Brit. J. Phil. Sci. **60** (2009), 303–344

How Science Textbooks Treat Scientific Method: A Philosopher's Perspective

James Blachowicz

ABSTRACT

This paper examines, from the point of view of a philosopher of science, what it is that introductory science textbooks say and do not say about 'scientific method'. Seventy introductory texts in a variety of natural and social sciences provided the material for this study. The inadequacy of these textbook accounts is apparent in three general areas: (a) the simple empiricist view of science that tends to predominate; (b) the demarcation between scientific and non-scientific inquiry and (c) the avoidance of controversy—in part the consequence of the tendency toward textbook standardization. Most importantly, this study provides some evidence of the gulf that separates philosophy of science from science instruction, and examines some important aspects of the demarcation between science and non-science—an important issue for philosophers, scientists, and science educators.

doi:10.1093/bjps/axp011

Advance Access published on March 31, 2009

Is scientific method a part of science? Yes and no. Yes, if you hold that this method is a defining characteristic of science; but no, if you consider that this method is obviously not a part of physics, for example, in the way that thermodynamics is a part of physics.

The objects of study of a science include a variety of natural phenomena that scientists seek to explain. Scientific method is not one of these phenomena—unless you hold the view that the question of the nature of this method is best treated by cognitive science. Yet even this would not entail that scientific method is an object of science *per se*, and it would have to deal with the complication that cognitive science would then find itself in the peculiar position of being a discipline whose method is an object of itself.

Physicists study such things as quarks and quasars, but not the method that enables them to conduct this study. The fact that they skillfully employ such a method does not mean that they know at a deeper level what they're doing. While I hold that it is philosophers of science that are primarily responsible for having developed rich explanations of this methodology (with support from other investigators for whom the practice of science is an object of inquiry, such as cognitive scientists, sociologists, historians, and educators), this claim is not the point of this discussion. Rather, I want to investigate the nature and adequacy of the accounts of scientific method that we regularly find in basic textbooks in the natural and social sciences: what they say and don't say, what inconsistencies may exist among them, and what debatable positions they may espouse, knowingly or unknowingly.

Why is this important? First of all, such a study provides us at least some substantial data on the spectrum of treatments we find in these texts.

Second, philosophers of science can contribute some measure of critical thinking with regard to these treatments. Inconsistencies, confusions, and omissions can be identified. Methodological distinctions not apparent to more empirical investigators can be provided. This can aid textbook authors in the improvement of their presentations. Siegel ([1991]), for example, recommends that philosophy of science be utilized by science educators to focus more explicitly on critical thinking, that is, on 'reasons and evidence' in science (p. 53). Further, philosophers of science may in turn find in these texts certain concepts important for scientists, but neglected in philosophical analyses. Many of these texts, for example, highlight the intelligent revision of hypotheses, and therefore at least implicitly provide a more realistic picture of testing and response to testing than the more eliminativist approaches inspired by Popper.

Third, because the philosophy of science still takes scientific inquiry to be a complex phenomenon not completely understood or fully characterized, it can help science-text authors avoid communicating to their students the erroneous impression that the question of the nature of scientific method is settled and that little controversy remains.

Finally, such a study may encourage more communication among scientists, science educators, and philosophers regarding, for example, what demarcates scientific from non-scientific investigation and the possible commonalities of method shared by scientific and non-scientific disciplines.

1 Scientific Method in Science Textbooks

These accounts, found at the beginning of such textbooks, are designed to assist the beginning student in getting an idea of how scientific investigation may differ from inquiry in other areas. Rudolph ([2005]) has provided an excellent historical account of the origin of the step-by-step algorithm ('observe', 'hypothesize', 'test', and its many variants) we find in these descriptions, tracing the basic idea to John Dewey's *How We Think*, published in 1910. The irony is that Dewey presented this pattern as one that characterizes any careful, thoughtful inquiry (despite the fact that it was referred to as a 'scientific' method), while authors of science texts quickly appropriated it in a way that restricted it to the natural sciences.

