realist argumentation is the simple, incontestable fact that observational evidence is logically inconclusive with respect to the truth of theory. That is, the falsity of any theory T is logically consistent with the truth of all observation statements O_i used in assessing T. T is, in this respect, *underdetermined* by the evidence; indeed, by all possible evidence. It follows that in affirming T on the basis of evidence, one must reason ampliatively; one must use forms of reasoning that are not truth-preserving. Eliminative induction, a kind of process of elimination in which the range of potential contending hypotheses is unrestricted, is such a form, as is inductive generalization. Another is abduction, in which the explanatory power of a hypothesis counts as evidence for its truth. In practice, scientific argumentation exhibits all the rationally cogent ampliative forms that philosophers have identified.

But this reliance on ampliative reasoning is not immediately an objection to affirming theories. For such reasoning is endemic to, and ineliminable from, ordinary inference that grounds common-sense beliefs about the observable world. Without it one could not ground even the belief that a plainly observable object continues to exist when unobserved. This belief is certainly abductive. Without ampliation, one gets not anti-realism about science but a sweeping skepticism that no party to the dispute over realism accepts.

The question for realism, therefore, is whether, *allowing* the rational cogency of standard forms of ampliative inference, theories are *still* underdetermined. It is usual to formulate this thesis of ampliative underdetermination, the only kind worth considering, by asserting the existence of rival theories to T that are equally well supported by the evidence. It is then said that theory-*choice* is underdetermined: a preference for T over its rivals must have some nonevidential, pragmatic basis.

Is ampliative underdetermination credible? Notice that the usual formulation of this thesis is not its minimal formulation. Minimally, the claim is that no body of observational evidence $\{O_i\}$ warrants any theory T; neither truth-preserving inference nor rational modes of ampliative inference can get you from $\{O_i\}$ to T. The existence of evidentially equivalent rival theories is a *further* assertion. It is easy to see why this further assertion is made. Without it, there is no particular reason to suppose, quite generally and abstractly, that T is not rationally reachable. Embedded into the very formulation of underdetermination, then, is an argument for it: there will always be alternative theoretical options that fare as well on the evidence as T does.

But now we discern a serious equivocation. Once the legitimacy of ampliation is conceded, the existence of the evidentially equivalent rivals can have no logical or otherwise *a priori* guarantee. For the sort of rivals to T whose existence the logical gap between $\{O_i\}$ and T guarantees are surely not defensible by ampliation. For example, the construction $\sim T \& \Pi\{O_i\}$, while both inconsistent with T (and to that extent a "rival") and consistent with the evidence for T, is certainly not supported by this evidence. ($\Pi\{O_i\}$ is the conjunction of all the propositions in the set $\{O_i\}$.) While the O_i might fall short in warranting T – this is the possibility that underdetermination declares realized invariably – the O_i can hardly be supposed to warrant $\sim T$. The situation at issue is not one in which T faces counter-evidence or disconfirmation, but one in which the evidence bearing on T is allowed to be as supportive as one likes; the claim is that *even then* T is underdetermined because some rival is supported equally well. $\sim T \& \Pi\{O_i\}$ is merely logically consistent with the evidence; it is not *supported* at all. What would support $\sim T$ is the *negation* of some O_i. Thus an immediate reason to declare $\sim T \& \Pi O_i$ unconfirmable is that confirmation of its first component requires refutation of the second.

This argument assumes that if the O_i support $\sim T \& \Pi\{O_i\}$ then they support $\sim T$. What makes this assumption reasonable is not some holistic conception of confirmation according to which any evidence for a theory supports equally, or even to any extent, all distinguishable components of the theory or all consequences of the theory. Such holism is implausible on its face and proves detrimental to realism. Rather, in the particular situation depicted, $\sim T$ is the only bit of theory around to be the subject of support. The O_i are suppliers of support, not objects of it. It makes little sense to speak of a relation of support between $\{O_i\}$ and itself. It could only be in virtue of supporting $\sim T$ that the O_i support $\sim T \& \Pi\{O_i\}$, as they have been hypothesized to do.

Of course $\sim T \& \Pi\{O_i\}$ makes empirical commitments and is truth-valuable. Why isn't this enough to make it confirmable? Although the O_i are not self-supporting, are they not confirmed by the facts? That is, do not the observations, as opposed to the observation sentences that report them, confirm $\sim T \& \Pi\{O_i\}$?

The answer is negative. We are not concerned with choices among rival theories that the evidence refutes. The underdetermination thesis applies to theories that the evidence supports; it is these that are supposed to have equally supported rivals. Therefore, the O_i must either be supposed true or supposed to instantiate generalizations of supposed truths if the supposed rival committed to them is even to be formulated. And a theory is not confirmed by an observation presupposed in its very formulation. Observational results obtainable from a theory only by defining the theory to include them do not support the theory. For it is not even logically possible for the

theory to get them wrong, whereas a result can support a theory only if it is unlikely to be obtained if the theory is false.

