REVIEW SYMPOSIA

What is This Thing Called Philosophy of Science?*

Alan Chalmers, *What is This Thing Called Science?* Third edition. Milton Keynes: Open University Press; and St Lucia: University of Queensland Press, 1999. Pp. xxii + 266. A$19.95 PB.

By John Worrall

Alan Chalmers' book has been the best introduction to philosophy of science since it appeared in 1976. It remains so in 1999 courtesy of this newly published third edition. Translated into fifteen languages, the book has been a major force worldwide for straight thinking about science. The book, reflecting its author's own intellectual biography, is especially valuable as a way of introducing graduates in science to its philosophy. Drawing liberally on examples from the history of science, it reveals the central philosophical and methodological issues, not as dry, 'merely academic' puzzles, but as exactly the things that a reflective practitioner ought to worry about. Chalmers writes in a clear, direct, entirely pretension-free style, taking—unusually for an introductory account—a clear stand on virtually every issue he raises, and never being afraid to say that he finds the motivation for some of the opposing positions incomprehensible. While some will object to this feature, the book, in my view, gains from its directness and vigour more than it loses from any lack of evenhandedness.

The third edition contains a somewhat reworked version of the material from earlier editions on observation and experiment, induction, falsification, Kuhn, Lakatos and Feyerabend. A substantially reworked version of the second edition material on "unrepresentative realism" now becomes an extended account of the ongoing realism/anti-realism debate (Chapter 15). The concession to Feyerabend—made slightly mutedly in edition two—that there is no single universal method for science is now highlighted and

* The editor is *very* grateful to Peter Ansty for organising this *excellent* symposium.

explicitly argued (Chapter 11)—more on this issue shortly. Finally, some entirely new chapters aim to bring the book up-to-date as an introduction to the current state of the field—these are on personalist Bayesianism (Chapter 12); and on the "new experimentalism" (Chapter 13), which in turn relates to a re-consideration of the nature of natural laws and of causality (Chapter 14).

Despite my overall admiration of the book, I do find some of this new material problematic. Although the chapter on Bayesianism has its heart in the right place (subjective Bayesianism leaves *too much* scope to subjective opinion to give an adequate account of the nature of scientific reasoning), neither the exposition of the account (relying altogether too heavily on Popper's entirely discredited claim that the 'natural' position is that all universal generalisations have zero probability) nor the argument against it is altogether convincing. (Also, although the chapter is very largely a reaction to it, all references to the book by Howson and Urbach are to the first (1989) edition rather than to the—rather substantially amended—(second, 1993, edition.) Moreover, *neither* the doubts that I have about the extent to which the 'new experimentalism' lends any genuinely new insights, *nor* those I have about the value of analysing nature in terms of dispositions or capacities, were at all laid to rest by Chalmers' treatments.

But rather than deal with these relatively detailed niggles, or indulge in some 'in-house' differences over the interpretation of Popper's or Lakatos' work, I want to focus here on one rather striking feature of the third edition where there is at least the appearance of outright and significant disagreement between Alan Chalmers and myself. As just mentioned, the second edition contained a rather muted concession to Feyerabend to the effect that there is no single, universal, ahistorical scientific method and hence no single category "science". It also contained a brief (though not in my view very cogent) reaction to the suggestion that this concession makes the title of the book, and hence the whole project, somewhat problematic. In the third edition, however, the concession and its consequences are brought considerably closer to centre-stage. There is now a whole chapter (Chapter 11) arguing for "methodical changes in method" and so against a universalist, ahistorical view; and the Epilogue (Chapter 16) is largely taken up with the issue of whether the denial that there is "a general account of science and scientific method...that applies to all sciences at all historical stages in their development" (p. 247) means that the whole project, reflected in the title of the book, is nugatory. And indeed Chalmers, while of course continuing to insist on the importance of the project, does now explicitly admit that "[t]here is a sense in which the question that forms the title of this book is misguided" (p. 247).

This concession would put the book amongst elevated company—for example, Popper's *Logic of Scientific Discovery*, as has often been noted, explicitly denies that scientific discovery *can* be a matter for logic. However, it is the concession not the title that is misguided: Alan Chalmers' title-question has an answer and most of his book consists of a pretty good shot at pointing towards it. Moreover, and more importantly, *if* the concession were necessary then his attempts to insist on the continued importance of the project and on the avoidability of relativism would be hopeless.

Chalmers seems at first to hold a very strong version of the revisability of method thesis, declaring himself "happy to join Feyerabend in regarding the idea of a universal and ahistoric method as highly implausible and even absurd" (p. 161). He believes that that idea is in fact susceptible to straightforward historical refutations: Galileo, for example, according to Chalmers, effected a change in the 'standards of science' when he rejected the reliance on naked-eye observation as "a criterion of science itself" and overruled some such data on the basis of the evidence supplied by his telescope (pp. 163–8). Where Chalmers differs from Feyerabend is in resisting the suggestion that allowing that standards change is tantamount to embracing relativism. If theory *A* is better than theory *B* according to set of standards one, but *B* is better than *A* according to set of standards two, then it would seem that which theory is 'better' will depend on which standards happen to be in force. But in fact, Chalmers claims, we can make sense of the idea of a given set of standards being an *improvement* on another. Rather than facing a relativistic impasse in the case sketched, if set of standards number two is an improvement over set of standards number one, then a (doubly) non-relativistic, 'objective' judgment seems justified: that *B* is the better theory according to the better standards. (This is my reconstruction rather than Chalmers' own but I believe it captures his view.) The Galileo case is exactly one in which the change in standards constituted a clear improvement. Feyerabend was right to reject universal method, but his inference to no method at all is invalid, because the assumption that these are the only two alternatives is false: "[a] middle way would hold that there are methods and standards in science, but that they can vary from science to science and can, within a science, be changed and changed for the better" (p. 162).

Let's put on temporary hold the idea that standards may, at a given time, differ from particular science to particular science; and concentrate, as does Chalmers himself, on the idea that within a given science standards may improve over time. Doubtless there is a sense in which the history of science reveals not only the development of ever better science, but also the development of better ideas about *how to do* science. But the obvious rejoinder—one that I endorsed earlier (Worrall 1988) in discussing Larry Laudan's version of this view—is that it is only if there are some *fixed* general standards that we can make sense of the idea of improvement of more specific standards. Are judgments of 'improvements' in standards adventitious or are they principled? If the former then the position is relativism in disguise. If the latter then the principles that somehow inform judgments of improvements must themselves, on pain of infinite regress, be fixed principles of rationality.

Alan Chalmers in fact cites my position here and takes it to amount to the assertion of the existence of universal 'superstandards' governing alleged changes in more specific methods (pp. 162–3). If this means *meta-level* standards of the form "set of standards one is better than set of standards two if and only if . . .", then it is not an accurate representation of my view. It *is* certainly my view, however, that the only two options are relativism, on the one hand, and the endorsement of some general, *fixed* methodological principles on the other. Since Chalmers cites my view and continues to assert his own apparently quite contrary one, he clearly believes he has some fairly knockdown response. Part of that response (perhaps most of it) is clearly meant to lie in the analysis of the Galileo example. But that example is, I suggest, entirely ineffectual.