These textbook accounts usually occupy no more than three or four pages and rarely indicate that there is any problem in characterizing this method. While these texts are intended as introductions to specific sciences (and not to the philosophy of science), the simplification we typically find in them can have an unfortunate effect. Scientific method, taken as a logical, epistemic, and cognitive process, is certainly at least as complex as, say, the theory of evolution. We do neither of these phenomena justice by failing to appreciate how puzzling they can be.

Philosophers are not the only critics of these textbook accounts. A number of educators specialized in science instruction have taken exception to their stilted and often confused structure and have advocated using the resources of the history and philosophy of science to improve them. Ennis ([1979]) provides a helpful summary, including an outline of how earlier authors (John Dewey, James B. Conant, Joseph Schwab, Max Black, Henry Margenau, Michael Martin, and Ernest Nagel) addressed the issue of defining a distinctive 'scientific method'. Hodson ([1991]) provides a more contemporary account, noting that science educators' conceptions of scientific method often lag decades behind contemporary philosophy of science (while mixing viewpoints from diverse perspectives). Matthews ([1994]) provides a very extensive discussion of the whole issue, including its historical background and contemporary criticisms, and makes a number of very concrete recommendations for improvement.

All these authors (as well as the many whom they cite) agree that philosophy of science should play a greater role in science education than it has. My analysis here, however, makes no specific recommendations for the improvement of pedagogy in this area. Rather, it provides two types of information that may aid these educators in this task: (a) data from 70 fairly current science textbooks, yielding a more-than-anecdotal profile of their treatments; and (b) a detailed philosophical criticism of these treatments with some recommendations for their improvement. My criticism is distinctive because of the approach I have taken in my own work. For example, I have argued that the 'method' employed in the natural sciences may also be found in systematic empirical inquiry in the humanities; and also that a very important component of this method is to be found in a 'logic of correction' that has been neglected by contemporary philosophy of science.

Martin ([1976]) ruefully noted that, in science textbooks, 'rather complex and difficult scientific theories are presented with some sophistication along with simple-minded views about the confirmation of these theories' (p. 296). Yet while Martin and Ennis ([1979], pp. 141–2), each give an example from a science text, these are used as singular cases rather than as parts of a more extensive database in which more general patterns might be detected. There is at least one other recent study that used such a database of science textbooks as a foundation for a similar analysis. Reiff *et al*. ([2002]) based their analysis on 40 science textbooks, and found that what they call a 'feedback' among all the various stages of this method was the rule rather than the exception. They, like other educators, are sensitive to the more naturalistic practices of inquirers, which never seem to mirror the linear step-by-step schemes that many textbooks offer. They did not, however, provide a very extensive range of issues beyond this in their analysis.

I too want to criticize these step-by-step schemes, but not because of their failure to correspond to real-life practice. *Any* formal or logical approach to an analysis of the methodology of inquiry will produce some sort of idealized scheme, with various distinguishable stages and steps. Even the simplest of logical operations will rarely appear in concrete reasoning in the form in which we find them represented in logic texts. If we are analyzing someone's extended prose argument, the logical form we may use to represent that argument will leave much of the concrete detail behind and adopt an 'unnatural' linearity. This is acceptable because we realize that an abstract logical form may have many alternative concrete embodiments, and that a formal 'train of thought' or 'line of argument' is not meant to represent concrete inquiry as such. So too, in scientific inquiry, there are many stops and starts, reversals, jumps, and simultaneous intuitions; but this does not mean that such reasoning must be tied to a purely naturalistic account and never distilled into a logical form. Such a form is not a 'real-time' description, but a more normative characterization.

