The Relevance of History to Philosophy of Science

Robert G. HUDSON

Received: 2005.10.11
Accepted: 2005.12.09

BIBLID [0495-4548 (2006) 21: 56; pp. 197-212]

ABSTRACT: My task in this paper is to defend the legitimacy of historicist philosophy of science, defined as the philosophic study of science that takes seriously case studies drawn from the practice of science. Historicist philosophy of science suffers from what I call the 'evidence problem'. The worry is that case studies cannot qualify as rigorous evidence for the adjudication of philosophic theories. I explore the reasons why one might deny to historical cases a probative value, then reply to these reasons on behalf of historicism. The main proponents of the view I am criticizing are Pitt (2001) and Rasmussen (2001).

Key words: historicized philosophy of science, induction, reflexivity, flux, methodology, evidence, Joseph Pitt, Nicolas Rasmussen.

1. Introduction

Since the work of Thomas Kuhn, philosophers of science have been more assiduous in ensuring that their work is well-grounded in historical practice. In contrast to the abstractness of positivist analyses of science or the hagiographic accounts of science complained about by Harry Collins and Trevor Pinch in *The Golem* (i.e., those accounts that appeal to 'Mount Newton' and 'Mount Einstein'), it has been realized by philosophers that historical case studies have to be taken more seriously if fidelity to *real* science is to be maintained.

But there is a important problem to contend with if we admit historical cases to play a probative role in the philosophic examination of science. It is in fact a problem that attends to philosophic work generally. For in adjudicating philosophic disputes, whatever their topics, we need to have a resource of informative data. For example, in ethics we need access to exemplary ethical decisions, in the philosophy of mind we need unquestioned examples of things with minds, in aesthetics we need examples of art, especially of good art, in epistemology we need authoritative instances of knowledge, justification and good reasoning, and so on. Thus, philosophic debate often proceeds by the formulation of a theoretical hypothesis, where one might initially argue for this hypothesis on the basis of a more general theoretical approach, followed by the testing of this hypothesis against various case studies. For example, suppose someone formulates an ethical theory according to which performing such and so act is deemed right. And let us say it is granted by most people that such an act cannot be right. We would then conclude that there is a problem with this ethical theory. Similarly, it might follow on some theory of mind that a particular sort of thing lacks a mind. If we are inclined to concede such a lack, we are then motivated to approve of the theory of mind that entails it. Examples of this sort, where cases adjudicate philosophic theories, easily multiply. But perhaps now we see the hazard in this form of argument. For from where does one derive one's set of adjudicatory case examples? That is, what justification do we have for using this set of case examples rather than some other set? Is the justification here theoretical? If so, the relevant justificatory theory better not be the one under scrutiny! But from where else are we to derive the relevant guidance? If we are going to use case studies to test the adequacy of a philosophic theory and if these case studies are to be authoritative, then they need to be philosophically well-grounded —but that grounding can only come from the theory under evaluation if this theory purports to be philosophically fundamental.

At any rate, this is how we might express the problem as a paradoxer. It is an example of the ancient problem of analysis: to wit, how can one analyze the nature of something without already knowing, i.e., having the correct analysis of, this very thing? It is also an example of the worry raised by the naturalistic fallacy: given that we want to analyze the concept *X* (say, the concept of a good act), our goal is to pronounce on what one should mean by *X*, for this will be its correct analysis. But what guidance do we have for this correct analysis? One obvious source is a set of examples of *X*. Suppose then we pick from this set; we find an example of what *X* is. But then we are not so much interested in what *X is*, but what *X must be*, since we are analyzing *X* not just describing it. So just having examples of *X* is not good enough for analysis, and to infer from these examples to the nature of *X* is a fallacy. I take it that instances of this sort of fallacy are especially salient in value theory —it is commonly noted that having instances of, say, good acts, does not unproblematically allow one to draw any conclusions about what a good act is, since we are not sure if these instances of good acts really are good acts to begin with, absent a completion of our investigation into the nature of a good act, absent, that is, the analysis of 'a good act'. Thus, in arguing for the immorality of a kind of act, one cannot simply point out instances of these acts and ostend their immorality, for one can rightfully complain that the immorality of these acts is the very question at issue.

In this paper, my concern is not this philosophic problem in its full generality, but solely with the form it takes in philosophy of science which is informed by historical case studies. In this regard, the 'evidence problem' (as I shall call it) has been given recent expression in the work of Joseph Pitt (2001). Let me paraphrase Pitt's concern as follows. In examining scientific practice we either go into this examination with certain preconceptions about what science is about or we do not. In the former case, our historical research into science filters through historical cases and weeds out those cases which do not conform to our preconceptions. Yet, by filtering in this way, we are challenged with the problem of how our preconceptions can be tested by this historical research, if these preconceptions filter the case examples to begin with. So let us consider the alternative —let us avoid the use of substantive preconceptions in guiding our choice of historical cases. But now it seems we are served no better, for without any preconceptions about historical cases how do we know which historical cases to take seriously and which ones to dismiss? For example, what right do we have to be assured of the quality of Galileo's reasoning whilst righteously debunking Lysenko's logic? What reasons could we give to justify the guidance we derive from the work of reductionist physiologists at the same time as we rebuff the reasoning of animists? Without a filter of preconceptions, we are rudderless in our choice of historical data and it is difficult to see what constructive work will be performed in the philosophy of science from a historicist point of view. So, in dilemmical fashion, it seems we are logically in trouble whether we advert to historical cases, or not.