First, what exactly was supposed to be the standard that (partially) defined science ahead of Galileo and was rejected as a consequence of his work? Not surely that naked eye observations are totally sacrosanct as true reflections of reality. Everyone, even Aristotelians, knew about 'mis-observations' through drunkenness or bad visibility—but these could, as Chalmers suggests, be dismissed as observations made under 'abnormal circumstances'. But could anyone serious, even long before Galileo, have believed even the modified view that the senses are always direct and accurate reflections of reality *in 'normal' circumstances*? Aristotle knew, for example, about the (apparently) bent oar—this is after all a repeatable effect where all normal observers will assert that the oar *appears* bent, yet surely no one believed that it mysteriously became bent when immersed in water (and bent exactly at whatever point it meets the water's surface) but miraculously regains its rectitude when withdrawn from the water. I am no historian of pre-modern thought, but it seems impossible to believe that the ability to create 'fire-writing' on a dark night with a lit torch was anything other than a generally known phenomenon (just give a [gloved] three-year old a sparkler on Guy Fawkes night—no instruction needed!). Yet again everyone must have known, without of course necessarily explicitly articulating it, that the rings and curlicues we all—perfectly normally and regularly—observe are not 'real', in the sense that they are relational effects, dependent on the features of our physiological apparatus as well as 'objective' features of the external world. We did not need movies to establish this point. (I am aware of course that historians have contradicted this. But historians seem to be committed, by trade union rules, to—often especially naive versions of—historical relativism and are not necessarily to be trusted here. Remember that if underdetermination of theory by evidence applies in physics, it applies with knobs on in history.)

Galileo by 'overruling' naked eye, in favour of telescopic data was not inventing a new 'method', but extending the range of methods that were already known. Moreover, the reason why—as Alan Chalmers correctly insists against Feyerabend—Galileo's case was (gradually and cumulatively) compelling was exactly because no new standard or principle of method was involved. The outcome of Chalmers' (brief but nicely turned) analyses of Galileo on the telescope and moons of Jupiter and on the telescope and the apparent size of planets is in effect simply this: (i) Galileo's claims about the telescope's 'veracity' were in both cases *independently testable* and independently *confirmed*; while (ii) any claim that judgments about planetary sizes based on naked eye observations are invariably accurate is demonstrably empirically inconsistent (that is, claims of accuracy about naked eye observations made in one set of circumstances logically conflict with claims of accuracy about naked eye observations made in others).

The conjunction of the two claims that Jupiter has moons and that the apparent multiple splodges associated with looking through the telescope at the relevant section of the sky are produced by those moons predicts for example (against the background assumption that at least some other planets do not have moons) that similar splodges will *not* be observed when viewing at least some other planets. This, as was pointed out by Galileo himself, is independently confirmed. Moreover—did Galileo himself ever explicitly point this out?—once observations had been made of the disappearance and reappearance of the splodges (attributed of course on the hypothesis that the moons exist due to their periodically disappearing behind the main planet) the theory again makes independent and confirmed predictions about when particular splodges will later disappear and reappear.

Or consider the other case analysed by Chalmers—that involving the apparent sizes of some of the planets. Copernican theory (along in fact in this case with its rivals) predicted, again against the background of commonly held 'background knowledge', that the apparent sizes of Venus and Mars, for example, should vary in a certain well-defined way over the course of time. Naked eye observations of the apparent size did not conform to these predictions; when observed through the telescope the apparent sizes over time were (more or less) right. Must this be left as a stand-off; or can the situation be independently resolved? As Alan

Chalmers in effect shows, it *was* resolved by Galileo's development of a further theory about the source of the inaccuracy of the naked-eye observations—the (initially alleged) phenomenon of irradiation (the extent of which, he further conjectured, depends on the relative brightness of the viewed object compared to its background)—and by his producing *independent* evidence (that is, evidence you had to accept whether you held or denied Galileo's theory of the superior accuracy of the telescope) that this alleged phenomenon is indeed real. Moreover (it really is another aspect of the same methodological feature), Galileo showed that the assumption that naked-eye observation is an accurate reflection of the size of illuminated objects leads to logical inconsistencies. (Venus appears very small before sunset, much larger thereafter; a glowing torch appears much larger at night than it ought to do given its actual [of course, in this terrestrial case, measurable] size and its distance from the observer.)

Of course you *can* at a stretch tell this story as a case of the replacement of one method (roughly rely on naked-eye observation) by another (roughly rely on telescopic observation); and then, in order to defeat the relativist, argue for the superiority of the 'new' method. But there is a much more direct (and universal and much less misleading) way of telling the same story. Sense data are *never* reliable, if that means automatically accurate reflections of whatever objective reality they are supposed to mirror. Any such connecting claim is inevitably theory-dependent. However, the sense data *are* fixed (you can't choose what size Venus or the lit torch appears to have to you) and are always in need of explanation. The explanatory theories can, if we are lucky, be independently tested (against the *raw* sense data). Generally rival explanatory theories are either inconsistent with the data or can be made consistent with data only through *ad hoc* (not further, independently testable) manoeuvres.

Fundamental assumptions—seemingly obvious yet of enormous power —about the importance and significance of successful independent tests, about the unacceptability of any theory that, given the data, is internally inconsistent, not to mention the underlying, and clearly fixed, principles of deductive logic, are unvarying principles that govern this scientific advance, and I suggest all others. Strangely enough, Chalmers eventually concedes this point. He suggests that I "and like-minded people" (please put me in touch!) will respond to his Galileo case-study by pointing out that "an appeal to some higher, more general standards is involved" and that "[o]nce we have spelled out these general assumptions...then it is [they]...that constitute universal method"; and indeed that "without such a backdrop...you cannot argue that the change is progressive" (p. 171). Quite so; and Chalmers seemingly agrees, "I concede that there is a universal commonsense method" (p. 171).

This concession, however, has a kick in its tail. The word 'commonsense' is to be taken very seriously and rather dismissively: although there are such fixed underlying principles, they are so trivial—amounting to not much more than "take argument and the available evidence seriously" (p. 171)—that they are no more than commonsense. If articulating them were all that philosophy of science was about, then it would put him (and me) "out of business, since [they] are hardly the kind of thing that it takes a professional philosopher to formulate, appreciate or defend" (p. 171).

Now, Chalmers misidentifies the crucial, general principles which although general are much meatier than he suggests. The sort of things I have in mind include the principles of deductive logic, and those underlying *both* the appreciation of the status and power of independent evidence *and* general injunctions like "always test your theories against plausible rivals". It is in fact not at all obvious that these *are* parts of 'commonsense', at least if that means what is as a matter of fact commonly believed. On the contrary, fallacies like *post hoc ergo propter hoc* that clash with those principles seem sadly to be endemic, even in 'well-educated' populations. But if the crucial assumptions are, in themselves, fairly minimal, then all the better (though I do not believe that defending them is as easy a task as Chalmers suggests). I of course accept, indeed emphasise, that their immense power is revealed largely when they are conjoined with specific information (that will therefore of necessity have been acquired at a particular time). So, for example, as I showed in my (1988), the powerful *specific* methodological principle that clinical trials should be conducted double-blind and placebo-controlled is a simple consequence of the universal principle that one should always test theories against plausible rivals together with the specific substantive empirical discovery of the placebo effect. This is of course why, despite Chalmers' 'light-hearted' suggestion, philosophers of science would not be out of a job even if (and it is a big if) we all agreed on the basic general principles of good science: there would still be the important task of showing how these general principles are informed by changing background knowledge through history to produce changing *specific* methodological principles. (It is also why his account of my position as involving meta-level 'super-standards' is askew.)