This is where a philosopher's criticism may help balance that of an educator. An excessively simplified account of scientific reasoning is like a strawman argument, leaving itself open to easy criticism, both by friends and enemies: The friends may retreat from logical characterization into naturalistic description—a strategy adopted by many educators as well as philosophers of science encouraged by Thomas Kuhn's influence. The enemies, also in part utilizing Kuhn, may feel justified in smudging the differences between science and less rigorous forms of inquiry.

Naturalistic accounts of scientific inquiry are not substitutes for formal analyses. If scientists and science educators still want to maintain a real difference between the form of rigorous inquiry found in science and that found in speculations that are considerably less constrained, a logical/methodological analysis will be indispensable, even if deemed insufficient.

This perception is also important in maintaining a significant difference between the multiple 'local' methods that various individual sciences may employ and a more 'global' method common to all empirical science. I develop this point in *Of Two Minds* ([1998], pp. 150–1) in discussing Thomas Nickles' ([1987], [1990]) defense of local methods.

I also recognize, of course, that many philosophers have questioned even the existence of a unique 'scientific method', especially in the aftermath of Thomas Kuhn's analysis. We should remember, however, that even Kuhn distanced himself from the 'against method' movement he helped to stimulate and emphasized that he had not claimed that a formal account of method was no longer necessary, but only that it was insufficient. My own view, which I acknowledge is in the minority, is that it is indeed correct to say that there is no unique scientific method—not, however, because there is no helpful formal account of this method, or that this account would not apply to empirical science in general, but rather because this method can be found in any inventive area of inquiry, even in the humanities. This is the view I developed and defended in *Of Two Minds*.

What I provide here is a preliminary investigation of textbook treatments of scientific method—not a full, methodologically constrained empirical study. I have not quantified my results in a way that might reflect, for example, possible differences relating to year of publication, development across successive editions, inheritance and/or borrowings by textbook authors from other textbook authors, and level of education (pre-college, lower-level and upper-level college). And, while I have not attempted to quantify precisely the many differences in the way that various natural and social sciences treat scientific method, I will, in the conclusion of this study, provide some such data. I would also expect that the sample size of this study (70 texts) is sufficient to provide a more accurate empirical basis for our (sometimes erroneous) general impressions of how science texts treat scientific method, even if that sample may not be large enough to make more precise statistical judgments.

1.1 Textbook selection

I relied on students in my undergraduate philosophy of science classes, most of whom were natural science majors and resided in the Chicago area, to locate recent science textbooks both at the high school and college levels. This yielded a total of 70 different texts, as follows:

Number of texts	Area
21	Chemistry (including biochemistry)
16	Biology (including botany, anatomy/physiology)
9	Physics (including astronomy)
9	Psychology and behavioral science
6	Environmental science (including earth science and geology)
3	Social science
3	General science
2	Anthropology and archaeology
1	Nursing

Thirty-nine texts were published since 2000, 17 between 1995 and 1999, and 14 before 1995. One-sixth were high school texts; all the rest were undergraduate college texts. (In what follows, I will refer to individual texts as, e.g., ‘T38’; this corresponds to the numbered texts in the references).

1.2 Topic frequency

The first and most general observation to be made regarding which specific topics come under discussion in science textbooks is that these texts are for the most part content to provide only a very general outline of method, without getting into the special issues that make such a discussion methodologically interesting. Each of these texts invariably mentions ‘observation’, ‘hypothesis’, and ‘experiment’/‘testing’. About two-thirds move beyond this to more specialized topics such as inductive reasoning, falsification, theories versus laws, etc. Yet any given special topic was discussed in no more than a third of the texts (and usually less). What follows is a list of these topics (suggested in many cases by the texts themselves) in the order of frequency in which they were discussed. This frequency is given both in terms of the percentage of the total number of texts in which each topic is discussed and the absolute number of texts in which it’s discussed (in parentheses):

