Testing Philosophical Claims about Science[1]

David Hull

Northwestern University

With respect to the role of evidence in testing statements both *within* science and *about* science, four combinations are possible. Logical empiricists such as Hempel (1966) insist that evidence plays a crucial role in the sort of testing that goes on in science. In their own discussions of science, logical empiricists also include occasional examples drawn from science, both current and past, but these examples function only as illustrations of the points that they are making about science, not as tests. A common view among philosophers of science, and not just logical empiricists, is that no connections exist between what scientists actually do and the sorts of claims that philosophers make about science. Even if no scientist ever explained anything by deriving it from a law of nature, the covering-law model of scientific explanation would remain untouched. Deduction is deduction, and nothing about the conduct of science can touch that. According to these philosophers, evidence may play a crucial role in testing the empirical claims that scientists make about the natural world but no role whatsoever in testing the sort of meta-level claims that philosophers of science make about science. Such philosophical claims must be supported in some other way.

According to one prevalent reading, Kuhn (1962) advocates the opposite position with respect to the role of evidence in science and the study of science respectively. To the extent that evidence plays any role whatsoever in scientists choosing between different paradigms, it is never decisive. However, Kuhn urges a greater role for the history of science in the choice between different philosophies of science. Incommensurability between scientific theories precludes a decisive role for evidence with respect to theory choice within science, but the even greater incommensurability that characterizes different philosophies of science somehow does not preclude a decisive role for the history of science as evidence in choosing between these meta-level theories.

Some of Kuhn's social constructivist disciples have carried his position to even more extreme lengths. As Collins (1981a, p. 218) sees it, advocates of the radical program in the sociology of knowledge, "must treat the natural world as though it in no way constrains what is believed to be." However, sociologists of science should "treat the social world as real, and as something about which we can have sound data" (Collins 1981a, p. 217). Other social constructivists have pursued what they take to be Kuhn's views to their logical conclusion—total relativism. Evidence plays no role

PSA 1992, Volume 2, pp. 468-475

https://doi.org/10.1017/S027086470000953X Published online by Cambridge University Press

in either science or meta-level investigations of science. Just as the natural world does not constrain our interpretations of the natural world, texts do not constrain our interpretations of these texts (Woolgar 1988).

The fourth alternative is that evidence can play a significant role in both science and the study of science. Evidence is not easily brought to bear on general claims in science. Showing the relevance of evidence in testing meta-level claims is even more difficult, especially since students of science make very different sorts of claims about science. For example, philosophical claims about the adequacy of operational definitions are different in kind from sociological claims about the disproportionate effect that the work of a very few scientists has on the course of science. However, those of us who see a need for testing meta-level claims made by students of science insist that even the most philosophical claims about science can be interpreted so that evidence can be brought to bear on them, albeit sometimes quite indirectly. Certainly the sorts of claims made by sociologists of science can be tested empirically.

1. Testing Meta-Level Claims

If the theory-ladenness of even the most observational of terms really does pose the insurmountable problems in choosing between alternative theories that Kuhnians claim, then we should see the results of this incommensurability in science, both past and present. If different paradigms are incommensurable, then some fairly obvious conclusions follow about how successfully scientists who hold the same and different paradigms can communicate with each other. Scientists do, not infrequently, talk past each other, but does this failure in communication covary with adherence to different paradigms?

One would think, given the huge literature on the subject of incommensurability, that numerous authors would have attempted to test this apparent implication of Kuhn's thesis in a systematic way. Lots of case studies have been presented, some showing how incommensurability did pose a serious problem in communication, some showing how it did not. If the correlation between incommensurability and communication is taken to be universal and the Popperian asymmetry between verification and falsification is taken seriously, then a single contrary instance should refute this meta-level claim. Several refuting instances have been presented.

For instance, Ruse (1979) has detailed the controversy over the age of the earth that took place in the second half of the 19th century between evolutionists and geologists, on the one hand, and physicists, on the other hand. Evolutionists and geologists thought that the earth was extremely old, while physicists insisted that it was relatively young. Neither side was able to come up with extremely precise figures, but the physicists thought that the earth has been around for at least twenty-five million years, possibly as long as a hundred million years, while the evolutionists and geologists insisted that it had to be much older—hundreds of millions of years.