Finally, it is *also* why the difference between Chalmers and myself on this issue might appear to be, in the end, merely one of emphasis. Perhaps so; but then it is an important matter of emphasis. Only when the emphasis is in the right place does his defence of science against the (social constructivist) 'levellers' (p. 172) work. Moreover the mistaken emphasis leads to significant outright mistakes. Crucially, Chalmers slips easily from "changing methods within one science (physics)", to "changes in methods *from science to science*". And he holds that one can make this concession too without surrendering to relativism. Allegedly (p. 248), those who defend 'creation science' are really claiming that the methods employed in that field are similar in certain respects to those involved in physics. But that similarity claim can be assessed by inspection of the two fields: "No universal account of science is necessary" (p. 248). But what if the claim were, as it sometimes indeed is, that the methods of 'creation science', although different from those of physics are 'equally valid'? That the methodological injunction that no theory can be satisfactory if it clashes with a 'literal interpretation' of *Genesis* is 'just as valid' as the methodological injunction that no theory can be satisfactory if it is massively *ad hoc* (while a non-*ad hoc* rival for the same range of phenomena exists)? It is this sort of suggestion that reveals that defenders of scientific rationality *must* assert firmly that there are general principles governing not just physics but all sensible empirically-based attempts to acquire knowledge. It is because it contradicts those general principles that so-called creation science is pseudoscience.

There is *a* thing called science. Despite its many imperfections in practice, blessed be its name! Reading Alan Chalmers' book will give students a good start towards understanding what that thing is (though they should resist taking Chapters 11 and 16 seriously).

Centre for Philosophy of Natural and Social Science,
London School of Economics,
Houghton Street,
London WC2A 2AE,
UK.

By Deborah G. Mayo*

The more philosophers of science have turned their attention to historical episodes in science and to the complexities of actual scientific practice, the more they have come to see the inadequacies in all philosophical accounts of scientific evidence, inference, and hypothesis testing. Attempts to set out formal rules or logics relating statements of evidence and hypotheses by logical relations of confirmation, support, corroboration, and the like either fail to capture actual scientific inference or lack any normative force—or both. Nor have these problems been solved by the attempts to look away from logics of evidence to developing, instead, methodologies for large-scale changes in paradigms, research programmes, and the like. The question now is: Should philosophers of science give up on what has long been held as their primary task: to understand and justify scientific methods for assessing hypotheses on the basis of empirical data? And if they should, what job is left for the philosopher of the epistemology of science? What is this thing called philosophy of science?

Anyone seeking a clear, sophisticated and impressively concise tour of the developments that have led contemporary philosophers of science to this predicament will find Alan Chalmers' third edition of his *What is This Thing Called Science?* a treasure. After taking the reader through the twists and turns of the logics of confirmation and falsification, and the attempts to locate scientific rationality in large-scale theory change, Chalmers takes up this question directly: "Since I have denied that there is a universal account of science available to philosophers and capable of providing standards for judging science...it might be concluded that the views of philosophers of science are redundant and that only those of scientists themselves are of consequence. It might be thought, that...I have done myself out of a job. This conclusion (fortunately for me) is unwarranted ...Scientists are not particularly well equipped to engage in debates about the nature and status of science...such as are involved, for example, in the evaluation of creation science" (p. 252). But after finishing the book, the reader is left in the dark as to how Chalmers' philosopher of science could engage this task. Those who have grappled with it appeal to some kind of general criterion for what counts as a science (for example, falsifiability, testability), and Chalmers claims that (like Feyerabend) he denies there are general "standards that all sciences should live up to if they are to be worthy of the title 'science'" (p. 161).

At most, Chalmers allows, there are "commonsense" universals as exemplified by such general principles as "take argument and the available evidence seriously" (p. 171), but, as he concedes, it hardly takes a professional philosopher to formulate such bland generalities. How then does Chalmers think philosophers of science can help adjudicate "controversies about the nature and status of science"? They can do so, he proposes, by describing historical episodes from acknowledged sciences (for example, physics) in the right ways—the ways that emphasise the epistemological aspects of the episodes. Noting "the similarities and differences" between the disciplines, Chalmers claims, gives us "all that we need for a judicious appraisal" of claims that such and such [for example, creation science] is a science. But does it? How could one pinpoint the relevant features that must be exemplified by an enterprise before it can pass our test (of whether to count it as a science), if Chalmers is correct to deny such things exist? He nowhere answers this question. What leads Chalmers to deny there are overarching principles above and beyond trivial commonplace generalities is that once we demand more detail, "then those details will vary from science to science and from historical context to historical context" (p. 172). But, if this is so, then the mere fact that there are dissimilarities between the enterprise in question, call it x, and the particular historical episodes from physics that Chalmers' philosopher of science describes, cannot carry any weight.

Even though enterprise x has (or lacks) features found in Chalmers' favourite episodes from physics, they may well be features found (or lacking) in some perfectly good science, or in a different episode of physics—or so a defender of the status of x could rightly argue. Without an adequate account of 'good evidence' and what is required to 'take evidence seriously', I see no 'judicious' way to rule on the scientific merits of enterprise x. Fortunately, the assumption that leads Chalmers into this predicament—that the variety and context-dependencies of actual inferences preclude non-trivial norms—has much more to do with the fact that philosophers have not developed adequate accounts of evidence and inference than to a lack of non-trivial norms. Unfortunately, Chalmers does not consider this possibility. For the balance of this review, I will raise and address the following: the example of transgenic pollen, subjective Baysianism, the new experimentalism, error-statistical testing and expecially the matter of severe tests of hypotheses.

So, to cite an example, it has recently been reported that there is evidence that transgenic corn pollen (pollen from corn genetically altered to control certain pests) harms the larvae of Monarch butterflies. Here is a data report: "Larval survival after four days of feeding on leaves dusted with [transgenic] pollen was significantly lower than survival either on leaves dusted with untransformed pollen or on control leaves with no pollen" (*Nature 399*, May 20 1999, p. 214).

In particular, the observed difference in survival rates was improbably far from what would be expected by chance variability alone. This improbability, called the p-value (or statistical significance level), is given as 0.008. By contrast, had the observed difference been fairly likely even if the mortality rates were unaffected by the transgenic pollen, if, say, the p-value had been 0.4, then the data would not be good evidence of an increased mortality rate, even in the conditions of the laboratory. The former method is, and the latter method is not, a fairly reliable indicator of a genuine difference in mortality rates. While the availability and applicability of methods of interpreting data will vary this does not alter the properties of the methods (for example, reliable or not), which are objective, empirical ones. Were statistically significant increases in mortality rates interpreted as evidence that there is no risk posed—perhaps on the grounds of a very strong prior degree of belief that transgenic corn poses no increased risks to untargeted species—then that would be an example of not taking the evidence seriously. (I am distinguishing this inference about the laboratory Monarchs from an inference as to what risks are posed in the field—something that requires further errors to be ruled out.) I see nothing historically or contextually relative about this and similar principles. Although this is just a particular illustration of a method from standard statistical practice (and I can only be very sketchy in describing it here), it illustrates what is true for methods and strategies for obtaining and interpreting data in general.