Accordingly, $\sim T \& t \Pi\{O_i\}$ is not confirmed by $\{O_i\}$. The only possible epistemic route to $\sim T \& t \Pi\{O_i\}$ is indirect, via some further theory inconsistent with T that delivers all of T's observational consequences. But this further theory cannot be generated algorithmically, as $\sim T \& t \Pi\{O_i\}$ was generated, by operating on T. Not even the existence of such a theory, let alone its confirmability, has any *a priori* guarantee.

Strictly speaking, of course, the anti-realist's claim is not that the O_i confirm $\sim T \& t \Pi\{O_i\}$, but only that they support $\sim T \& t \Pi\{O_i\}$ as well as they do T, that T and its rival are *equally* confirmed by the evidence. This common amount of confirmation cannot, however, be zero without repudiating ampliation. More generally, it is a reasonable constraint on theories that the thesis of underdetermination invokes as rivals that these be at least *amenable* to evidential support. A propositional structure that could not in principle be confirmed violates this constraint. Such a structure, crafted solely for logical consistency with the observational consequences of an existing theory, will not be entertained as an alternative to the existing theory because there is nothing to be done with it; it is rightly dismissed as dead in the water because its only possible support is derivative from some further, independent theory that would itself be the proper object of confirmation. Once theories are required to be defensible by ampliation, conditions such as confirmability in principle, explanatory power, and generality – conditions that standard forms of ampliative inference select for – become reasonable constraints on theoretical status.

Similar problems befall other candidates for T's rivals. Let T' affirm that T holds whenever observations are made but not otherwise, the universe instantaneously reverting, when observations recommence, to the conditions that, according to T, would have then prevailed had observation been continuous. Certainly T' is unconfirmable. It has a confirmable component, but not only do confirmations of this component not confirm the other component; *nothing* could confirm the other component. So again, it does not take holism to rule T' unconfirmable. But if T' is unconfirmable, it cannot be equally well supported as T, which *is* confirmable. Of course, T, as much as T', carries a commitment to conditions that prevail in the absence of observation. But T's commitment, unlike that of T', is ampliatively defensible; stability is a paradigm of ordinary, unproblematic ampliation. Moreover, the component T^*, common to T and T', that T holds during observation, is confirmable and exempt from the underdetermination that T' is supposed to create. T' offers no rival to T^*.

The equivocation of the underdetermination thesis, then, is this: while only the ampliative form of underdetermination challenges realism, it is the merely logical form of underdetermination that supplies the rival theories invoked to establish the underdetermination thesis. For the challenge to get off the ground, there will have to be some substantive, independent basis for supposing the rivals to exist at all.

What could this be? There are examples of theoretical rivalries that resist adjudication, but they do so for reasons that presuppose a substantial body of theory that further evidence could undermine. Thus, the possibility of an epistemically principled choice cannot be precluded. Within Newtonian theory, rival attributions of motion to the center of mass of the universe are unadjudicable, but Newtonian theory could prove to be wrong about the detectability of absolute motion. Any theory that both

fixes one of its parameters and declares it unmeasurable, or defines a parameter but leaves it unspecified, generates rivals. But these rivals share the theory's fallible ontological and nomological commitments.

Even if such examples were compelling, argument by example is unlikely to motivate a perfectly general thesis about all theories and all possible bodies of evidence. Moreover, there are plenty of examples that point the other way, cases in which scientists are unable to produce even a *single* theory that makes sense of the mystifying empirical regularities their experiments have revealed. Astronomy has many such anomalies, such as the motion of stars at the periphery of the Milky Way, which is greater than known gravitational forces allow, and the "great chain" of galaxies, which violates the large-scale uniformity required by big bang cosmology. Far from unavoidable, a multiplicity of theoretical options may not even be the norm.

I submit that any challenge to realism from underdetermination is vastly underdetermined.

5.3 Superseded Science

Lots of successful theories turn out to be wrong; the entities that they posit nonexistent. Why is currently successful science any more entitled to credence than the once successful science we now reject? Does the history of theorizing not provide ample reason to distrust theories, regardless of the evidence that supports them?

An obvious answer is that past theories ultimately proved unsuccessful whereas current theories have not. This is why their temporal status differs. But with further testing and theoretical developments, might current theories not prove unacceptable, just as their temporarily accepted predecessors did? Indeed they might; realism admits the defeasibility of all theory. The question is whether there is reason to forecast this development. The major fundamental theories of current physics – general relativity, the basic laws of quantum mechanics, the standard model of elementary particles – are certainly the most severely tested theories ever, and they have proved flawless to a precision uncontemplated in the assessment of their predecessors. These theories are not thought to be the final word; their very multiplicity reflects limitations that a more fundamental, unifying theory will overcome. But there is no reason to expect a unifying theory to require their rejection. Why should the fact that earlier theories failed count against the different, better-tested theories that we have now?