If any two groups of scientists ever held incommensurable paradigms, the Darwinians and the Kelvinians did. They deployed different symbolic generalizations, employed different methodologies, shared different professional values, and most importantly extrapolated from very different exemplars. However, in spite of all these differences, these two groups of scientists were able to disagree with each other just fine. Perhaps they meant something slightly different by "age" or the "earth," but such slight differences in meaning were overridden by the magnitude of the differences in the age of the earth implied by these two paradigms. When Darwin estimated that the denudation of a single stratum took 250 million years and the physicists had a hard time coming up with that figure for the entire duration of the earth, a contradiction clearly existed, incommensurability notwithstanding.

Although case studies are in principle sufficient to *refute* a general thesis, in practice they rarely do so. Too many objections can be raised to their relevance, applicability, construction, execution, etc. They are not even in principle sufficient to *confirm* a general thesis. Rarely, however, are theses about science presented in a universal form. Usually they are hedged here and there. For example, sometimes Collins (1981a, p. 218) portrays the radical program in the sociology of knowledge as requiring that the "natural world in no way constrains what is believed to be." At other times, it requires only that the "natural world has a small or non-existent role in the construction of scientific knowledge" (Collins 1981b, p. 3). Systematic, preferably quantitative, studies are required to test claims such as these (Cole 1992).

No one to my knowledge has even attempted such a study with respect to the effects of incommensurability on success in communication. In my own research, I have studied these effects in a semi-systematic way (Hull 1988). I found no clear correlation. Confusion was as common within groups of scientists holding the same paradigm as between groups holding different paradigms. I realize that this lack of correspondence is impossible, but as far as I can tell, it is actual. Either there are so many other causes for failure to communicate successfully in addition to incommensurability that they swamp the effects of incommensurability, or else incommensurability does not present the insoluble problems that holistic semantic theories seem to imply they they should.

2. Idealizations

One obvious response to the above comments is that the connection between philosophical analyses and science is not as simple as I make it out to be. Philosophers discussing the problem of incommensurability are not talking about science as it is practiced but about idealizations of their own construction. Idealizations play legitimate roles in science. Perhaps they play equally legitimate roles in our analyses of science.

In order for two theories actually to contradict each other, they must be presented in complete, totally precise, possibly axiomatized form with all meanings sharpened to a fine point by sufficient conceptual analysis. Only then can the two theories be shown to be incommensurable. Scientists do not present their theories in such an ideal form. Nor do they evince any interest in doing so. At one time philosophers set themselves the task of producing such ideal versions of scientific theories. However, such undertakings are now decidedly out of favor. Instead, we are asked to consider problems that *would* arise *if* we had two perfectly formulated theories. For such ideal formulations, incommensurability would be a problem.

Scientists test ideal laws within science by seeing how real systems behave as they approach the ideal. Some inclined planes exhibit less friction than others. When actual surfaces are ordered according to their degree of friction, the results approach the ideal. Even if incommensurability characterizes only those theories that are completely and perfectly formulated, actual theories can be ordered to see if incommensurability becomes a greater problem as this ideal is approached. With respect to the theories that scientists actually produce, deciding which observation statements follow from these theories and which do not is far from easy. All sorts of approximations and simplifications have to be introduced, just the sorts of approximations and simplifications needed to derive commensurable observation statements from different theories. As a result, attempts to test one theory in isolation do not look all that different from attempts to test two theories by inferring incompatible and, hence, commensurable observation statements from each. As science proceeds and theories in a particular area become better formulated, the issue of incommensurability should become even more prominent. So far no one has attempted a study to see if incommensurability becomes increasingly more evident as scientists make their theories increasingly precise.