There is plenty of evidence that Chalmers agrees, and it is much to his credit that he shows (in Chapter 12) that the subjective Bayesian emperor has no clothes. Especially when they are responding to criticism, subjective Bayesians stress the extent to which both the prior probabilities and the evidence which needs to be fed into Bayes' theorem are subjective degrees of belief about which the subjective Bayesian has nothing to say. But to what extent can what remains of their position be called a theory of scientific method (p. 192)? "A good theory of scientific method...will surely be required to give an account of the circumstances under which evidence can be regarded as adequate, and be in a position to pinpoint standards that empirical work in science should live up to" (p. 191).

Indeed. But for this criticism to have any weight, Chalmers needs to show how we can pinpoint (normative) standards—the very thing he suggests we are in no position to do. Can an appeal to error statistical methods rather than to Bayesian ones be the basis for a more adequate account of scientific method and inference? In discussing the "New Experimentalism" (Chapter 13), Chalmers occasionally hints that it might.

A theme running through the New Experimentalism chapter is that in place of the familiar logics of evidence (confirmation theories and inductive logics) we should focus on how experimental knowledge is actually arrived at and how it functions in science. Promising as much of this work has been, nothing like a general account of evidence and inference has been forthcoming. The reason the New Experimentalists have come up short, it seems to me, is that the aspects of experiment that have the most to offer in building an account of evidence and inference are still largely untapped: designing, generating, modelling and analysing experiments and data, activities that receive structure by means of standard statistical methods and arguments. The New Experimentalists (while hardly a homogeneous group), seem, by and large, to be too haunted by the ghosts of probabilistic logics of evidence to appeal to statistical methodology altogether. But the methodology they are forfeiting is crucially different from the logics of evidence. For one thing, rather than start with given evidence these methods direct themselves to the tasks of generating modelling and analysing data to obtain evidence in the first place. Second, in striking contrast to the logics of evidence or confirmation, probability arises in these statistical methods, not to measure degrees of credibility or support to hypotheses, but to characterise properties of tests and estimation procedures: how reliably a test is able to detect a given type of error, namely, the test's error probabilities or error characteristics. An account of evidence and testing based on error probabilities may be called an error-statistical account.

Although Chalmers looks favourably on the fruits of the error statistical account (for solving a variety of problems) the reader is not told that there is a well-worked-out battery of statistical techniques that serves as the fundamental basis for those methods. This is very unfortunate: there is a pressing need to bring these methods more squarely into philosophy of science—not just for their value to the philosophical tasks of evidence, but also, in 'the other direction' as it were, to help disentangle a host of philosophical conundrums faced by users of these methods (especially in sciences where the uncertainties are greatest). Given the alleged commitment of contemporary philosophers to the actual practices of science, it is especially surprising to find philosophers overlooking a standard set of inferential methods rather than trying to understand when and why scientists find them so useful. The time is ripe to remedy the situation: the conglomeration of statistical techniques (Neyman-Pearson tests and confidence intervals, Fisherian tests, non-parametric methods, data analysis and others) is the place to look for erecting an adequate philosophy of evidence and inference.

Granted, using these techniques to build a philosophy of evidence requires a good deal of work above and beyond any statistical texts. Most broadly put, the task for philosophers of science is to consider how to relate statistical hypotheses tests, and methods of data generation and modelling, to substantive scientific hypotheses and actual, messy, data. There is an overarching goal that may guide us in articulating these statistical-substantive links, the desire for severe tests, for severely learning from error. Impressively, Chalmers arrives at the key idea with non-technical ease: "[a] key idea ... is that a claim can only be said to be supported by experiment if the various ways in which the claim could be at fault have been investigated and eliminated" (p. 199). Keeping this central and informal idea in mind can avoid many perplexities that others generate when they take the formal statement of severity out of context and rush to find (alleged) "counterexamples". Nevertheless, this informality can result in some misunderstandings. In order for a claim or hypothesis to have passed a severe test, Chalmers writes, it "must be such that the claim would be unlikely to pass it if it were false" (p. 199). This is correct, but one must be careful not to leave off the requirement that for hypothesis H to pass the test with outcome e, H must "fit" e for an appropriate notion of fit. Chalmers omits the requirement, I think, because he is very sensitive to the fact that there are many cases where e severely passes H even though $P(e|H)$ is low, as he discusses in his appendix to Chapter 13. True, H's passing a severe test with e does not require $P(e|H)$ to be high, but it must be higher than $P(e|\text{not-}H)$.[1] But how does one assess severity? Again, failing to allude to statistical methodology leaves the reader without an answer.

Although I do not wish to limit severity assessments to formal statistical hypotheses, the statistical framework gives crucial guidance for both formal and informal severity assessments. Most importantly, it teaches us that it is impossible to assess reliability or severity with just statements of data and hypotheses divorced from the experimental context in which they were generated. Minimally, we need to consider three main elements of experimental inquiry which we can represent as three types of models: models of primary scientific hypotheses, models of data, and models of experiment that link the others by means of test procedures. The primary question in our transgenic corn example above might be: does transgenic pollen harm Monarch butterflies in the field? It is tackled by asking a specific statistical hypothesis about a sample of larvae in a given experimental set-up: Is there a statistically significant increase in mortality among the sample fed transgenic pollen (treated) in contrast to the sample fed non-transgenic pollen (controls)? This is probed by considering, within a model of experiment, a standard null or error hypothesis, *Ho*, any observed difference in mortality rate between treated and control larva are 'due to chance'.

The data are modelled as the difference between mortality rates in treated and non-treated larvae. However, for the statistical inference to go through, we need to determine if they were generated in such a way that the treated and non-treated larvae are 'like random samples' from a data generation procedure where, at the start of the experiment, both groups have the same probability of mortality. By means of an interconnected set of inferences and checks, the statistical claims that are severely passed can teach about the primary hypothesis of interest.

Where tests are appropriately severe it is possible to learn from rejections and falsifications: one obtains real effects that will not go away, and in this way experimental knowledge grows. To violate the severity requirement is not to 'take evidence seriously', and if a given enterprise is regularly unable or unwilling to develop sufficiently controlled inquiries so as to distinguish different sources of error, real effects from artefacts, signal from noise, *etc.*, then it will be hindered or prevented from making progress in knowledge and its scientific credentials will be rightly questioned. In several places throughout this book, Chalmers endorses these ideas about learning from error and severe testing, and one would have thought he would put them to use in the task he regards as central for philosophers of science.