The issue here is the status of second-order evidence. The challenge to realism is that our methods of developing and evaluating theories, our standards for investing credence, are demonstrably unreliable. They have led us to judge Newtonian gravity, phlogistic chemistry, and the electromagnetic ether to be firmly established, as well confirmed as a theoretical commitment could be expected to be. In the nineteenth century, Maxwell famously considered the ether the best-confirmed theoretical entity in natural philosophy. Lavoisier declared the material theory of heat to be no longer a hypothesis, but a truth (*Mémoires de Chimie*, vol. 1, section 2). The phenomena of heat, wrote Lavoisier, are inconceivable without "admitting that they are the result of a real, material substance, of a very subtle fluid, that insinuates itself throughout the molecules of all bodies and pushes them apart" (*Traité de Chimie*, vol. 1, sections

1–3). Chastened by such misjudgments, we distrust the methods that licensed them. Distrusting our methods, we distrust the theories that they now recommend, however much these theories excel under them. First-order evidence supports current theory, but there is second-order evidence against reliance on first-order evidence. The priority of second-order evidence challenges realism.

But why is second-order evidence privileged? As there could be no second-order evidence without first-order evidence to learn from, the relation would appear symmetric. We cannot very well infer from the conclusions to which first-order evidence leads that the conclusions to which it leads are untrustworthy, for if they are untrustworthy then they are no basis for inference. Yet the trustworthiness of conclusions we draw from first-order evidence depends on the evidential warrant of the standards of evidence and modes of inference used in drawing them. If the trustworthiness of each level of evidence is presupposed in assessing that of the other, neither is privileged.

Giving priority to second-order evidence raises quite general problems in epistemology. Not only do theories prove wrong; ordinary, paradigmatically justified beliefs that ampliative reasoning must be allowed to license prove wrong. That systems of ordinary beliefs have proven to contain errors is second-order evidence for the erroneousness of current belief systems, none of whose component beliefs is currently individually impeachable. Am I to induce, from my record of fallibility, that some of my present beliefs are false, although the evidence favors each of them and I have no grounds to doubt any? If so, I am lodged in paradox. For in addition to believing that some of my beliefs are false, I am entitled to believe that all of them are true by the principle that epistemic justification is closed under conjunction. If each of these propositions is justified, so, by further application of this closure principle, is their self-contradictory conjunction, which is absurd.

Partly for this reason, the closure principle for justification under conjunction is disputed within epistemology. But an anti-realism that purports to rationalize scientific practice cannot afford to dispute it. Without this principle, rational inference does not in general transmit epistemic warrant. For in general it is only in conjunction, not individually, that premises provide a basis for inference. And science grows as much by forging new inferential connections – by relating new ideas to what is already known – as by introducing new theories, hypotheses, empirical laws, and experimental results. Inference, often without prospect of independent empirical confirmation, is a frequent basis for extensions of science.

The presumption that second-order evidence trumps first-order evidence gives too simple a picture. Methods of evaluation depend on substantive developments in theory. What we expect of theories responds to what our best theories achieve. As current theories are more severely tested, current methods may reasonably be supposed more reliable. As our knowledge of the world improves, so does our knowledge of how such knowledge is obtained. It might be difficult to prove that methods improve, without assuming, impermissibly, the superiority of current theories. But the burden of argument here is squarely on the anti-realist. A challenge mounted from history is not entitled to presuppose that methods are stable. Unless the anti-realist can show that rejected theories were once warranted by the highest standards that the best current theories meet, the challenge fails.

It is also unclear that the failures of past theories are epistemic failures. Not only do we learn *from* our mistakes; it is an epistemic advance to learn *that we have been mistaken*. That a posited theoretical entity does not, after all, exist, or that a posited theoretical mechanism is not responsible for a certain effect, is important theoretical information. It is not clear that a consistent anti-realism can allow for this information. What makes it any more trustworthy than the information, equally a conclusion from first-order evidence, that a theoretical entity *does* exist? The indispensability of auxiliary hypotheses in generating observable predictions from theories belies the apparent logical asymmetry between verification and falsification. If empirical evidence cannot establish theories, neither can it refute them. For the refutation of a theory requires that theoretical auxiliaries assumed in testing it be independently established.