Topic	Frequency
Experimental controls	34% (24)
Scientific method in everyday life	33% (23)
Flow charts of steps of method	33% (23)
‘The’ scientific method	30% (21)
Scientific method as common sense ‘codified’	27% (19)
Falsification	24% (17)
Theories	23% (16)
Inductive generalization	21% (15)
Science versus non-science	21% (15)
Deduction (‘if–then’ reasoning)	20% (14)
Generation of hypotheses	19% (13)
Historical origins of science	16% (11)
Revising/correcting hypotheses	16% (11)
‘Broken lamp’ example (and the like)	14% (10)
‘Only a theory’	13% (9)
Models	11% (8)
Ad hoc/alternative hypotheses	10% (7)
Description versus explanation (hypotheses/laws)	10% (7)
‘Verification’	10% (7)
‘Hypothetico-deductive’	9% (6)
‘Confirmation’	6% (4)
Novel prediction	6% (4)
Theory-laden observations	6% (4)
Assumptions underlying induction	4% (3)
Simplicity of hypotheses	3% (2)

Because so many of these topics were discussed in so few texts, a much larger sample size would obviously be needed to establish the few patterns that seem to emerge.

My discussion of the content of these treatments is divided into three parts:

(a) the general, preliminary ideas used to demarcate scientific from non-scientific investigation, including distinctions derived from a consideration of the historical origins of science;
(b) the various methodological components that these texts identify as essential for understanding scientific inquiry, such as deduction, induction, hypotheses, theories, and laws; and
(c) the actual steps of inquiry that form what these texts take to be the 'scientific method'.

Part I: Preliminaries

2 Science versus Non-science

The great majority of texts emphasized the testability of hypotheses as a hallmark of a distinctively scientific method, but 15 of them also attempted to characterize non-scientific investigation (beyond simply saying that such investigation failed to be testable). The following four general categories are helpful in sorting out what characterizes a 'non-scientific' inquiry:

2.1 Subjective experience/bias

T53 (a college text in psychology), relying on M. H. Marx's ([1963]) taxonomy, characterizes non-scientific approaches as 'intuitive', 'biased, subjective', 'ambiguous', 'inaccurate, imprecise', 'not valid or reliable', 'uncritical, accepting' (p. 6). T36 claims that (in addition to theological assertions) moral judgments and 'judgements about what ideas of things are valuable, beautiful, or likable are not subject to falsification by hypothesis testing' (p. 5). The authors do not make clear, however, whether this limitation derives from such judgments' normative status (a status that some philosophers might claim can also be objective) or from their bias and subjectivity.

T60 observes that

> [...] science is unable to answer such philosophical questions as 'What is the purpose of life?' or such religious questions as 'What is the nature of the human spirit?' Though these questions are valid and have great importance to us, they rely on subjective personal experience and do not lead to testable hypotheses. (p. 10)

This text is not alone is using such questions as typical of philosophical or theological inquiry. Such characterizations leave us in the dark as to how the authors would treat such a philosophical question as 'What is science?' That is, there seem to be lingering stereotypes of philosophers and theologians (the two often enough lumped together) as dealing principally with the unmanageable 'big picture' questions. Philosophers (in this context) are rarely treated as epistemologists, for example, despite the fact that, sometimes in the same textbook, they reappear in discussions of reasoning, proof, falsifiability, and even Kuhnian 'paradigms' of scientific inquiry.

Theologians also lack any champions in these texts and are usually portrayed as inquirers into non-observable objects (as we shall see below). Yet a case might be made even for the scientific respectability of some theological analysis—in textual criticism, for example, or in systematic investigations into the variety of human experiences of the 'holy'. I found no acknowledgment of such methodology in any of the texts examined.

These texts are careful in maintaining a sharp methodological fence between science and religion, going so far as to affirm 'no conflict' between the two and granting the 'legitimacy' of faith (e.g., T24, p. 6). While many scientists embrace such a 'two-world' approach, it is possible that some significant number of these authors take this stand simply to avoid controversy.