Perhaps the reason he does not is to be found in his last chapter. Here, Chalmers questions whether these ideas serve as the basis for a general account of scientific inference, because he thinks: 1) "the emphasis on experimental manipulation involved in the New Experimentalism renders that account largely irrelevant for an understanding of disciplines, especially in the social and historical sciences" (p. 250); 2) it is incomplete until it is augmented "with a correspondingly updated account of the role or roles of theory in the experimental sciences" (p. 251). Let me consider these in turn:

1) The first allegation is quite unwarranted, at least insofar as it is being alleged of the experimental account I recommend. From the very start I say "I understand 'experiment'...far more broadly than those who take it to require literal control or manipulation. Any planned inquiry in which there is a deliberate and reliable argument from error may be said to be experimental" (Mayo 1996, p. 7). The whole point of appealing to statistics, as I emphasise repeatedly, is that it enables us to model "what it would be like to control, manipulate, and change in situations where we cannot literally" do any of these things (Mayo 1996, p. 459). Nor are these empty promises: I make use of numerous examples that rest upon, not literal manipulation, but computer simulations, manipulations 'on paper', and other tools of the statistical trade for obtaining reliable data and severe tests by analogy to what literal controls afford. For example, if one can distinguish, through analysis, the factors responsible for a given effect, one is not hampered by being unable to hold each fixed. That is why it is so important for philosophers wrestling with problems of method to understand statistical methodology. By giving short shrift to the statistical component of the error statistical account of experiment, Chalmers completely overlooks its key features.

2) As to his second caveat, I happily accept Chalmers' urging to augment the error statistical account so as to relate "the life of experiment" to the "life of theory", and experimental knowledge to theory testing, but I would reject his suggestion that I should embrace the comparativist account of theory testing he recommends. "[Mayo's] argument for scientific laws and theories boils down to the claim that they have withstood severe tests better than any available competitor. The only

difference between Mayo and the Popperians is that she has a superior version of what counts as a severe test" (p. 208). For example, Chalmers claims I must implicitly be endorsing the position that it was warranted to accept the General Theory of Relativity (GTR), as a whole, on the basis of the eclipse results—until such time as an alternative gravity theory was available. But I do not endorse such a position. Granted, since a large-scale theory may, at any given time, contain hypotheses and predictions that have not been probed at all, it would seem impossible to say, about such a large-scale theory, that it had severely passed a test as a whole. But if one were to allow, as Chalmers recommends, that we nevertheless regard the large-scale theory as well tested, simply because no known competitor does better, one would forfeit the very fruits of the piecemeal account of severe testing that leads Chalmers to regard it as superior (for example, to Popper's). In particular, it would take us back to the problem of Popper's account of testing—namely being unable to say what is so good about the theory that (by historical accident) happens to be the best tested so far?

We can give guarantees about the reliability of the piecemeal experimental test, but we cannot give any guarantees about the reliability of the procedure: go from passing a hypothesis H, a proper subset of theory T, to passing all of T. Indeed, this is a highly unreliable method—anyway, it is, entirely unclear how one could ever assess this. By contrast, we can apply the severity idea because the condition "given H is false" (even within a larger theory) always means given it's false with respect to what it says about this particular effect or phenomenon. I am not denying that there may be licence to go from one severely tested claim to others; the ability to do so is a very valuable and powerful way of cross-checking and building on results. However, whether these connections are warranted is an empirical issue that has to be looked into on a case by case basis, whereas the comparativist is saying we are licensed to do this so long as theory T is the best tested so far.

The second important feature of the severity account that is given up by the comparativist (in Chalmers' sense) is that of stability. Suppose an experimental test is probing answers to a question: What is the value of this parameter? Then if a particular answer or hypothesis is severely passed, this assessment is not altered by the existence of a theory which gives the same answer to this question. More generally, in the error-statistical account of testing, if two rival theories, T_1 and T_2, say the same thing with respect to the effects or hypotheses that are being severely tested by experiment E, then T_1 and T_2 are not rivals with respect to E—no matter how much they may differ regarding domains or concepts not probed by E. Thus, a severity assessment can remain stable through changes in 'higher level' theories. By contrast, as soon as an alternative theory comes to light that does as well as *T* does on this (and other tests), the comparativist would regard *T* as no longer severely tested.

Eager as Chalmers (and other comparative-holists) are to license accepting an entire large-scale theory, they ignore what for our severe tester is the central engine for making progress, for getting ideas for fruitful things to do next, to learn more. Rather than asking, Given our evidence and theories, which theory of this domain is the best? we ask, Given our evidence and theories, what do we know about this phenomenon? Far from allowing ourselves to say the full theory GTR is well-tested, our severe tester would set about exploring just why we are not allowed to say that GTR is severely probed as a whole in all the arenas in which gravitational effects may occur. Even without having full-blown alternative theories of gravity in hand, we can ask (as they did in 1960): How could it be a mistake to regard the existing evidence as good evidence for GTR? To this end, a set of related experiments was modelled within what was called the parametrised post-Newtonian, or PPN, formalism. The PPN framework sets out a list of parameters that allow a systematic articulation of violations of, or alternatives to, what GTR says about specific effects. These alternatives, by the physicist's own admission, were set up largely as straw men with which to set firmer constraints on these parameters. Whereas it's not even clear, from the comparativist point of view, what motivation there would have been for deliberately erecting rival theories to GTR in 1960—after all, GTR was not facing anomalies—it is motivated from the point of view of getting more experimental knowledge about gravity, for this was the only way to extend the regions that could be said to have been severely probed. It was the only way to learn more about gravity.

Suitably massaged results of astronomical observations, organised into appropriate data models, supply the measured values of those parameters which could then be compared with the different values assigned to it by the diverse theories of gravity. In this way, in each particular solar system experiment the same PPN model of experiment mediated between the data and several alternative primary models, based on GTR and its rivals within the class of (metric) theories. What is most interesting and most deserving of greater attention are the strategies by which a primary theoretical parameter about the given post-Newtonian parameter (for example, the deflection of light by gravity) is probed by turning it into a claim or hypothesis about a statistical parameter (in the experimental model). Then, inferences from (statistically modelled) data to these statistical distributions were used to learn answers to the primary questions. Putting together the interval estimates, they constrain the values of the PPN parameters and thus squeeze the space of theories into smaller and smaller volumes. In this way they could rule out entire chunks of theories at a time (namely, all theories that predict the values of the parameter outside the interval estimate). By getting increasingly accurate estimates, more severe constraints are placed on how far theories can differ from GTR, in the respects probed. Although we may not have a clue what the final correct theory of the domain in question will look like, the experimental knowledge we can obtain now gives us a glimpse of what a 'correct' theory would say as regards to the question of current interest, no matter how different the full theory might otherwise be.

There are plenty of important philosophical issues in these inferential and modelling strategies that cry out for philosophical elucidation, but one thing is for sure: by turning a blind eye toward the methods and models of statistical data analysis, modelling, and inference, this thing called philosophy of science will continue to wring its hands and lament its emasculation in the face of controversies about the nature and justification of scientific knowledge.

* I gratefully acknowledge the support of an NSF Scholars Award during 1998–9. I also benefited greatly from lengthy correspondence with Alan Chalmers.