As much as we see once successful theories rejected, we see once unsuccessful theories resurrected. Why is the anti-realist more impressed by the fall of an admired theory than by the rise of one scorned? Heliocentrism, the vision of ancient Greek astronomers eclipsed by Aristotle, and the checkered history of Prout's hypothesis are as compelling examples as phlogiston and nested spheres. Yet the correctness of the information that a posited theoretical entity does not exist after all is presupposed in pronouncing past theories wrong. There is a certain commitment to the correctness of current theory in impugning past theoretical commitments, for it is current theory that corrects them. Yet the anti-realist impugns past theory *so as* to induce that current theory is unfounded.

Anti-realists have several tactics for finessing this problem. Their essential theme is that inconsistencies among theories – either among historical theories or between a historical theory and current theory – guarantee that, one way or another, there will be examples of successful but failed theoretical entities. Thus we need not assume that current theory is right and past theories wrong; it suffices to assume that not all can be right.

The obvious rejoinder is that appeal to inconsistency is unpromising as a source of data for induction. Inductive strength depends on the preponderance of evidence, whereas inconsistencies necessarily generate evidence in different directions. Less obviously, the anti-realist is to be pushed as to the status of the inconsistencies he discerns. It is easy to take inconsistency for a purely formal relation identifiable independently of substantive empirical judgments. But inconsistencies that relate rival or successive scientific theories are rarely so straightforward. Theories may be thought inconsistent, only to be reconciled by a radical new idea. The anti-realist may be forced to endorse a particular theoretical perspective even to diagnose the inconsistency that he wants to supply his inductive evidence.

More reasonable than to read scientific successes as license to diagnose failures is to regard the record of theory change as constructive. In being rejected, theories are improved upon. Because science has been able to identify its mistakes and rectify them, its current commitments are all the more trustworthy.

5.4 Selective Confirmation

What are these commitments? What, for that matter, are the commitments of any successful theory? I have argued that because the ontological status of theoretical posits is inconstant and disputatious within science, a philosophy that imposes uniformity

assumes the burden of dissenting from science. But if the status of theoretical entities is disputatious, what is one to be realist about? In particular, are the once successful posits, from whose eventual rejection the anti-realist induces the epistemic unreliability of current science, ones to which a realist should have been committed in the first place?

Success has many forms. Not only is scientific appraisal unstable; it is multidimensional. Theories and the entities they posit are assessed for their utility, heuristic power, explanatory value, mathematical tractability, experimental manipulability, and cohesion with background knowledge. Epistemic justification is neither the only, nor necessarily the most pressing, concern. Because of this complexity, the epistemic commitments of science cannot simply be read off of scientific practice. Whether or not the failures of past science are failures of a realist view of that science depends upon realism's criterion for deciding what theoretical posits to treat realistically. Deference to scientific practice is not a definitive criterion.

This problem is intractable if one takes a holistic view of confirmation. All manner of theoretical posits whose roles in a successful theory are not such as to accrue epistemic warrant then go along for the ride. Is the spacetime interval in Minkowski's interpretation of special relativity supposed to be a real entity, to whose existence the empirical evidence for relativity attests, or is it but a mathematical invariance of dispensable convenience? Are quarks a purely formal method of classification for hadrons, or are they their physical constituents? Is the mechanical ether an intuitive aid to picturing what happens in the space between the locations of charged bodies, or is it an entity that Maxwell's equations require to exist?

To treat the empirical success of a theory as confirmation of the theory as a whole is to obviate such questions; it is then the theory as such that evidence confirms, not this hypothesis or this theoretical posit over others. The loss in discrimination may appear innocuous from an overview of scientific practice, because the pragmatic goals of prediction and control are as well advanced by an entity's conceptual utility as by its existence. But the questions are pressing for the realist, who must discriminate entities whose existence is established by the evidence from those that can come or go with impunity.

To do this, the realist must adopt some criterion beyond mere participation in an empirically successful theory to identify theoretical posits to be treated realistically. There are both positive and negative criteria, and some of them are intuitively obvious in the abstract, if problematic in application. For example, a theoretical posit gets no epistemic support from the successful prediction of a result that it was artificially contrived to yield. Nor is it supported by a result that the theory predicts independently, a result the theory does not need it to obtain. The positive criterion would seem to be that the posit be used *essentially* in achieving the theory's empirical success, and that this success be unexplainable without it.

It may have been inconceivable to Maxwell that electrical phenomena could proceed without a mechanical medium to propagate electromagnetic waves. But his equations alone, without the ether hypothesis, generate the predictive success of his theory. Newton believed that the apparent motions, which his laws of mechanics governed, presupposed the existence of absolute frames of space and time. But these posits have no role in the use of Newton's laws to generate observable predictions. The igneous fluid that Lavoisier thought was necessary to push the molecules of a heated

substance apart was not required to account for the phenomena of heat; molecular motion itself was the operative mechanism.