Like T60, T59 provides some representative questions, such as 'Why do we exist, for what purpose?' and concurs that such questions derive from subjective experience 'as a consequence of all those variable factors shaping the consciousness of each individual' (p. 25). But T59 does not conclude that the non-scientific character of such inquiry derives from its subjectivity or bias *per se*. This leads us to a distinctive new criterion for non-scientific investigation.

2.2 Too many unmeasurable variables

Immediately following the comment on the 'variable factors' shaping subjective human consciousness, T59 makes the observation that 'because these factors can be infinitely variable, they do not readily lend themselves to scientific analysis' (*ibid.*). That is, it is not subjective bias as such that is the decisive factor in blocking scientific analysis, but the unwieldy number and perhaps unquantifiable nature of these variables.

The same point is made more explicitly by T26 in a section entitled 'The Limitations of Science'. It asks: 'But why can't the procedures of the scientist be applied to social, political, ethical, and economic problems? Why do scientists disagree over environmental, social, and political issues?' The answer is: 'Often the disagreement results from the inability to control **variables**' (p. 6). Accompanying this statement is a very curious graph, mapping 'numbers of variables' (measured in powers of ten) onto 'disciplines', running as follows: math, physics, chemistry, earth science, biology, social sciences (p. 7). At the math end of this range, we have fewer than ten variables, while in the social sciences, we have 10^8+ variables. This in an interesting approach to the too-many-variables problem, but the graph itself, which lends a scientific air to this idea, seems more a guess or metaphor than a pattern based on real data. Besides, such complexity may not be an obstacle in principle. ‘Complexity theory’ may be in its infancy, but it may prove to be one of the most important new frontiers of scientific investigation.

Subjective experience/bias and excessive/uncontrolled variability: These two characteristics are similar to some others mentioned in these texts. As we have seen, T53 stresses the merely ‘intuitive’ nature of non-scientific investigation, meaning, for example, that such inquiry relies on uncritical hunches and guessing, that is, on what is superficial ‘common knowledge’. T56 argues that what distinguishes science ‘is its insistence on rigorous methods to examine a problem’ (p. 16). T25 singles out, in addition to unbiased hypothesis-testing, the reliability that derives from reproducibility (p. 15)—a point also made by T26 (p. 5) and T42 (p. 3). T20 notes that ‘we cannot [...] repeat history under controlled conditions to determine whether historians’ theories are correct, because no two situations are ever the same’ (p. 4). T46 seems to combine the too-many-variables and subjectivist arguments, observing that ‘politics, so strongly linked to human behavior, economic conditions, international affairs and unforeseen events, cannot be studied by the scientific method’ (p. 5).

2.3 Non-phenomenal objects

It is not surprising that many texts also note that investigations into matters that involve no possible empirical experience cannot be scientific. T60 observes that ‘science has no way of verifying testimonies involving the supernatural. The term *supernatural* literally means “above nature.” Science works within nature, not above it’ (p. 10). T36 makes the same point about the statement ‘there is a God’ (p. 5). T53 (a psychology text) notes that a hypothesis is untestable if it appeals to ‘ideas or forces that are not recognized by science’ (rather circular reasoning). It continues:

> To suggest that John Hinckley shot President Reagan because he was under orders from the Devil is not testable because it invokes a principle (the Devil) that is not in the province of science. Such hypotheses might be of value to philosophers or theologians but not to the scientist. (p. 18)

They provide no examples of any philosophers or theologians who might today find the Devil-hypothesis ‘of value’. Contemporary psychologists may embrace a hard empiricism more completely than their colleagues in other sciences (given their special concern with human bias and their suspiciousness of theory), but it does not serve their purpose to defend the scientific character of their discipline by using contrasts that demonstrate their ignorance of the disciplines that they claim to be non-scientific. Fortunately, this is a rare case.