1. Some points about my notation: $P(e|H)$ is not the usual "conditional probability" but rather the probability of outcome e under the assumption of, or according to, the assignment given in statistical hypothesis H. There is no prior probability assignment to H. "Not-H" is not the so-called catchall hypothesis. It is not even a disjunction of hypotheses. It is the denial of a specific hypotheses H, for example, if H asserts a parameter is less than m, not-H asserts it is greater than m.)

Department of Philosophy,
Virginia Polytechnic Institute and State University,
Blacksburg, VA 24061,
USA.

By J. J. C. Smart

Alan Chalmers' book is a truly excellent introduction to the philosophy of science. He writes lucidly and with a charming (but I think excessive) modesty, and he is able to make use of his early experience when he was an experimental physicist. A large part of the book is a fine critique of the ideas of Popper, Kuhn, Lakatos and Feyerabend, and the last three of them at least were heavily concerned with illuminating and testing ideas about the nature of science and its relation to reality by means of historical examples over the ages. This is one *genre* of writings on the philosophy of science. There is another *genre* of introduction to the subject which may be epitomised by C.G. Hempel's influential *Philosophy of Natural Science* (perhaps now a little dated) which is more directly problem oriented. Popper's department in the London School of Economics has produced much work in which the history of science is put to the service of philosophy of science. It seems to me to be a little bit curious that the author of *The Poverty of Historicism* should have had such an effect among those who had most contact with him. In fairness, however, I need to say that Lakatos expressly rejected such a charge of historicism. Lakatos had a questionable distinction between 'external history' and 'internal history'. Kuhn (1971, p. 143) characterised the latter as "not history at all, but philosophy fabricating examples".

Scientific Realism

Chalmers defends what may be regarded as a weakened form of what is called "scientific realism", the view that the entities postulated in physical theory really exist and that physical theories are not mere computational devices for predicting observations from observations. I myself would add to the connotation of 'scientific realism' the contention that plausibility in the light of total science is an important guide to metaphysical truth. An important argument for scientific realism is that if the theoretical entities did not exist it would be a cosmic coincidence that the facts on the observational level should be as they are, namely just as *if* the theoretical entities *did* really exist. A variant of this argument is that realism provides the best explanation of the success of theories at the observational level. At the top of p. 238 Chalmers characterises 'scientific realism' roughly in this way. However lower down the page he speaks of the testability of realism against the history of science.

Now I wonder whether this kind of testability is needed. I wonder whether this does not smack of historicism in Popper's sense, as in his *The Poverty of Historicism*. Would it not be better simply to look at science as it is at the beginning of the twenty-first century and see how likely it is that much of it will go the way of Ptolemaic astronomy or the phlogiston theory, or whether Newtonian mechanics and gravitational theory has been overthrown by Einstein? Certainly Newton's laws had to be modified by special relativity, since they are not Lorentz invariant, but this is hardly an ontological difference, and Newton's mechanics is approximately correct in the domains to which they are applied. Moreover special relativity strengthened Maxwell's theory (though not Maxwell's hypothesis of the ether) because Maxwell's equations are Lorentz invariant.

One trouble about testing scientific realism in the light of history is that science has changed so much over the centuries. The Ptolemaic cosmology is so anthropocentric and foreign to us now. We know the distances of sun, stars, and galaxies so that it does not seem at all surprising that the Ptolemaic theory *was* overthrown. This should not suggest that all contemporary physical theories will one day be overthrown. Indeed we can see that the predictive success of the Ptolemaic theory was not very much of a coincidence, because the theory was cosmically parochial and moreover it could save itself from refutation by continually adding epicycles. It hardly presents a cautionary tale for contemporary physics. In the eighteenth century only one of the four fundamental forces, namely the gravitational force, was at all well understood. Nowadays there is so much physics that we can plausibly think will never be overthrown. And insofar as scientific realism is a *metaphysical* theory, plausibility is the most that we should claim for it. It should be noted that 'plausible' is an epistemological epithet, not the name of a third truth-value. What is plausible is most likely *true*.

History of Science and Unjustified Scepticism

Too much concentration on the history of science can lead to unjustified scepticism and even relativism. There is a great body of scientific assertions that are very unlikely to be overturned. If a philosopher is sceptical about them we should be even more sceptical about the reasons he or she gives for scepticism, even reasons derived from the history of science. Vulgarised Kuhn (not necessarily Kuhn himself) has done much damage, not so much in philosophy departments as in social science departments and other humanities ones. Chalmers will have no truck even with more sophisticated relativism.

Here are some assertions that I think will never be overturned. Water contains atoms of hydrogen (or one of its isotopes). Electrons have a charge of approximately 1.602×10^{-19} coulombs. The transparency of glass is partly due to the fact that it is not crystalline. Neutrinos exist. $E=mc^2$. Space-time is curved near massive bodies. One could go on and on. In 1966 the physicist Gerald Feinberg published an article about what he called "The Thales Problem" (Feinberg 1966). This is the problem of the properties of ordinary bulk matter, that which makes up tables, chairs, stones, planets, rivers, seas, and so on—the things with which Thales was familiar. We could add on the properties of other things such as the upper atmosphere and certain plasmas. These properties can be explained by the physics of the electron, proton, neutron, neutrino and the photon. Feinberg takes the intra-nuclear forces as given phenomenologically: to explain these it is necessary to postulate exotic and transitory particles such as can be produced experimentally at very high energies. Indeed in order to understand protons and neutrons, physicists have to go deeper and postulate quarks. Feinberg says that we do not have to go deeper to solve the Thales problem, the problem of the nature of ordinary matter.

We should also suppose that 'ordinary matter' does not necessarily mean 'the commonest matter'. It may well be that the commonest matter is the seething mass of virtual particles, continually coming into and out of existence in what we commonly think of as the vacuum of intergalactic and interstellar space. No matter, this has little to do with the Thales problem and gives us no reason to believe that the physics of ordinary matter will be overturned. So Feinberg argues. Anything further, such as quantum field theory and the physics of exotic particles or the attempt to unify the four forces by means of string theory, is extra. If electrons are made up of strings this does not mean that they do not exist. If J.A. Wheeler's beautiful but unfortunately unsustainable conjecture that everything was made up of the ends of wormholes in a multiply connected space-time had been correct (how Spinoza would have loved it!), this would not have meant that electrons and protons did not exist—they would exist as ends of wormholes. On this way of looking at things, advances in physics by and large (though of course not invariably) add to knowledge of the universe by going deeper without overthrowing what has gone before.

"Hold on", the anti-realist might protest, "there is no agreement about the interpretation of quantum mechanics. So all we have now is a purely instrumental theory". I am slightly suspicious of the word "interpretation" here. I would prefer "understanding". I do not see that better understanding would involve rejection of the entities ostensibly referred to in the earlier discourse. I read in Steven Weinberg's popular book *Dreams of a Final Theory* (Weinberg 1993, p. 115), that a particle can be in a state which is neither definitely electron nor definitely neutrino until there is a measurement of a property such as electric charge which distinguishes the two. (A bit like the cat paradox.) This does not mean that we are wrong to talk about electrons. Consider J.J. Thomson (who in a sense discovered the electron) and Dirac talking and using the word "electron" in their conversation. Dirac had a lot of ideas about the electron that I presume that Thomson might not have understood. Nevertheless there would be a great number of sentences in common to which they would both assent and this would be enough for us reasonably to say that they were talking about the same entities. It would be misleading to say that they meant different things by the word "electron": we should have a Quinean distrust of the notion of meaning here. What we should say is that the predicate "is an electron" as used by Thomson and Dirac respectively has the same extension. Again if in the face of the problem of the Bell inequality, hidden variables could be restored in quantum mechanics by means of backwards causation (temporally reversed correlations) as suggested by Huw Price (Price 1996, Chapter 9) this would alter our understanding of quantum mechanics but not our belief in the reality of the particles in question.