A pattern is evident in such examples. Like the hidden variables of intuitively picturable interpretations of quantum mechanics, an entity or structure is introduced to make physical sense of the laws used to predict empirical phenomena. But the particular properties attributed to this entity do not matter to the use of these laws in successful prediction. The center of mass of the universe is supposed, by Newton, to be at rest in absolute space, but it makes no difference to the use of Newton's laws to give it a positive constant absolute velocity. Maxwell thought there had to be some sort of ether to propagate waves, but he was free to give it all sorts of mechanical properties without affecting his laws. Lavoisier was vague by default as to the physical process by which his igneous fluid flowed. Entities such as phlogiston or the nested spheres of a geocentric universe are rejected because they give the wrong theoretical mechanism. They have identifiable successors in later theories – oxygen and gravity. In contrast, presuppositional posits in the conceptual background of successful laws may be rejected simply because their existence proves inconsistent with subsequent theory. They have no successors because there is no predictive role to continue to fill. We teach ourselves to regard them as metaphysically superfluous.

Thus, conceptual involvement, however fundamental, in a successful theory is not sufficient for the reality of a theoretical entity. What scientists believe the world must be like and how they make sense of the empirical phenomena owe too much to heuristic concepts that underlie their theory and to the entanglement of pragmatic among epistemic ends. Rather, the criterion must be that the theory owes its success to this entity. With this criterion, the realist can reject historical counterexamples of successful but nonexistent theoretical entities, and argue that entities meeting his criterion survive in current science.

5.5 An Argument from Novelty

Despite their inconclusiveness, the challenges to realism have pressured realists into independent lines of argumentation. For anti-realism has its own pedigree. It is embedded within a long-ascendant tradition of empiricist epistemology, and from the perspective of this tradition realism's epistemic commitments are excessive. Empiricism is certainly the philosophical root of anti-realism's indiscriminate suspicion of theoretical entities. If experience is the only possible source of knowledge of the world, beliefs as to the existence and nature of unobservable entities are inherently suspect and require a special defense. The predominant line of defense is explanationist: If theories owe their empirical success to unobservable entities, then we need realism to explain why theories are empirically successful. If there is no truth to theory, if theoretical entities are not real, then the predictive accuracy of theory is a coincidence too cosmic to accept.

The argument I shall construct is a descendant of this line. Lots of empirical success needs no explanation, and lots of it has nonrealist explanation. Lots of theory is unsuccessful. The problem is to identify a specific form of success that realism alone explains, and then to show that this virtue of realism is epistemically justificatory.

The form of success I propose is *novel* success; the successful prediction by a theory of an empirical result that is novel for it.

Intuitively, "novel" means "new," not just temporally, but also in the sense of different or unusual, and "unexpected." These attributes pose an explanatory challenge that it will take realism to meet. How does a theory manage to predict unusual and unexpected results correctly? That its predictions are correct is a matter of experience. The answer to the question *why* they are correct is that this is the way the world is found to be. That they are *its* predictions is a matter of deduction. The theory yields its predictions because of the inferential resources of its semantic content. Realism is not involved in answering these questions, but the question I have posed does not reduce to these. It is a second-order question about the success of the theory. That it be just the correct results that the theory predicts is not, if these results are novel, explained by the fact that the theory sustains certain logical relations to certain observation sentences which simply happen to be correct. The explanation must appeal to some property of the theory, something distinctive in its content that enables it reliably to forecast the unfamiliar and unexpected. Unless this content is interpreted realistically, the theory's novel success appears purely accidental.

To make clear why realism is required to explain the novel success of theories, let me be more precise about novelty. Classic examples of novelty are the prediction from general relativity of the gravitational deflection of starlight, and the prediction from Fresnel's transverse wave theory of light of the bright spot in the center of the shadow cast by a circular disk in spherical diffraction. These results were new, unknown, surprising, unanticipated independently of the theory predicting them, uninvolved in constructing this theory, and unlike results supporting rival theories. They present all of the features intuitively associated with novelty. But some of these features do not require realism and some need not be present for realism to be required. A result could be new and unknown, and yet instantiate a general law presupposed in constructing the theory that predicts it. Then the success of the prediction gives no epistemic support to the theory. A result could be well known yet unexplained, even contrary to the predictions of extant theories. Then its prediction by a new theory carries probative weight. The use of a result in constructing a theory might have been inessential, such that the theory would have predicted the result without its use. Even a theory expressly motivated by the need to explain a result can receive epistemic credit for doing so, if the result is not involved in its construction. The core conditions for an analysis of novelty are those under which novel predictive success depends on the existence of the mechanisms theorized to produce it and the accuracy of their description.

Two such conditions will be jointly sufficient for the prediction of a result R by a theory T to be novel for T. The more basic intuitively is *independence*: R must not instantiate any general law used essentially in constructing T. This captures the idea that R was not built-in, whether expressly or inadvertently, such that T would automatically have predicted R whether T is true or not. Secondly, the prediction must be *unique*: no viable rival of T provides an alternative basis for predicting R. For otherwise, R's evidential status is ambiguous. The uniqueness condition speaks to the intuition that a novel consequence of T must differ from the empirical consequences of other theories.