2.4 Falsifiability

The great majority of the 70 texts stress that genuinely scientific hypotheses must be falsifiable (although only 24% examine falsification as a logical principle–see Section 8). I will discuss this important concept and its limitations later, but there is one consequence of a strict Popperian interpretation of this idea that has made its way into a handful of science texts and is relevant here. T24, for example, claims that:

> [...] the hypothesis 'intelligent life exists on other planets somewhere in the universe' is not scientific. Although it can be proven correct by the verification of a single instance of intelligent life existing somewhere in the universe, there is no way to prove it wrong if no life is ever found. (p. 5)

The same argument appears in T36 relative to the hypothesis 'the sun will rise in the east tomorrow morning' (p. 5).

This appears to be an application of Popper's ([1934]) point that 'existential statements' (as opposed to 'universal statements') cannot be falsified (pp. 69–70). This is a symmetry that Popper insisted on. Universal statements cannot be verified because it would be impossible to examine all possible singular instances of the phenomenon described by these statements; and singular (existential) statements cannot be falsified for exactly the same reason.

Yet the science texts that use this example fail to note two very relevant points. First, Popper himself had to qualify his claim that existential statements were not falsifiable. He gives the example of the discovery of the element Hafnium, whose existence was predicted from the pattern discerned among known elements in the periodic table—an apparently existential claim. Popper observed that the element wasn't discovered until several of its properties were deduced from Bohr's theory: 'but Bohr's theory and [...] its conclusions [...] are far from being isolated purely existential statements' (p. 70). The same might be said for the prediction of extraterrestrial intelligent life: Scientists can make judgments about the frequency of earth-like planets, the likelihood of organic chemistry on such planets, and the processes of evolution that might lead to complex higher organisms. These judgments all rely on theories from which the 'existential' prediction would be made.

Second, although Popper insisted that falsifiability should be the defining characteristic of a genuinely scientific methodology, we can argue that, if he were to be faithful to his claims of symmetry, the verifiability of existential statements should be no less a valid component of method than the falsifiability of universal statements. I do not make that argument here (for I think it may be reasonable to restrict scientific claims to universal statements), but authors of science texts sometimes give the impression that the Popperian view is unchallengeable.

3 Scientific Method in Everyday Activities?

The authors of most science textbooks seem concerned with avoiding two extremes. One is the idea that there is no distinctive scientific method and that the claims of science should be treated with no more respect than other claims. The other extreme is the idea that scientific method is the esoteric possession of a group of isolated professionals. This second view is rejected by those texts that affirm that scientific method may be found even in the problem-solving we encounter in everyday activities. Yet if this method is indeed found in this more mundane context, should we really be calling it the 'scientific' method at all? Let's look at a few examples.

First of all, 19 texts hold that 'scientific method' is little more than 'common sense' codified or formalized.

T35 observes:

> Because many experiments discussed in this book were performed by people called scientists, you might think that science is a special process used only by certain people and useful only under special circumstances. That is not true at all. We all use the scientific method every day! (p. 11)

T5 notes that 'the scientific method is not confined to scientific laboratories. It can be used profitably anywhere a person wants to understand something that isn't immediately obvious' (p. 5). T59 points out that the processes of scientific inquiry 'are not difficult to learn. They are not professional secrets. They can be used in searching for answers to questions in everyday life as well as in methodical research' (p. 22). T65 observes that 'science is not the exclusive preserve of white-coated laboratory workers. The scientific method of experimentation and refinement of theories is practiced by all of us in an almost reflexive way' (p. 27). T27 suggests that the difference between the employment of this method in science and in non-scientific experience is only a matter of degree:

> You may appreciate the scientific method as a useful procedure in your everyday life. The primary difference between your everyday usage and that of a scientist in the laboratory is the care with which a scientist constructs experiments and the caution he or she uses in not jumping to conclusions. (p. 8)