As it happens, I believe there are clear parallels here to the situation we are faced with in empirical research, generally speaking. It is indeed simply the problem of the theory-ladeness of observation. For example, with preconceptions about what granite is, we go out into the world and find samples of granite. But do we then make objective determinations of what is really in the world? That depends on what the stuff we are calling 'granite' really is, whether it really does satisfy our preconceptions of it. But do we have no access to *that*? Not directly —we only have direct access to what we *think* granite is, to our preconceptions of it, and how could such a thing test its own validity? Analogously we are faced with the same problem in historicist philosophy of science, except that the data are social scientific: we are observing the activities of a special group of people —scientists—, and where our focus is epistemological we are examining the status of these people as justifiers of beliefs, as knowers. As such, we observe how they go about defending empirical claims and theoretical hypotheses, and we use these observations to adjudicate our philosophical understanding of science. At this stage, it would seem that the theory-ladeness problem we face in such historicist philosophic inquires is potentially quite grave, compared to what we are faced with in natural scientific observations, in that we are utilizing observed instances of the 'good reasoning' exhibited by scientists to assess theories of good reasoning, in a case where our modus operandi is, we hope, *good reasoning itself*. It is this special form of circularity that is of special concern to us here. Generally, one wonders how epistemologists can be said to non-circularly justify their conclusions about justification —this seems to be a special challenge for them. A good bulk of this paper is directed at addressing this sort of problem, specifically as it arises in historicized philosophy of science.

From one perspective, at any rate, perhaps the use of such social scientific evidence in assessing epistemological theories is not so much of a worry compared to the sorts of theory-laden inquires made in the natural sciences. For consider again our geological example. In categorizing rocks into certain kinds, and then empirically studying these kinds, the 'theory-ladenness problem' points to the difficulty we have in ensuring that our categorization really does reflect the true nature of the world. For, once again, we cannot see the world 'as it really is', we cannot 'get out of our skins' to see the noumena behind the phenomena. By contrast, in the historicist use of social scientific evidence to decide between philosophical theories of science, perhaps our problem is alleviated somewhat since, in the process of researching scientific reasoning we can take into consideration the fact that we are reasoners as well. We can, in a sense, see the noumena behind the phenomena in investigating scientific reasoners for in our theory-laden investigations into how people reason, using as case examples instances of good reasoning exhibited by scientists, we have another source of information to guide us, and that is our own awareness of how each of us reasons, and reasons (we think) well. Now it is important to note that this alternative source of information is quite fallible. No one is suggesting that we have solid *a priori*, rationalist access to the nature of good reasoning. The point is simply that we have here an alternate source of information that we lack in the natural sciences. It is as though in our research in geology we could routinely experience the world as the rocks we are studying. No doubt, one's experience of the world, *qua* rock, would be highly confused. But I think it would be foolish to ignore such a source of information; *prima facie* it sounds quite valuable, assuming such a circumstance to be conceivable. Similarly, in our empirical research into good reasoning in science, it is valuable (but not decisive, of course) to factor in our reflective assessments of the case studies which are presented to us as evidence, that is, to factor in our own experiences of such reasoning, whether we view such reasoning positively or not. This is a source of information that it would be foolish to ignore.

Ultimately, it is because we have such a two-pronged attack into the nature of scientific work —examining such work from the outside in the form of case studies, and examining it as well from the inside through considered reflection— that I am generally optimistic about historicist philosophy of science. And what I wish to do in the remainder of this paper is to explore how this two-pronged strategy can be used in defusing the Pittean dilemma described above, thus perhaps setting the stage for a general resolution to the meta-philosophic problem of how philosophers can non-circularly use case studies in resolving philosophic controversies.

2. The Process of Using Case Studies in Historicized Philosophy of Science

What procedure should we use in accommodating the results of case studies in advancing claims on the philosophy of science? Here is my provisional recommendation. Our first step is to draw an initial partition of scientific cases into 'good science' and 'bad science'. For many philosophers this step is an easy one. For example, we find good science and good reasoning in Newton's argument for Universal Gravitation in the *Principia*, and bad science and bad reasoning in Lysenko's anti-Mendelian genetics; we find Darwin espousing good scientific logic in *The Origin of Species* whereas Henry Morris flounders logically in *Scientific Creationism*. There is no problem in performing such provisional judgements; and once we have settled on them it is appropriate to then let these judgements guide our philosophic inquires into science. But not inexorably. Whatever provisional judgements we start with, there are circumstances under which they should be reconsidered. For example, what if a form of reasoning a 'good scientist' commends never occurs in the work of other 'good' scientists or occurs in the work of 'bad' scientists, or what if a form of reasoning a 'bad' scientist commends occurs in the work of 'good' scientists. Under these conditions, I think we should be circumspect in drawing epistemological lessons from such contestable good (or bad) scientific work. For I think it would be highly surprising if a purported good scientist were to adopt a successful investigative strategy that remained hidden from view and beyond the power of invention by other scientists. Surely scientists start out equally equipped intellectually, though of course not equally equipped technologically, and if an investigative stratagem had been found to work in one case, its merits would in due course be made available to all, perhaps in independent discoveries. It is of course conceivable that a procedure is so specialized that other scientists in other fields see no use to it or that a procedure is so new that others have not had a chance to become aware of it. The philosopher needs to recognize these possibilities. But barring such circumstances we should be suspicious if, say, only Newton used a particular kind of investigative method. No doubt he is an exceptional thinker. But did not subsequent generations of researchers find any use for his method? If not, of what use could this method be? Conversely, suppose in examining Henry Morris's work we find him to be reasoning in a way common to the work of presumed 'good' scientists; he adopts a reasoning strategy many others have found successful. Then, for all our wish to debunk Morris, it will not do to dismiss his logic when it does not serve our ideological vent, and to approve the same logic when cast for more approved goals. Surely considerations of fairness require that we not use Morris' reasoning as exemplary of bad reasoning, and use the badness of such reasoning in an historicist fashion to justify a philosophic pronouncement, when it is the sort of reasoning used by many 'good' scientists in different historical contexts.