Classic examples of novelty exhibit these conditions. Young's law of interference could be used to obtain the positions and intensities of diffraction bands, but not Fresnel's predictions for spherical diffraction. And certainly no corpuscular theory of light could be made to yield the unexpected bright spot. A gravitational influence on light could be based on Newtonian theory, but only by suspending major theoretical developments since Newton: the incorporation of the wave theory of light into electromagnetic field theory, and the role of light in relativistic mechanics. No viable rival to Einstein's analysis of deflection was available.

My explanationist argument for realism is now straightforward. A theory's sustained record of novel predictive success is only explainable by supposing that the theory has correctly identified and described the entities or processes responsible for the observations predicted. The realist explanation of the theory's success is then epistemically unrivaled. If this success is uncompromised by failure, if the theory is free of disconfirming results and conceptual problems, then the realist explanation of its success is also epistemically undefeated. But an explanation that is neither rivaled nor defeated is justified, on pain of skepticism with respect to ordinary beliefs whose provenance is unavoidably abductive.

5.6 Consequences and Clarifications

Independence is a historical attribute; whether *R* satisfies the independence condition depends on how *T* was developed. If we imagine *T* having alternative provenances without overlapping reliance on common results, or generalizations of them, then all of *T*'s predictions satisfy independence, for the use of none is essential to *T*'s construction. This is appropriate, for alternative provenances constitute a form of *over*determination, which should be as much of an epistemic advantage as underdetermination is an epistemic liability. A theory's provenance provides reasons for thinking it plausible and taking it seriously as a potentially acceptable account of some domain of nature, pending empirical testing. Additional provenances represent additional grounding.

But, of course, predictions novel for *T* do not automatically support *T*. They must be established by observation with quantitative accuracy. A theory rich in novel consequences is rich in opportunities for epistemic support, none of which might materialize. Independent provenances are rare historically, and even if present are no guarantee that *T* merits realist interpretation.

Although novel for *T*, *R* might come to be explained or predicted by a rival T' of *T* that is developed after the fact. Rather than preempt *R*'s novelty, the effect of the advent of T' is to challenge *R*'s epistemic weight. The uniqueness condition is to be read as temporally indexed; it is not novel status that varies with historical developments but its epistemic significance. This interpretation accords with philosophical as well as scientific practice. The anti-realist wishes to credit superseded theories with empirical success. The most important form of empirical success is novel success, and it must be possible to diagnose such success however further developments affect *T*.

However, T' does not necessarily undermine *R*'s support of *T*. If *R* would not have been novel for T' even if *T* were unavailable – if, for example, T' is expressly designed

to yield *R* and provides no independent theoretical basis for explaining *R* – then *T* is still favored. For the correctness of *T*'s explanation of *R* is still undefeated as an explanation of *T*'s ability to predict *R* successfully.

What if, but for *T*, *R* would have been novel for *T′*? What if *T′* only *happens* to come later, its provenance owes nothing to *T* or *R*, and its prediction of *R* is epistemically as impressive as *T*'s? Then, of course, *R*'s support of *T* is undermined; *R* is no longer a reason to interpret *T* realistically. And if such sequences are historically common, the anti-realist has a new basis for skeptical induction.

Accordingly, realism predicts that such scenarios will *not* occur. It predicts that viable rivals to epistemically warranted theories will not arise. It predicts that theories that record novel success will continue to be successful, that the existence and properties of the theoretical entities they invoke to explain novel results will be upheld through further developments in science. As novel success is realism's standard of epistemic warrant, realism, if it is to be warranted, must itself meet this standard. The importance of the predictions in question is to deliver on this requirement.

No philosophical rival to realism provides a basis for these predictions, nor are they involved in the argument for or content of the realist thesis. Accordingly, they satisfy the independence and uniqueness conditions for novelty. Realism is self-referentially consistent, as any naturalistically defensible theory must be.

Scenarios that defeat realism leave us with no explanation of *T*'s ability to predict *R* successfully. We must allow that some novel success may simply be chance. And of some novel success, the correct explanation may not be epistemically warrantable by us. But these possibilities do not prevent realism from being warranted where the explanation it offers is undefeated. Rather, they register the unavoidable defeasibility of realism and suggest that realism be embraced only where a substantial record of sustained novel success has been achieved. As a philosophical interpretation of the epistemic status of *T*, realism requires retrospective evaluation of the evidential situation; we should not expect to read realism off of scientific practice in real time.