Chalmers' Realism

In this sort of way I would defend realism about the theoretical entities of physics against Kuhnian examples from the history of physics. Chalmers' own position in defence of realism in physics is rather a complex one. He has a correspondence theory of truth. I am persuaded by Donald Davidson's early article "True to the Facts" (Davidson 1984) that what hooks language on to the world is not picturing or correspondence but is Tarski satisfaction of predicates by objects or sequences of objects, as "red" is satisfied by a ripe tomato or "loves" may be satisfied by the ordered pair (John, Mary). But this is still in the spirit of the correspondence theory. Chalmers defends a sort of realism that he calls "unrepresentative realism" and which is similar to a position that John Worrall has called "structural realism".

The main point seems to be that when one physical theory is replaced by another much of the mathematical strcuture is retained. My worry about this is that equations by themselves do not state laws. Furthermore a realist should want to preserve reference, for example to electrons or neutrinos. The mathematics in laws contains reference only to real; and complex numbers, vectors, tensors and so on. For simplicity, and ignoring the fact that force and acceleration are vectors, consider Newton's second law of motion. This is stated as an identity between the real number which is the force in newtons and the product of the real numbers which are the mass in kilograms and the acceleration in metres per second per second. Thus the law states a contingent relation between real numbers which cannot be understood purely from the mathematical equation. If one accepts Quine's naturalistic Platonism here, we have an ontology of real numbers and other mathematical entities, but for a meaty realism one must bring in reference to non-mathematical physical entities. Perhaps the best that I can do to save realism is on the lines suggested by my fiction above about J.J. Thomson and P.A.M. Dirac. I think that this goes beyond unrepresentative or structural realism, at least as stated in the present book. I also concede that there is some ontological difference between Newton's gravitational theory and Einstein's insofar as Newton's theory postulates forces and Einstein's is more geometrical. Nevertheless, the notion of gravitational force could be replaced by that of the curvature of space-time correlated with the masses of the gravitating bodies, but there would still be important differences in the mathematical structure. So I concede that there is a problem for me and perhaps Chalmers here.

I conclude that not only is Chalmers' book a fine introduction to philosophy of science but is also a particularly challenging one. As an attempt to give a necessary and sufficient set of conditions for something to be called a 'science', it ended, as Chalmers recognises, with failure, just as Plato's great dialogue, the *Theaetetus*, ended in an instructive failure to define the word 'know'. In both cases this failure should really be seen as success. In the present case I suggest that the reason for the impossibility of giving necessary and sufficient conditions for something to be 'science' is that the word is what Wittgenstein called a 'family resemblance' one.

Philosophy Program,
Research School of Social Sciences,
The Australian National University,
Canberra, ACT 0200,
Australia.

By Barry Barnes

When one encounters a lucid, informative and well-organised textbook on whatever subject, it is appropriate to celebrate the rare event, and not to pick fault. Or so it seems to me, as someone inclined by temperament to make the most of what we have, rather than to single things out for criticism: no doubt I should have made an indifferent philosopher of science. In this review symposium, however, other contributions will surely do justice to the many indubitable merits of Alan Chalmers' book, and this makes it possible to dwell upon problems and reservations here, without the risk of conveying a false overall impression.

The very title of the book constitutes a significant problem. I remember chancing upon it, I think in 1980, whilst searching for teaching materials, and can recall my disappointment that so little on the "thing called science" was to be found in it. I did eventually find what I had been looking for elsewhere. John Ziman's wide-ranging natural history of the thing called science was another fine text for students that became available at about this time. Oddly, however, this book was entitled *The Force of Knowledge*, even though it said relatively little about what might make knowledge claims credible, or compelling, or rationally justified—which is the central concern of Alan Chalmers' book. It is pointless now to suggest that Ziman and Chalmers might usefully have traded the titles of their books, but it is relevant and interesting to ask what the connection is between the "thing" (essentially, a collective enterprise) described by Ziman, and the philosophical problems (essentially problems of individual reasoning and individual inference) discussed by Chalmers.

Chalmers' text said nothing of the nature of this connection, and perhaps it was wise to refrain from doing so, since the matter is complex and has given rise to vexatious controversy in the contemporary literature. On the other hand, other issues both more complex and less important were dealt with in the text, and arguably it did deserve to be brought to the attention of readers together with the dilemmas to which it gives rise. Was Chalmers seeking to discuss arguments and forms of inference acknowledged as exemplary by scientists themselves, and in that sense part of "the thing" he sought to describe? If so, in what sense was his project philosophical, rather than historical, descriptive and empirical? Or was he seeking to set out how scientists ought to reason and what kinds of inference they ought to permit themselves to be moved by? In this case the question becomes why it was felt necessary to make any reference to the "thing called science" at all, and why the resulting philosophical account should not have been made the basis for a radically critical independent appraisal of the actual practice encountered in the "thing called science".

In truth, Chalmers' text was neither subservient to the history of science nor independent of it. It was a review of what might reasonably prompt the acceptance [or use] of theories, backed partly by formal arguments and partly by exemplary accounts of the history of "the thing". And, crucially, these accounts were not selected by any method or principle designed to make them either representative or reliable. Indeed they were not selected according to any disciplined procedure at all, such as could be explicitly described. The author had effective discretion in the selection of these accounts of exemplary scientific practice, and could use it both to identify what he found valuable and worthy of admiration in the history of "the thing", and to associate the authority of "the thing" with favoured philosophical doctrines and perspectives. To this extent, his textbook in philosophy of science had something of the character of a textbook of natural science, as famously described by Thomas Kuhn. It gave an account of exemplary practices and procedures, to readers who could not themselves as neophytes hope independently to evaluate them, and who were accordingly assisted by appropriate associations with the authority of (the thing called) science.

There was, of course, an important difference between Chalmers and the natural science textbooks that Kuhn described. The latter presented just one authoritative perspective, whereas Chalmers reviewed, and criticised, a number of them. Indeed the constant interaction of exposition and criticism constituted the dramatic form of the book, with marvellously plausible arguments being propounded one after another, only to be undermined one after another and sent crashing down. And it was surely this, together with the insistent retention of a touch of rationalist hopefulness through all adversity, that gave the book its particular charm. Nonetheless, the book was not by any means a disinterested review of philosophical doctrines: there was a discernible bias toward the theories of Popper and his followers: "the Popperian approach is infinitely better than the approach adopted in most philosophy departments that I have encountered" (p. xii). (It is worth bearing in mind here, in defence of what on the face of it is an unduly strong contrast, that where the probability of any theory of a thing being true is zero, as Popper proposed, multiplication by infinity can be an important operation.)