There are further circumstances under which we should re-consider our provisional categorizations of good and bad scientific work —specifically when a 'good' scientist appears to be reasoning in an 'intuitively bad' way, and when a 'bad' scientist seems to reason in a 'sensible' way. These considerations are less straightforward and more contentious, for quite frankly, it sounds aprioristic to qualify the probative force of historical cases by reference to our 'intuitions' or our 'sensibility'. Nevertheless, it is a good idea to include such factors in investigating the reasoning of scientists since, as we noted above, we do not examine this reasoning from afar as though we were examining a piece of insensible rock. We are after all reasoners as well, so it cannot be a bad idea to include our intuited, sensed evaluations into the picture. Moreover, a key mitigating point to consider here is that whatever pronouncements we arrive it in this process, there is always the contrary force of what we find in our historical case studies, in what scientists actually do and approve of in terms of methodology. For instance, suppose our a priori reflections suggest to us a rule of reasoning governing empirical research, and upon an application of this rule we are prompted to judge presumed 'good' scientific work as flawed. For example, suppose we judge, in our a priori wisdom, that the standard of error in science should be maximally low and that we should strive for infallibility. Intuitively, such a requirement sounds eminently prudent. Turning then to scientific work, we naturally find that *no* such work meets our standard. And we find that scientists, when presented with this standard, are universally dismissive of it. Such a standard is for them infeasible and unrealistic. What should our conclusion be? My claim is that, as philosophers, we need to rethink our a priori notions and perhaps discount them. Indeed, this is exactly what attends to the extreme

Cartesian scenario we just painted, for no one doubts the need to question such a standard. But I think there are less obvious and quite challenging cases of this sort of historical adjudication. That is, there are cases where, though rationalist intuition presses the reasonableness of an epistemic standard, a standard not seemingly too radical, scientists do not seem to follow suit. An example I have in mind concerns the question whether in retrieving empirical data it is worthwhile varying the experimental methods by which this data is produced. That is, is the epistemic quality of empirical data improved for its having been generated by two or more experimental regimens involving differing physical processes? Many philosophers have defended this approach —but I do not see scientists adopting it, in the main. It may be that scientists have something to learn from philosophers. But as a historicist philosopher my approach is to prompt us to question our a priori intuitions in such cases, if the historical facts do not bear then out.

Again, I have suggested that there is value in filtering our historical cases through our rationalist intuitions —these intuitions provide an important vantage point in examining case studies. But such intuitions are controvertible when they fail to capture the work and attitudes of practicing scientists. Moreover, I would hazard to say that such intuitions are not a particularly fertile source of information about scientific, and especially experimental methodology. The problem is simply that rationalist philosophers have not been working at science and experimentation as long as scientists and experimenters have. In my own historical research, I have found scientists to be in possession of many remarkable strategies for wresting facts from nature. As a case in point, consider that Ptolemaic astronomy was invented almost two millennia ago and is still used today in guiding ships navigating the world's oceans. Clearly there is substantial knowledge here about celestial patterns, the result of eons of careful observations and repeated testing of the principles of celestial motion. How was this accomplished? What normative principles did pre-Copernican astronomers use to establish their findings? I think it would be a momentous job to determine, from a priori scratch, what the relevant principles of reasoning were, what observations were made, what mathematical rules were deployed, to arrive at this monumental feat of cosmology. And I think similar comments apply to many fields of science; the past work of scientists has been characteristically remarkable and ingenious. This leads me to contend that scientifically-inclined epistemologists would do well to study these patterns of reasoning, not reinvent them. In other words, no matter how prescient our rationalist intuitions might be and despite the fact that they play a critical role in an effective historical philosophy of science, there are far too many important insights to be gained though the historical study of scientific work to cruise solely along a rationalist path.

Hopefully it is becoming clear how I propose to resolve the Pittean dilemma described above. In brief, our choice of historical data in the philosophic investigation of science is not an untutored choice, for we filter our choice through our preconceptions of the nature of good science. But these preconceptions are not the sole driving forces to our work. We need also consider the views of scientists themselves on the quality of the science being explored, and to factor in what it is feasible to hold as a standard for scientific work, given the finite limitations of scientists as *humans*, whilst reflecting in the background on how ingenuous and intellectually remarkable scientific work often is. All in all, I am proposing a balanced approach, one that garners information both rationalistically and historically, drawings conclusions that are at once intuitively sound and substantiated in case studies.