5.7 The Importance of Novelty

I have rebuked anti-realism for insensitivity to differences in the status of theoretical entities. The positive argument that I have constructed for realism reveals a related failing. Anti-realism lacks the resources to distinguish novel predictive and explanatory success from a theory's routine empirical applications. There is no question that novel applications are especially compelling epistemically. From the use of Newtonian theory to discover the outer planets of the Solar System, to Mendel's backcross test of the genetic hypothesis, to the application of atomic theory to Brownian motion, to the bright spot discovered in spherical diffraction, to the conversion of matter and energy, to the gravitational deflection of starlight, novel results have provided a warrant for theories that mere conformity to observation cannot signify.

To account for this difference requires treating not just the observable results themselves, but a theory's success in predicting them, as a proper object of explanation. Where, but only where, the results are novel, the explanation of their successful prediction must be that the theory has correctly identified and described the theoretical

mechanisms of their production. To treat all theoretical entities with agnostic indifference preempts this explanation. Anti-realism may acknowledge the possibility that a theory has posited the right entities. But in insisting that this possibility be epistemically inaccessible, anti-realism in effect reduces the theory to a mere predictive instrument: a theory's only permissible endorsement is that it is, unaccountably, a good predictor. And with respect to this attribute, all observable phenomena, from novel discoveries to programmed outcomes, are epistemically on a par. All of a theory's observed consequences are equal confirmation of its predictive accuracy.

The concept of novelty gives realism a criterion for epistemic commitment; it identifies the conditions under which theoretical belief is warranted. This criterion undercuts the holistic conception of confirmation, which is as unacceptably indiscriminate as anti-realism's dismissal of all theoretical entities. The portions of a theory that are to be interpreted realistically and expected to survive future theory-change are those responsible for the theory's novel success. The realist does not presuppose the correctness of current theory in allocating epistemic commitments to past theory. He identifies what was novel in the successes of past theories and determines how that novel success was achieved. His criterion of epistemic commitment is the same for past science as for present.

Nor does the realist presuppose the legitimacy of whatever inference is needed to close the logical gap between theory and evidence. He endorses a specific ampliative move to support a theory with novel success over its algorithmically generated rivals. I argued that *T*'s rivals are not confirmable, while, by hypothesis, *T* is. The basis for the hypothesis was simply that ampliative inference as such may not be disallowed on pain of skepticism. Allowing ampliation as such is enough to make *T* a proper object of confirmation, but not *T*'s rivals.

The positive argument that I have since constructed for realism affords greater specificity. The ampliative principle that warrants interpreting *T* realistically is an abductive inference from *T*'s novel success. No algorithmically generated rival to *T* can possibly claim novel success. The rivals cannot make novel predictions at all, because their observational consequences violate both requirements for novelty. All their consequences are already consequences of *T*, which violates uniqueness. And all their consequences are used essentially in their construction; their semantic content is determined by specifying what their consequences are to be. This violates independence. It follows from my positive argument for realism that the rivals are ineligible for epistemic support. Their availability cannot, therefore, establish the underdetermination of *T*. This reasoning presupposes nothing more than the rejection of a sweeping skepticism that would deprive ordinary, paradigmatically unproblematic beliefs of their necessary grounding in explanatory inference.

5.8 Observation and Theory

To this point, I have assumed a clear and absolute distinction between observational and theoretical propositions. Observational propositions formulate the evidence by which theories are judged, and the question has been whether this evidence is ever, in principle, sufficient for epistemic commitment to theories. The distinction is nec-

essary to engage the anti-realist, for without it anti-realism collapses into skepticism. Anti-realism requires an observational level exempt from problems of underdetermination and historical inconstancy to circumscribe its range of incredulity.

Realism, by contrast, makes no such requirement. Although traditionally formulated within the problematic of vouchsafing the inference from observation to theory, realism's essential message is that epistemic justification is not confined to the observable. Realism does not require that there even be an essential division of observation from theory; it requires only that if there is, it does not divide the justifiable from the unjustifiable. In particular, my own positive argument for realism does not depend upon the observational status of novel results. Results novel for a theory are consequences of it that satisfy independence and uniqueness. Any result, whether classifiable as observational or not, could be novel and, if epistemic justification is not restricted to an observational classification, could be independently established and evidentially probative.

As anti-realism requires an epistemically privileged category of observational propositions distinguished from theory and realism does not, it is appropriate to ask whether the requirement can be met. I am inclined to concede to the anti-realist at least a rough-and-ready distinction between observation and theory, although I think it is more contextual and variable than he can tolerate. The issue I wish to press is that of epistemic privilege: Could there be justified observational beliefs if no theoretical beliefs are justifiable?