Despite this generous view of method, I found no text that acknowledged the employment of the scientific method in the sorts of problem-solving one might find in humanistic disciplines or in legal inquiry, although T47 does at least strike an analogy with the process of legal investigation:

> As in a court of law, every effort is made to have [scientific] observers objectively examine the logic of both procedures and conclusions. Courts sometimes make wrong decisions; science, likewise, is not immune to error. (p. 20)

In contrast to the 'scientific-method-in-everyday-activity' approach, T53 claims that 'the scientific method is something abstract. It is an approach to knowledge that is best described by distinguishing it from what might be called nonscientific or everyday approaches to knowledge' (p. 5). Because this last text is in psychology, one can understand the reluctance of its authors to soften the line between 'intuitive' approaches ('folk psychology') and scientific methodology.

Many of these texts use concrete everyday examples of problem-solving to illustrate scientific ways of thinking. I will explore in some detail the cases of the 'broken lamp' and their cousins later, in the context of exploring the dynamics of method described by these texts.

4 When Did Science Begin?

While most texts discuss some important scientists of the past, only 11 specifically address the question of the historical origin of the scientific method, some more directly than others.

4.1 Greek science?

According to the majority of those texts that raise the question, Greek 'science' fails to meet the criteria of real science because it failed to experiment, that is, to subject its ideas to test. T5 singles out Aristotle, telling us that 'in ancient times, this is the way most thinkers behaved. If a statement sounded right they believed it without testing it' (p. 3). T24 notes that Aristotle's false mechanics 'was held to be true for nearly 2000 years because of Aristotle's compelling authority' (p. 3). (T26 mentions this 2000-year stretch as well, but neither text notes that Aristotle's and most other ancient Greek texts were lost for about half of this period and that Aristotle exercised little or no 'authority' until they were rediscovered after the late antique and early medieval dark ages.) T26 observes that Aristotle sought to 'understand nature through logic' and that 'experimental science did not triumph over Aristotelian logic until about C.E. 1500' (p. 3). (They do not mention Aristotle's biological works.) T46 tells us that Aristotle simply 'assumed' the four-element theory and that 'chemists of several centuries ago (more commonly referred to as alchemists) tried, in vain, to apply the four-element theory to turn lead into gold' (p. 5). (They do not mention Newton's extensive alchemical investigations).

An exception to this pattern is found in T36, which observes that 'the testing of hypotheses is an ancient discipline' and can be found in the writings of Herodotus (p. 17). This text also credits non-Western civilizations with developing 'sophisticated systems of medicine, metallurgy, and other technologies on the basis of a solid scientific foundation' (*ibid.*).

There is also the interesting objection that even if Greek 'scientists' did often enough base their conclusions on empirical fact, their method was *deductive*, using established principle (and established fact) to determine specific new facts, rather than generating new principles. Both T46 (p. 5) and T48 (p. 6) strike this contrast between non-scientific deduction and scientific induction. T48 uses the case of Eratosthenes' determination of the circumference of the earth to illustrate this deductive method, which 'is the reasoning of mathematics and philosophy and is used to test the validity of general ideas in all branches of knowledge' (*ibid.*).

Granting that Eratosthenes' geometrical constructions were simple, it still seems reasonable to include his empirical determination of this astronomical fact with that group of legitimate scientific inquiries that seek to determine specific physical values, such as fundamental physical constants. We encounter such investigations in applied science. Thomas Kuhn included many examples in his discussion of 'normal' scientific exploration, which he took to be the lion's share of what scientists do. Is Eratosthenes' 'deductive' determination of the circumference of the earth really different methodologically (factoring in the difference in sophistication) from what was accomplished in the Millikan oil-drop experiment? Aren't most if not all experiments of this type 'deductive', that is, seeking to determine specific values (via clever experimental designs) on the basis of already established principles?