Having presented in rough outline how I plan to defend historicized philosophy of science, my task from here will be to further this defence by responding to some objections to my view. Three particular objections come to the fore: I call these 1) the problem of 'few data', 2) the problem of 'reflexive consistency', and 3) the problem of 'flux'. By responding to these objections I hope to enhance and defend the approach to historicized philosophy of science I have just set forth.

3. The Problem of Few Data

In studying historical cases and drawing from them philosophical lessons, our project is inductivist and should be evaluated using the principles of good induction. Now, so far in our discussion above we have focused on one particular problem attending induction, the problem of 'sample bias'. This is the problem we face in selectively choosing those empirical, historical facts upon which to induce —there is the prospect here of choosing those facts that preferentially suit our theoretical goals. Part of our response to this problem is to suggest that philosophers in studying the nature of scientific reasoning have, in a sense, an 'inside track' since they, too, are reasoners and so have a way of adjudicating in a nontrivial way judgments about what constitutes good, scientific reasoning. Moreover, let us add that when philosophers come to decisions about the quality of scientific research, their decisions are not made in a vacuum; particularly, such decisions impact what sorts of conclusions can be acceptably drawn about other scientific episodes. If, for example, an experimental strategy is applauded in one particular scientific case, then all else being equal the same experimental strategy should be applauded in the next scientific case. Requiring such consistency is an important constraint, one that has the capacity to quickly expose the flawed nature of some historically-based philosophies of science which occasionally err on the side of irrationally upholding the merits of one science and diminishing the merits of another, despite the fact that these sciences adopt kindred methodologies.

But sample bias is not the only problem confronting induction; there is also the issue of 'sample size'. How can it be feasible to generalize about all of science on the basis of inspecting one or two cases studies? Indeed, how is such a generalization legitimate when the case studies themselves are presented with necessarily limited information. For example, one might cite the work of Galileo as exemplifying good scientific practice —but there is a lot that is covered under the heading, 'the work of Galileo'. He wrote many things, on many topics, endured many influences, had various foibles, and so on; so to pick from his work an example of 'good scientific reasoning', and then to generalize to all Galilean science, or even more dramatically to all science whatsoever, is presumptuous, to say the least. Consider, in addition, that there are many different kinds of sciences to consider, from the natural to the social, the theoretical to the experimental, the ancient to the modern, the Eastern to the Western, the established and the frontier, and on from here. With all this data to consider, and factoring in the inevitable complexity and ambiguity of this data, how can an inductive move on the basis of one or two diluted case studies be anything more than speculation?

I believe these sorts of comments are overly pessimistic. To illustrate, I want to consider an analogous argument considering the recent SARS epidemic. As the epidemic raged, researchers attempted to identify the virus underlying this outbreak. But, as we remember, many people had the disease and they came from many different places. Some are female and some are male. Some are rich and some are poor. Some people died from the disease; some survived. Some passed the disease to others and some did not. Though our current level of understanding concerning how viruses work is formidable compared to what we knew in the past, there is still much more we need to learn about viruses. In all likelihood, researchers of the future will amuse themselves with the rudimentary nature of our understanding. Nonetheless, should these considerations lead us to be skeptical about our work with SARS? Is our investigation futile? Clearly not. Our current research is only a step to the future and there is no reason to resist formulating hypotheses about the nature of this disease and to pronounce sincerely about the etiology of SARS. To fall back on pessimism regarding our ability to understand this disease, a pessimism nurtured by the enormous scope of the data at hand and a realization of the dire complexity of viral pathogenesis, would make little sense to the medical community and would be unacceptable to the general public. So why not adopt the same attitude with regard to our inductive work in the process of understanding scientific practice? Is the highly complicated nature of a subject matter and the limited scope of evidence we have regarding it a reason to deny the value of induction in coming to understand it?

At this stage, the sceptical, anti-historicist critic can respond by drawing disanalogies between empirical research in epidemiology and a similar kind of research in the history and philosophy of science. For, with SARS cases, we know that certain factors can be ignored. Given the characteristic symptoms of the disease, researchers have a fair guess that it is the same disease in Canada as it is in China and Hong Kong, and is caused by the same virus. That is, the legal jurisdiction in which one suffers the disease is not a factor in the identification of the virus, nor is one's hair colour, one's religion, one's economic status, one's reading preferences, and so on. In other words, scientists have a lot of opportunity to piece through the cases, removing irrelevant factors and connecting relevant ones, much simplifying their investigation. Finally, they are (or will soon be) able to isolate a unique biochemical signature for the virus that can be used to track it with fair consistency and without much controversy. And from there the SARS virus will end up being listed in the published literature as one more catalogued, microbiological phenomenon, just as well understood as any other virus.