Chalmers' text was by no means that of a devoted Popperian, and indeed it was at variance with a rigorously falsificationist approach, both in its use of historical examples to confirm and support philosophical theories and in its own theoretical preferences. Chalmers was clearly more attracted by the philosophy of Imre Lakatos, with its more realistic attitude to inductive inference. It is an enthusiasm which remains in the present edition, which offers an improved formulation of Lakatos' idea of scientific progress, as expressed in his well-known methodology of scientific research programmes: "a [research] programme is progressive to the extent that it makes natural, as opposed to novel, predictions that are confirmed" (p. 141). Even so, it remains fair to say that it was the story of the philosophy of science as seen from within the Popperian tradition that the text presented, and that the strongly rationalist and anti-empiricist perspective of that tradition profoundly conditioned its selection, interpretation and evaluation of materials.

The 'best seller' status of the text is a clear indication of the important role it has played in the education of readers of many kinds, in many different settings and cultures. Naturally, opinions on its educational value must vary, simply because they vary on the merits of the perspective that informs it. Those who are not Popperians, and who do not recognise that position as infinitely better than their own, are scarcely likely to be as enthusiastic for the work, however finely written and skilfully ordered, as those who are. My personal view is that the book has been the basis of many valuable courses for natural science students. If the empirical psychologists are to be believed, natural scientists are not outstandingly skilled in matters of logic and secure inference, and reflection on the gaps and deficiencies in the kinds of inference they routinely use can be salutary. At the same time, however, and crucially, natural scientists have direct experience of the "thing called science", which indeed they constitute. And this can offer them some protection against the worst excesses of rationalist philosophy, at least until retiring age. Sadly, however, those in the arts, humanities, and "social sciences" are not so protected, and in these contexts there have been serious disadvantages to approaching the "thing called science" via the kind of philosophical perspective so well expressed by Chalmers.

These have nowhere been better exemplified than in the long series of debates on such issues as whether economics, or sociology, were or could have been sciences, or whether Marxism, or rational choice individualism, were genuinely scientific theories. What the participants in these almost invariably sterile debates tended to do was to hurl parodies of the philosophies of science of Popper, for example, or Kuhn, or Lakatos, at each other. What just conceivably might have helped them, and what they seemed to lack, was any proper sense of the "thing called science" as an activity. It goes without saying, of course, that neither Chalmers nor the philosophers he popularised were to blame for this mode of use, or misuse, of their thought; and indeed it should not be forgotten that philosophy of science has traditionally been taught in close association with courses on the details of its history. Those in the "social sciences" who debated the standing of their disciplines wholly in terms of philosophical abstractions did so out of preference and not necessity, and sometimes chose to ignore substantive work on science within their own fields in so proceeding.

It remains to take note of this new edition of the book, with a text substantially different from that original version to which the above discussion is largely related. There is space here only to comment on the overall effect of the various changes that have been made. It is clear that they are not designed to secure a closer engagement with the "thing called science". There is still, for example, far too little discussion of science as collective activity, dependent on shared knowledge and generally accepted procedure, and constrained in deeply interesting ways by that. And the style proclaims, more than ever, an affinity with the verbal culture of philosophy and the humanities, and alienation from that of the sciences, with their much richer and more diverse range of communicative resources. Read pp. 194–5, for example, and consider whether the discussion there does not scream for a diagram, so loudly perhaps that only a philosophical discussion would see fit to omit one. The entire book, indeed, continues to manage on just three diagrams, which is an extraordinary achievement in the worst sense.

What the changes have been designed to do is make the text more reputable, demanding and wide-ranging as a technical discussion of philosophy. And all that can be said against this perfectly legitimate objective is that a little in the way of coherence and textual unity has had to be sacrificed in pursuit of it. All the editions of the text are nicely structured in their earlier chapters, which move in a sweet dialectic from inductivist, to falsificationist, and thence to structural and holistic accounts of scientific theories. But all have difficulty making an ending. The first edition included a gesture toward Louis Althusser, quickly recognised as less than satisfactory. The second turned to the problem of realism, which, for all the interest of the topic, had the effect of opening up new issues and not of bringing the book to a natural close. The present edition now extends across a number of additional, apparently distinct and separate, topics, so that the closely connected chapters of the first part of the book are supplemented by more of a *pot pourri* in the second, designed to introduce some of the issues that have more recently come to interest philosophers of science. Specifically, there are now chapters on Bayesian approaches, and on the "new experimentalism" of Galison, Mayo and others, both of which will surely be found useful. And there is a chapter on scientific laws, that seriously disappoints by failing to come to grips with the interesting work of Nancy Cartwright that it cites as its inspiration.

Conscious perhaps of the multifarious character of the later part of this third edition, Chalmers has also added an epilogue to round off the work. In it he offers his own thoughts on some of the issues raised in this review, as well as a few brief critical remarks on the new philosophical approaches he has discussed. The criticisms, however, are piecemeal, and do nothing to allay the feeling that the book has come to an end in an unsatisfying way. And it is with this criticism that the present review would have come to an end, had I not chanced to glance, at this point, at the back of the cover. The book, so the blurb there tells us, is an "account of modern scientific attempts to dethrone empiricist thought". This final reminder of the rationalist bias of the text prompted the thought that there might conceivably be more of general significance in the recent philosophical developments Chalmers describes than his discussion is ready to acknowledge. Bayesian approaches, the new experimentalism, and some of the recent discussions calling into question the fundamental and universal character of scientific laws are all alike in that they raise once more some of the problems and concerns that formerly inhabited empiricist thought. Perhaps, taken together, they suggest the beginnings of a swing away from what had become in some contexts an overweening rationalism, and an increased readiness once more to acknowledge the virtues of empiricism in the philosophy of science. If so, it would not be before time.

Department of Sociology,
Amory Building,
Exeter University,
Rennes Drive,
Exeter EX4 4RJ,
UK.

Author's Response

By Alan Chalmers

The most sustained criticism of the views expressed in the new edition of *What is This Thing Called Science?* comes from Deborah Mayo and John Worrall on the issue of a universal scientific method, the existence of which I deny. They suggest that my denial implies a kind of relativism or scepticism that I wish to avoid. Interestingly the versions of universal method that Mayo and Worrall offer as an antidote are significantly different. When it comes to specific issues, the status and significance of Galileo's telescopic evidence, Arago's white spot as evidence for the wave theory of light, the significance of Perrin's experiments on Brownian motion, or whatever, then I suspect that the views of the three of us would pretty much coincide. That is, we combat relativism in a similar way when it comes to context-specific cases. It is only at the most general level that we diverge. My reluctance to see a primary role of the philosophy of science to be the formulation of a universal scientific method stems from the ease with which this strategy can lead, and has led, to an extreme form of relativism with respect to science. If some overly general positivist or falsificationist formula is taken as defining science then the opportunity is open for the "levellers" to claim that physics, say, does not qualify as science because it does not conform to the formula or that other areas, such as witchcraft or creation science, qualify equally well. The response of both Mayo and Worrall is to try to formulate a better account of science that does not suffer from such a deficiency. However, it seems to me that neither of my critics have