The issue is, as before, the testing of meta-level claims. Given a holistic semantic theory, incommensurability follows automatically. No evidence about the actual course of science is in the least relevant. On this view, semantic theories have nothing more to do with communication than the covering-law model of scientific explanation has to do with how scientists explain natural phenomena. Claims made by philosophers of science may sound as if they are about science and can be tested by reference to science, but in point of fact they express philosophical theses so abstruse that nothing so crude as evidence can be brought to bear on them. If this is the position that philosophers adopt, then detailed case studies are just so much deceptive window dressing. If all case studies are supposed to do is to illustrate a particular point, then brief gestures or silly science fiction examples will do. All the effort needed to set out real examples in all their complexity is wasted. Loading a philosophical discussion with detailed history of science may fool the unsuspecting reader into thinking that the author is talking about science, but that is all.

3. Studying Science

Those of us who are not inclined to take the a priori route are still left with plenty of problems. Bringing history of science to bear on general claims about science is extremely difficult. One of these difficulties is a meta-level version of a problem raised by philosophers in the context of science itself—theory-ladenness. Within science, observation terms are laden with the very theories that these observation statements are meant to test. If you approach the relevant data from the perspective of a particular theory, e.g., Darwinian evolution, which data seem relevant and how you construe these data will be strongly colored by your beliefs about the evolutionary process. Hence, you should not be surprised when your observations *support* your theory. However, you should be surprised when they *contradict* it, but contrary to a priori expectations, sometimes they do. No matter how strongly one's general views color one's estimations of data, sometimes these data can challenge the very theories in which they are generated. It should be impossible, but once again it is actual.

For example, T.H. Morgan's investigation of fruit flies eventually led him to abandon nearly every basic belief that he designed his experiments to support. He also reported no conversion experience as he abandoned one paradigm for another. Rather, he painfully modified one belief after another as the experiments that he and his students ran forced him to. Certainly career interests influenced Morgan the way that they influence all scientists, but if "interests" of a broader sort played a significant causal role, it is far from apparent. Morgan and his coworkers in the fly room were middle class, Euromales before they began their investigations; they remained so afterwards.

As Lakatos (1971) has pointed out, we are confronted by parallel problems in describing the course of science. We come to the study of history of science with all sorts of beliefs about history and science. These beliefs are only half-formulated and most not even explicit. To make matters worse, they usually do not deserve to be called a "theory" of history. Hence, their influence on the "data" that are generated are likely to be even more pervasive and elusive than the parallel situation in science.

https://doi.org/10.1017/S027086470000953X Published online by Cambridge University Press

As Richards (1993) points out in his paper, most practicing historians of science are crude inductivists when it comes to their own work, even those historians who reject inductivism as adequate for the practice of science itself.

The influence of the general beliefs held by historians on the stories that they tell are obvious. Recently, Columbus bashing has become fashionable. Did Columbus really discover America? After all, human beings already inhabited the continent when Columbus arrived. They got there long before via the Bering Strait land bridge. And Columbus may not have been even the first old-world person to set foot in the new world. Perhaps some Norseman or Viking may have made it there first. In addition, Columbus was not the only European on his boat. He was not the first to sight land. One of his men did. How come none of his crew members get any credit? They discovered America too. To make matters worse, Columbus did not think that he was discovering America. He thought he was arriving at the eastern shores of the Indies. Nor did he term his discovery "America." This name was coined much later, and on and on.

As trendy as the preceding discussion may sound, it introduces no problems not already familiar to historians of science. Did Mendel really discover Mendelian genetics? One can find numerous examples of three-to-one ratios in the works of his predecessors. Besides, Mendel did not think that he was discovering the laws of genetics, let alone Mendelian genetics. He thought he was investigating speciation by means of hybridization. William Bateson at the turn of the century was the one who was primarily responsible for transmuting Mendel's observations on peas into the science of genetics and making Mendel its patron saint.

As critical as I am of the general philosophical views of the social constructivists, they have forced us to see the bias that is introduced in the study of science by an overemphasis on "great men." For example, Desmond (1989) has shown how different the impact of Geoffroy St. Hilaire on Victorian science looks when viewed from the perspective of ordinary anatomists in medical schools rather than from the perspective of such big guns as Lyell, Owen and Darwin. The story of natural history in Victorian Britain reads very differently when it is written to include lesser lights along with the major figures, as different as the discovery of America looks from the perspective of Columbus's crew.