None of this, surely, can happen with the historical cases studies in science, the skeptic continues. This is because scientific episodes are human episodes, and understanding them requires, one anticipates, a fairly rich and deep understanding of human society and human nature, a feat reaching far beyond an understanding of such relatively simple things as pathogenic viruses. So perhaps some modesty is due the historicist philosopher of science if she contends that inductive leaps to the nature of 'science' are equal to the task. Empirically researching and comprehending the enormously complex phenomenon of scientific thought is a higher order task than understanding the etiology of a bad flu.

But once more the skeptic is excessively pessimistic. To preface my response, I am reminded of some comments made to me by an economist. I had queried him on the softness of economic data, on its pliability to differing interpretations. How does one examine a statistical trend on economic matters and draw any reasonable conclusions about these matters? One can read anything into the data, I was concerned. How can we be objective? His response was interesting and revealing, particularly because I do not think it was motivated by any conscious philosophical ideology. He remarked, with a dove-tailed swooping of his hand, that one must 'see through the data' to the objective reality behind. I was left with the image that he had some sort of x-ray machine that could peer through the surface of things to the reality underneath. But, as a description of a sort of scientific methodology, he was exactly right. Statisticians know this well: one is left with a set of data points, upon which one wants to draw a graph. The neophyte 'connects the dots'; the expert 'curves'. Comparatively, the doctors who first recognized the symptoms of SARS were confronted with a mess of cases, and then in a moment of inspiration saw a pattern —the disease's symptoms, its characteristic infectiousness, its surprising mortality— and then 'saw through the data' to the reality underneath, that there was a new condition here with its own disease pathogenesis. The point here is that good scientists do not always need 'lots of data' to perform effective inductions, and in fact the more one gets hunkered down and preoccupied with more and more data, the more we have an indicator that one's understanding of the situation is quite thin.

Now, without a doubt, I do not want to emphasize the reliability of this interpretive vision I am claiming scientists to have —it is fallible along with all the rest of our intellectual and observational regimen. The point is simply that the imperative —to collect more and more data— is fruitless if a researcher has no idea what to look for and has no idea where to start, and that this is just as true in the sciences as it is in historicist philosophy of science. Thus, to criticize historicist philosophers for 'jumping to conclusions' on the basis of a small data sample is, unless we are told more, unfair. Every researcher must take a chance on an interpretation, to start. The critical phase is what she does with that interpretation anon. That is, is her interpretation open to further testing, and is she willing to entertain further, different hypotheses? Does she recognize her initial interpretation as fallible? As long as her critical mind-set is in play, there is no problem with 'small data inductions', and rigorous natural science is fully aware of this.

4. The Problem of Reflexive Consistency

So let us grant, in performing historicized philosophical investigations, that we are not necessarily at a loss in only having recourse to a small data set. There is nevertheless, as I indicated earlier, something of a puzzle in the process of using historical case studies in science to advance claims about how best to reason, given that we are *reasoning* to these claims. The puzzle has two forms. First, suppose we start with style of reasoning *X*. Using *X* we investigate historical cases in science. The puzzles are, 1) in using *X*, could we learn that *X* is a flawed research strategy? And 2), could we learn that not-*X* is a strategy preferred by scientists?

With respect to 1), let us use as an example of *X* the process of induction, simply understood as the process in which we examine a number of cases, observe what is common to these cases and subsequently infer that all similar such cases have as a general property this common feature. Now, with *X* at hand, we attend to various case studies and observe whether they involve applications of *X*. But which cases should we use here? Those, of course, which exemplify what we take to be 'good science'. But presumably cases of 'good science', given our stated predilection to induction, will be cases where scientists actively engage in induction. When we see scientists reason in ways contrary to induction, we should, by our own lights, be sceptical about the value of these approaches. So our array of case studies will expectedly be populated by scientists using induction, or at least not populated by scientists flouting induction. And now the hazard should be clear, for how could we ever learn on the basis of our case studies that induction is *not* a strategy deployed by scientists? Ostensibly, we would be working from a population biased in favour of induction.

This is the problem of reflexive consistency. There seems to be an onus on a rational individual to approve of those standards of reasoning in others that conform to her own standard of reasoning. If one were to applaud reasoning in others that confuted one's own reasoning, especially where one's reasoning led one to conclude that others were reasoning in this way, then this would leave one in a 'reflexively inconsistent' position.

Problem 2) is the corollary to problem 1). Suppose one were to deny that one's advocacy of induction precluded one's ability to recognize 'good' scientific episodes as involving contrary-to-induction procedures. Observations of this nature are not so afflicted by one's preconceptions, one might suggest. And suppose now we examine various case studies to find that contrary-to-inductive measures are routinely authorized. Having sanctioned the use of induction for ourselves as historians, we now arrive at the conclusion, on the basis of induction, that scientists are contra-inductivists; and then, being historicists, we draw the conclusion that, to a first approximation, it is prudent to be contra-inductivist since scientists, after all, are contra-inductivists. But if, prudently, we then decide to be contra-inductivists, our main argument for concluding that scientists are contra-inductivists, itself based on induction, falls to the floor. So, overall, it does not seem that we can successfully argue on the basis of case studies that our provisionally preferred style of reasoning is a flawed one —supportive cases are too easy to come by, and the probative force of non-supportive cases is reflexively undermined.