Observation reports are certainly fallible, and often they are corrected or reinterpreted in ways that affect their evidential bearing on theory. In order to trust them as evidence, there must be reason to believe that conditions are not such as to undermine their reliability. But beliefs as to the conditions under which observation is reliable are not themselves classifiable as observational. We observe objects, not the accuracy of our observation of them. Should it be imagined that the accuracy of observation is itself somehow an object of observation, then it would become necessary to ask after the observability of the accuracy of observations of accuracy. That won't do. Unobservable entities and processes can interfere with our ability to represent objects accurately on the basis of observing them. We use theories about the nature of observation and the impediments to its veridicality in assessing the correctness and the evidential weight of observation reports. Only if these theories are justified are we entitled to our discrimination of trustworthy from untrustworthy reports.

Anti-realism assumes that epistemic justification has a rigidly foundationalist structure. There is a privileged class of judgments sanctioned by observation whose justification is unproblematic and automatic. This class must be identifiable independently of any theorizing, for it is assumed that its members are the common explananda of rival theories. Propositions outside this class are justifiable, if at all, only indirectly by inference from those inside. The more remote a proposition's content from paradigmatically observable situations, the more dubious is its epistemic status.

Given this foundationalist picture, doubts as to the existence of such a privileged class of propositions with the required properties are immediately skeptical doubts. If observation is fallible, if observational judgments may themselves be objects of justification, if their evaluation invokes judgments outside their class, then the entire structure of justification collapses. The anti-realist wants to be a skeptic only at the

level of theory. But his epistemology makes this restriction untenable. For his arguments against theory can be repeated at the level of evidential reports used to judge theory. Such reports are revised, reinterpreted, and underdetermined by experience. In view of these liabilities, the consistent foundationalist can only retreat to a yet more basic epistemic level.

If there is such a level, then with respect to it the anti-realist's arguments fail. There will be a foundation for belief not itself in need of epistemic support but capable of providing it. The evidential judgments of science will be justified in terms of it. But if this works for the evidential judgments of science, why not for the theoretical judgments inferred in turn? There will be no reason to arrest epistemic assent short of theory. If there is no epistemically basic level, then anti-realist arguments succeed everywhere. There is no place to arrest epistemic descent into a sweeping skepticism that anti-realism eschews.

We are left with two nonskeptical options: repudiate foundationalism or allow epistemic support to extend throughout the structure of our belief system. The first option eliminates the anti-realist's grounds for denying the possibility of theoretical knowledge. On this option, whether or not a belief is justifiable has no automatic connection to a classification of beliefs into ontological kinds. The second option provides for the justification of theoretical beliefs. Their epistemic status may be less secure than that of beliefs closer to the foundations, but the difference is at most one of degree. Either way, scientific realism wins.

Bibliography

Boyd, R. 1984: The current status of scientific realism. In Leplin (1984a), op. cit., pp. 41–83.

Churchland, P. M. and Hooker, C. A. (eds.) 1985: *Images of Science*. Chicago: The University of Chicago Press.

Fine A. 1984: The natural ontological attitude. In Leplin (1984a), op. cit., pp. 83–108.

Kitcher, P. 1993: *The Advancement of Science*. New York: Oxford University Press.

Laudan, L. 1981: A confutation of convergent realism. *Philosophy of Science*, 48, 19–49.

Laudan, L. and Leplin, J. 1991: Underdetermination and the empirical equivalence of theories. *Journal of Philosophy*, 88, 449–72.

Leplin, J. (ed.) 1984a: *Scientific Realism*. Berkeley, CA: University of California Press.

— 1984b: Truth and scientific progress. In Leplin (1984a), op. cit., pp. 193–218.

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

—Psillos, S. 1999: *Scientific Realism: How Science Tracks Truth*. New York: Routledge.

CONTEMPORARY DEBATES IN PHILOSOPHY OF SCIENCE

Edited by

Christopher Hitchcock

Blackwell
Publishing

350 Main Street, Malden, MA 02148-5020, USA
108 Cowley Road, Oxford OX4 1JF, UK
550 Swanston Street, Carlton, Victoria 3053, Australia

First published 2004 by Blackwell Publishing Ltd

Library of Congress Cataloging-in-Publication Data

Contemporary debates in philosophy of science / edited by Christopher Hitchcock.
p. cm. — (Contemporary debates in philosophy ; 2)
Includes bibliographical references and index.
ISBN 1-4051-0151-2 (alk. paper) — ISBN 1-4051-0152-0 (pbk. : alk. paper)
1. Science—Philosophy. I. Hitchcock, Christopher, 1964– II. Series.

Q175.C6917 2004
501—dc22

2003016800

A catalogue record for this title is available from the British Library.

Set in 10/12½ pt Rotis serif
by SNP Best-set Typesetter Ltd., Hong Kong
Printed and bound in the United Kingdom
by MPG Books Ltd, Bodmin, Cornwall

For further information on
Blackwell Publishing, visit our website:
http://www.blackwellpublishing.com