History of science cannot be written from no perspective whatsoever. It also cannot be written from all possible perspectives. All anyone can do is to be explicit about the perspective that one brings to a particular study. If our meta-level paradigms were so powerful that no observation couched in them could possibly refute them, then we would be in real trouble, but as in the case of science, students of science come up with observations about science that do not fit neatly into their own belief systems. For example, Popperians have attempted to test (or possibly only illustrate) Popper's views by recourse to the history of science. As biased as they may have been in favor of Popper's worldview, they were not always able to make the stories come out right—at least not without massive rational reconstruction. Bringing evidence to bear on ordinary empirical claims is not easy. All the problems that beset such efforts are only magnified in testing meta-level claims. Even so, these problems are not so hopeless that they cannot be overcome, if only we actually try.

4. Operationalizing in the Study of Science

In science theoretical claims have to be operationalized in order to be tested. Such operationalizations require reduction in scope as well as the introduction of rough approximations and particularizations. For example, biologists have long assumed that dinosaurs, like extant reptiles, were cold blooded. When the suggestion was made that they might have been warm blooded, biologists had to decide how such an hypothesis might be tested. Among extant predators, the ratio of predator to prey among warm-blooded predators is 1:50, while this same ratio among cold-blooded predators is 1:5. Since the difference is so great, just possibly the fossil record is good enough to choose between these two alternatives.

Of course, without even pausing to breathe, any red-blooded biologist can think of indefinitely many objections to the operationalizations required for this test (and some already have), but as sceptical as scientists are, they are much more tolerant of imprecision and possible error than philosophers are. The sort of argumentation that occurs in science does not come close to the extremely high standards that we set for ourselves. For example, using citations to gauge the impact of a paper on a particular area of science is a crude measure of importance, but it is the sort of operationalization common in science. I am sure that most philosophers would find the operationalizations that I used in assessing "success" in communication hopelessly inadequate, but it is this attitude that precludes philosophers from testing their beliefs about science by reference to science.

A common view expressed by philosophers about science is that the meanings of theoretical terms emerge only in the context of testing. Scientists cannot know in advance of their empirical investigations what they mean by their more general concepts. Meanings change as knowledge advances. The same should hold in our study of science if we are to provide theoretical definitions for our meta-level terms. We cannot possibly know what "testing," "experiment," and even "science" mean in advance of any and all empirical investigations. If we propose to test our knowledge of such general concepts as testing, we have to be willing to accept for the purposes of a particular study operationalizations that we all agree are crude, not good enough by half, etc. But this is the only way to improve upon these operationizations and, hence, our understanding of these general concepts. Conceptual analyses by intelligent ignoramuses can get us only so far.

Explications of theoretical terms in science by philosophers require our entering into the scientific process. If we are to assess the adequacy of a particular analysis of "species" or "gene," we have to persuade the relevant biologists to incorporate these conceptions into their own work and see what happens. That way they become theoretical definitions rather than conceptual analyses (Milliken 1984, Neander 1991). If conceiving of species as spatiotemporally restricted and located historical entities helps biologists improve evolutionary theory, then this conception must have something going for it. Such a procedure has the added virtue of allowing philosophers to support their theses by pointing to scientific usage.

What we need in science studies is theoretical definitions, not conceptual analyses, and theoretical definitions require theories. Although grand theories about the nature of science are currently out of fashion, I think that we need to rehabilitate them. We need to construct theories about science the way that scientists construct theories about fluids, gene flow and continental drift. To construct such theories, we need data, and our only source of data is the study of science, past and present. These histories will be theory laden. So? If scientists can use the data generated in the context of theories with which they disagree, why can't we? I do not agree with all the general views about science held by Rudwick (1985) and Desmond (1989), but I have no trouble using their histories to test my own general views about science. The stories that historians tell are theory-laden but not so theory-laden as to be useless. Needless to say, I have my own grand theory of science (Hull 1988).