So do these considerations negatively impact the ability to engage in historicist philosophy of science? If they do, I think it is only in a very narrow range of cases, those in which the method being used by the historicist philosopher is identical to the method being used by the subjects of the investigation, the scientists in the case study of concern. That is, in most circumstances reflexive consistency is not an issue because the scientists being investigated use methods completely unrelated to the methods used by philosophers —scientists pipette solutions, solve equations, adjust telescopes, and so on, all things philosophers professionally never do. So let us focus on the more narrow range of methods that potentially are common to both philosophers and their scientist subjects, the practice of induction (as simplistically described above) being one of these. Could philosophic proponents of induction, presented with an array of case studies involving induction, ever recognize induction as flawed? I think there is good reason to affirm this —philosophers can examine scientific inductive practice and, for independent reasons, recognize it as drawing the wrong conclusion. For instance, we are familiar with Russell's chicken who, upon being fed dutifully for many days, infers on the next day (alas, 'sacrifice day') that he will be fed to live another day. Historicist philosophers can easily recognize this happening in the sciences and it can inform their philosophic judgements, even if they are firm advocates of the style of reasoning being imputed. But in recognizing the failure of induction here, and perhaps in recognizing similar failures of induction in other cases, how could the historicist conclude *on the basis of induction* that induction fails? Easily, by seeing these failures of induction as exclusive to the sorts of cases being studied. For instance, induction where one is a farm chicken reflecting on mortality may well be erroneous, yet it is an altogether different matter to consider the value of induction with regard to the historicist study of science. Thus, one can coherently reject induction as an informative practice in some particular subject area while preserving the value of induction in a different area, such as in the historicist study of science.

Similar comments I believe apply to those cases where the inductivist philosopher is faced with cases where scientists avow contra-induction. If she practices induction she seems forced to confute induction if she is to generalize over what turns out to be contra-inductive practices amongst scientists. But once again maybe not if, when examining these studies more closely, one finds that in their advocacy of contra-induction scientists are arriving at the wrong answers. In fact, where contra-induction amounts to endorsing hasty conclusions without appeal to a broad range of cases, I think this is what we would find. Nevertheless, an inductivist philosopher could recognize on independent grounds that contra-induction succeeds in a particular set of circumstances, and could then argue inductively and cogently that contra-induction is epistemically preferable in these sorts of cases. This is possible so long as the cases under consideration are of a different sort than the ones reflected on philosophically. For example, contra-induction may work for cosmologists studying galaxy formation, and inductively the historicist may recognize this to be the case for galaxy formation research generally —nevertheless, the historicist's induction concerns specifically 'what scientists do', not 'what galaxies do', and so it is reflexively coherent for the historicist to performs inductions in this way.

So, in general then, it need not be the case that by advocating and deploying reasoning strategy X historicists will fail to recognize flaws in X. Even where reflexive consistency inclines one to reject challenges to one's form of reasoning, there are still ways to recognize flaws in this reasoning by examining historical case studies. This is because there are independent ways of determining that one's form of reasoning is flawed. Moreover, once a reasoning strategy has been isolated as flawed, one can, while advocating this reasoning strategy affirm this flaw so long as the subject matter of the case study differs from the subject matter of the historicist's investigation.

To conclude this section, let me introduce here a further sense in which historicist philosophy of science is threatened by the demand for reflexive consistency. Suppose, again, while in the process of examining historical cases we notice that scientists are deploying a reasoning strategy that we ourselves use in historical research. And again let the relevant reasoning strategy be the practice of induction as it is (loosely) described above. And suppose further that we observe that the advocacy by these scientists of induction is motivated irrationally; they do not seem to prefer induction for any other reason than they have been indoctrinated to do so, despite thinking that they have been critically evaluative in forming their methodology. Turning to ourselves, then, we seem compelled to draw the appropriate historicist conclusion, that perhaps we too propound our inductivist strategy on irrational grounds. After all, the best sorts of reasoners, scientists, are subject to these influences, so why not the rest of us? So acknowledging this, we are faced with a serious concern, for one might suspect that the accuracy of a method is dubious if such a method is adopted solely for non-epistemic reasons. For example, we might pass this judgment on the scientists we are investigating. And if we draw this conclusion, we cast doubt on our own reasoning since we too advocate induction as a reasoning strategy. But then it is induction that led us ultimately to the conclusion that scientists are improperly motivated. So if we suspect that induction is a flawed reasoning, this means that our assessment of our scientific subjects is inaccurate, casting doubt on our judgement that scientists are irrationally motivated. And if *this* judgement is flawed, our own reasoning is no longer a problematic target —which restores the quality of this reasoning and so restores our original judgement that scientists are irrationally motivated. And on we go in an endless self-defeating circle.

This further problem of reflexive consistency attends prominently to the sociological study of science, where conventional philosophers of science take heart in noting that if the sociologist's causal explanation of scientific behaviour is correct, then such a causal explanation can equivalently be made as regards the sociologist's own behaviour in arguing that scientific behaviour is caused, with the negative result that the sociologist's own reasoning falls back on itself with negative consequences. (See for example Bloor 1991, Chapter 1, for discussion of this issue.) But the defeat only comes if we infer that the causation of a style of reasoning implies the poverty of this reasoning —and this just is not true. Very good reasoning may have been inculcated in scientists, as well as in sociologists and philosophers. That is, the inference form 'this form of reasoning is caused (even uncritically)' to 'this form of reasoning is flawed' is invalid. Indeed, given that for most scientists their methodologies are learned by rote, even remarking that this inference has probabilistic value would be to misunderstand the value of scientific education. Thus I think this further problem of reflexive consistency should not pose a problem for historicists —the cogency of reasoning remains even after a causal examination of those who adopt this reasoning.

5. The Problem of Flux

There is one final concern I wish to deal with here, which I call the problem of 'flux'. Both Joe Pitt (2001) and Nicolas Rasmussen (2001) in speculating on the prospects for a historicized philosophy of science raise the possibility that scientific method is irremediably disunified, that is, it lacks a set of guidelines that are binding on all sciences at all times. In this respect Rasmussen describes scientific method as 'flux-like', whereas Pitt's description is 'Heraclitean'. Here Pitt and Rasmussen are suspicious of what one might call the 'One True Scientific Method' and to some degree we can share their suspicion as many candidates for the 'One True Method' have in the past been proposed only to be subsequently firmly refuted. For example, the Euclidean method of deriving all of science from some a priori true set of first principles was favoured by Descartes. However, at the crucial step of his deductive program Descartes needs to invoke the authority of God leaving many to doubt the relevance and value of his 'deduction'. Likewise, Carnap's method of reconstructing scientific language from remembrances of similarity and Russellian logic ambitiously sought to set in place a scientific method for grounding our knowledge of the world. Yet, as we all know, it was destined to fail and Carnap soon after gave up the project. Or again, some philosophers of science aspire to ground the cogency of scientific reasoning on Bayesian principles. Unfortunately for them there is little consensus that they will succeed at this (for example, see Mayo 1996 for a critique). So to this extent it seems that Pitt and Rasmussen's incredulity at finding the 'One True Scientific Method' is vindicated once we reflect on past philosophic attempts to identify this method.

Moreover, what do we learn when we turn to case studies drawn from science? Do we see a uniformity of method? This is, in fact, the main source of the doubt expressed by Pitt and Rasmussen regarding the existence of a univocal scientific method for, indeed, we find in science a multiplicity of techniques and strategies of reasoning. In the biological sciences cell specimens are chemically treated and examined under high-powered microscopes; in anthropology workers dig into the earth to uncover ancient artifacts; in astronomy researchers set up telescopes and, with cosmological models in hand, attempt to interpret the colours of distant galaxies. What do all these sorts of activities have in common apart from simple deductive logic and the use of observational data? Is there some special style of reasoning that underlies them all? Rasmussen (2001) finds nothing to tie them together and thus completely disinherits 'empirical philosophy of science' (as he calls it). Pitt, on the other hand, is less categorical —he recognizes there to be at least some provisional guidelines or 'heuristics' to be followed in science, though nothing he finds amounts to anything as precise as a 'method'. For both of them, an examination of scientific case studies reveal nothing that can pass muster of being a unique, solid method, attention to which scientists think will lead to the truth. Science is too 'messy' for this, the human-centred activity that it is.

The problem Pitt and Rasmussen are citing here cuts to the core of the concern of this paper. For, with a favoured a priori method in mind, we could probably dig in our heels and find this method used by all scientists, re-interpreting scientific activity where our interpretation demands this. But obviously this will not do, if we want to be objective about scientific practice. So we need to find some small set of core methods that can be observed to typify scientific reasoning, a set which scientists presume to be epistemically valuable in their inquiry into the physical world. And here I want to put forward such a small set. Abandoning rigour for another time, I take to begin with as the basis of my assessment a relatively cursory examination of the journals *Science* and *Nature*, the two gold-standard science journals in circulation today. Pick up any one of these journals and scan the styles of reasoning used by authors. Of course, the articles in these journals are highly technically sophisticated, yet this should not stop us from identifying the general structure of how these authors argue —and what we find is that the hypothetico-deductive (H-D) method is standardly used. In practically every article a hypothesis is set up and observations are sought to confirm or disconfirm this hypothesis. There are exceptions, granted. For example, there are those cases in which the author simply reports on some unusual findings and makes some passing remarks on the implication these findings have for some current scientific theory. Nevertheless, H-D reasoning is quite commonplace and, frankly, I was shocked to see this having been philosophically trained to dismiss a simplistic H-D view of science. Do not scientists know of the paradoxes of confirmation? Do they not realize for some hypothesis H and evidence e that, if H implies e, then H (plus some irrelevant claim) implies e, and so by H-D reasoning e confirms H (plus some irrelevant claims)? Of course, a scientist might wonder why we wanted to introduce an irrelevant claim to begin with. Still, I do not think that any such philosophic worries would dissuade scientists from persisting in H-D reasoning, and I suppose for the sake of scientific progress we should be thankful for this. Apparently scientists persist with method that warrants the scorn of many philosophers, and in historicist fashion we should heed their preference, recommending on the basis of our historical inquiry (which certainly needs more rigorous presentation) the merit of the H-D method.

I offer finally two more examples of styles of reasoning which are commonly adopted in scientific work (again leaving the rigorous historical proof aside), here looking more intently at experimentation. Both these methods are typically adopted by experimenters. Specifically, where one is deploying a technical apparatus in an attempt to observe a particular physical phenomenon, one should check on the reliability of one's technical apparatus, and ensure that this experimental result is found to occur under repetition in relevantly similar circumstances.