5. Normative Claims about Science

Not all scientific claims are simply descriptive. Some are also nomic. As difficult as it is to set out criteria that mark this distinction, I am still old fashioned enough to think that there are laws of nature and that a continuing goal of science is to discover these laws. Similarly, I hope that not all of the claims made by those of us who study science are going to be simply descriptive. Some I hope will turn out to be analogous to laws of nature—laws about the scientific production of knowledge. If the distinction between descriptive and nomic claims within science is so difficult to set out, the parallel distinction at the meta-level will surely be even more difficult. The notion of meta-nomic necessity may sound overly ambitious to many people, but it is one of the prerequisites for a successful empirical theory of knowledge acquisition in science.

Philosophers have also traditionally expressed normative claims about science. In general, such prescriptions are extremely difficult to test. One suggestion is to convince groups of scientists. Have them adopt one's views about how science *should* be conducted and see what happens. If science in such areas immediately grinds to a halt, then possibly something is wrong with one's normative claims. Conversely, if those scientists who adopt your views are even more successful in attaining their epistemic goals, then possibly there is something to be said for these norms. For example, scientists do not spend much time precisely replicating the work of other scientists. They adopt the results that support their own views without testing. They tend to reserve testing for those results that threaten their own findings, and these tests are rarely exact replications—whatever that might mean. For some, this lack of precise replication may seem a fault, as if scientists are somehow falling short of proper scientific conduct. Scientists *should* replicate all results before using them. However, I strongly suspect that if enough scientists adopted such a prescription, scientific progress would be sharply curtailed.

In emphasizing how important it is for scientists to understand and incorporate the views set out by students of science into their own work, I am not committing myself to the position that scientists are the final arbiters with respect to matters about science as well as within science. To the contrary, I think that philosophers are right about the fundamental inadequacy of operationalism as a philosophical thesis, regardless of what certain behavioral scientists may think on this score. They are mistaken and could easily discover their mistake by reading a few well-chosen papers on the subject. These papers need not be written by philosophers, but the issues will nevertheless be philosophical. However, I do think that the proof of the pudding is in the eating, and only if scientists come to incorporate explicitly formulated meta-level beliefs about science in their own work can we ever hope to see what effects that they have on science.

Notes

[1]Thanks are owed to Kim Sterelny and Todd Grantham for reading and commenting on an early draft of this paper.

References

Cole, S. (1992), *Making Science: Between Nature and Society*, Cambridge: Harvard University Press.

Collins, H.M. (1981a), "What is TRASP?: The Radical Programme as a Methodological Imperative", *Philosophy of the Social Sciences* 11: 215-224.

_ _ _ _ _ _ _ . (1981b), "Stages in the Empirical Program of Relativism", *Social Studies of Science* 11: 3-10.

Desmond, A. (1989), *The Politics of Evolution: Morphology, Medicine, and Reform in Radical London*. Chicago: University of Chicago Press.

Hempel, C.G. (1966), *Philosophy of Natural Science*. Englewood Cliffs: Prentice-Hall.

Hull, D.L. (1988), *Science as a Process: An Evolutionary Account of the Social and Conceptual Development of Science*. Chicago: University of Chicago Press.

Kuhn, T.S. (1962), *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

Lakatos, I. (1971), "History of Science and Its Rational Reconstruction", in *PSA 1970*, R.C. Buck and R.S. Cohen (eds.). Dordrecht: Reidel.

Milliken, R.G. (1984), *Language, Thought, and Other Biological Categories: New Foundations for Realism*. Cambridge: MIT Press.

Neander, K. (1991), "Functions as Selected Effects: The Conceptual Analysist's Defense", *Philosophy of Science* 58: 168-184

Richards, R.J. (1993), "History as the Necessary Foundation for Philosophy of Science", in *PSA 1992*, vol. 2, D. L. Hull, M. Forbes and K. Okruhlik (eds.). East Lansing, MI: Philosophy of Science Association.

Rudwick, M.J.S. (1985), *The Great Devonian Controversy: The Making of Scientific Knowledge Among Gentlemanly Specialists*. Chicago: University of Chicago Press.

Ruse, M. (1979), *The Darwinian Revolution: Science Red in Tooth and Claw*. Chicago: The University of Chicago Press.

Woolgar, S. (1988), *Science: The Very Idea*. London: Tavistock.