Now perhaps these methods are mundane and obvious. But then why do Pitt and Rasmussen overlook them as candidates for inclusion in the standard, unalterable method that is common to all (experimental) sciences? Again and again experimenters aspire to ensure the reliability of their interventionist regimes, and the dictum to produce "the same experimental result under relevantly similar circumstances" is the mantra of all cogent experimental work. Thus, these common and quite specific constraints on experimental work should guide and inform our philosophic investigation into experimental science.

Of course, historical proof of the centrality of these methods would require a thorough laying out of the evidence, something not done here. And in fact in demanding such a thorough proof we are advocating yet another norm of science, as secure as the preceding, to wit, in debating questions of empirical fact (the relevant question here being, "What methods do experimenters standardly use?") one should perform sufficient empirical research before drawing any conclusions. In particular, we need to examine a myriad of cases, in proper empirical fashion, before pronouncing firmly of the ubiquity of methods 1) and 2) above. But what of the requirement to always seek more empirical evidence? Is it a secure method of science? No doubt it is —and it would always be strictly adhered to but for restrictions on time and interest.

Nevertheless, the important point to make here is that, despite the appearance of flux and disorder in science, despite the presumably 'Heraclitean nature' of scientific methodology, we should not as philosophers rest content with the view that we thereby observe the true nature of science. For such an opinion advances a perspective on science that conflicts with the practice of science and the attitudes of scientists regarding scientific method. Scientists, we find, when confronted with a 'flux' of data hardly ever resign themselves to the conclusion that the phenomena they are examining are a product of random, chaotic occurrences. To argue thus —to concede that there is just 'flux' and no natural law-like order— would be, pretty well, to give up the scientific game. Recall with the SARS case above that scientists never entertained the conclusion that the disease vector worked on the basis of heuristics, or had a random changing nature. Recall that my economist colleague above, when presented with an overflowing array of data, sought to 'see through the data' to the reality underneath. It is the goal of science to bring order to confusion, whether the subject matter is virology, astrophysics, economics, anthropology, biochemistry, or whichever field. To be sure, it is still a possibility that, despite their ambitions, the natural world is found to be flux-like and chaotic after all —in fact this may be what quantum physics is suggesting. But it would be foolish to draw this conclusion now and too hastily, as was the wont of the original Heraclitus. Similarly, it should only be with trepidation that one should avow the doctrine of flux in historical philosophy of science. For it should only be after a substantial empirical inquiry has tried to reduce chaos to order and been found to fail that one should accede to the skeptical positions of Pitt and Rasmussen. Has such a substantial inquiry ever been performed? Actually, this inquiry has only just begun —there is a mammoth amount of science to consider, and few philosophers are having a serious look at it.

6. Conclusion

In arguing for the legitimacy of historical philosophy of science I have, in a subtle way, argued on historical grounds. For in responding to both the problems of 'small data' and 'flux' I have made references to the activities and attitudes of scientists: scientists, I have argued, are not deterred by a paucity of data —they will 'see through' to the (believed) reality underneath (and then argue hypothetically on that basis); moreover, they are not deterred by an apparent chaos of phenomena —they will seek rational law-like regularities, in any case, to explain their observations. On the basis of these considerations I have recommended that historicist philosophers of science not be deterred either, buoyed by the actions of their brethren subjects. But is this not itself to argue historically, and precisely in the context where it is the very legitimacy of historicist approaches that is at issue? Indeed it is, so my last appeal is to leave open what one thinks about the 'scientific attitude' with respect to small data inductions and flux. Assuming I am right in characterizing the scientific attitude with regard to these issues (and this is a big assumption, worthy of historical adjudication), is it an attitude to be approved? Such an attitude on my view is wonderfully optimistic and prudent. But perhaps the reader finds the attitude uncritical and naïve. It all comes down to, then, what we think we can learn about epistemology by examining scientists and their work.

REFERENCES

Bloor, D. (1991). *Knowledge and Social Imagery*. 2nd edition. Chicago: University of Chicago Press.

Pitt, J. (2001). "The Dilemma of Case Studies: Toward a Heraclitian Philosophy of Science", *Perspectives on Science* 9, 373-382.

Mayo, D. (1996). *Error and the Growth of Knowledge*. Chicago: University of Chicago Press.

Rasmussen, N. (2001). "Evolving Scientific Epistemologies and the Artifacts of Empirical Philosophy of Science: A Reply Concerning Mesosomes", *Biology and Philosophy* 16 (5), 627-652.

Robert G. HUDSON teaches Philosophy at the University of Saskatchewan. His main area of research is the history and philosophy of science, especially the epistemology of experimentation. He has recently published in such journals as *The British Journal for the Philosophy of Science*, *Studies in History and Philosophy of Modern Physics*, and *History and Philosophy of the Life Sciences*. He is currently working on the issue of realism as it applies to contemporary dark matter research.

ADDRESS: Department of Philosophy, University of Saskatchewan, 9 Campus Drive, Saskatoon, SK S7N 5A5, Canada. E-mail: r.hudson@usask.ca.