4.2 Seventeenth-century origins

Despite the tendency of the majority of examined texts to find scientific problem-solving in everyday activities, most texts resist the idea that science existed before 1600 (e.g., T24; T26; T48). They associate its emergence more with controlled experimentation than with developments in theory. Bacon is mentioned frequently, for example, while Copernicus, Kepler, and Descartes hardly at all. Further, they fail to point out that these latter three, and Galileo as well, all singled out applied mathematical reasoning as the key to this new way of looking at the world.

T46 apparently solves this problem by viewing such theoreticians as inductivists, characterizing induction as 'observations leading to a general statement or natural law' (p. 5), and taking even Copernican theory as a generalization from observations—never examining Copernicus' own claim that his theory was a simpler explanation of the same facts integrated into Ptolemaic theory and that it was this simplicity that made his demonstration 'stronger'. T36 notes that historians often date the beginning of modern science to the founding of the Royal Society in 1660, which for the first time provided an organized scientific community (p. 17). From this standpoint, Copernicus, Kepler, and Galileo were not modern scientists.

Part II: Components

5 Formal Logic

In about a third of the texts, before the discussion moved to specific consideration of various elements of scientific method (the testing of hypotheses, etc.), there was some attention paid to forms of logical inference in general, that is, to the basic distinction between *deduction* and *induction*. What is somewhat surprising here is how consistent these discussions are with one another—suggesting a kind of 'textbook inheritance' in which these ideas are simply passed from one text to another, even across scientific disciplines.

5.1 Deduction: 'if–then reasoning'

Deductive reasoning is presented in two contexts. In the first, it is described simply as the form in which we reason from the general to the particular. In the second, this abstract form is embodied in the way in which scientists deduce predictions from their hypotheses.

Only a few texts discuss deduction apart from testing, that is, apart from 'hypothetico-deductive' reasoning. We have already seen how deduction is treated in the case of ancient Greek reasoning about nature. T19, however, makes the more unusual claim that 'laws are deduced from observation' (p. 2). Since this is a physics text, we might conclude that this way of speaking was a version of Newton's 'deduction from the phenomena' and should therefore not be taken literally.

Many texts speak of deduction as 'if-then' reasoning. Yet instead of clearly demarcating the hypothesis from the predictions deduced from the hypothesis, some texts take the 'hypothesis' to be the complete 'if–then' relation. T40 explains that 'hypotheses are normally expressed in the form of testable if–then statements: If some set of conditions is present and observed, then a certain kind of behavior will occur' (p. 33). That is, in these accounts, the prediction is taken as part of what we mean by the hypothesis. In this same text, we also have: 'the researcher forms a hypothesis, which is essentially a prediction [...] ' (*ibid.*).

5.2 Induction

Most of the texts that discuss deduction as a species of formal reasoning discuss induction in the same way. Most, such as T36, also strike this familiar contrast: '**Deduction** is reasoning from the general to the specific. [...] **[I]nduction** (or more properly 'inductive generalization') [is] reasoning from the specific to the general' (p. 6). (No text pointed out that deduction may also reason 'from the general to the general'). T56 carries the symmetry further, referring to the 'hypothetico-inductive approach' (p. 17) (the only instance of this term I found). Most texts point out that induction 'never guarantees the truth of any consequences drawn' (*ibid.*), or, as this text puts it, that 'we must be sensitive to exceptions and to the possibility that the conclusion is not valid' (*ibid.*), but I found no text that addresses the issue of how an induction can have, logically speaking, any warrant whatsoever. T65 observes that 'philosophers of science debate the relative merits of these approaches [deductive and inductive reasoning], but in reality they are well-integrated parts of the scientific process' (pp. 26–7). The point, however, is whether and to what extent these types of reasoning do in fact contribute to 'the scientific process'.

Only a very few texts stand out as refreshing exceptions, acknowledging that inductive methods in science must rely on certain basic assumptions, including the uniformity of nature. T47 notes that 'the basis for the scientific method is the *belief* that the universe is orderly and that by *objectively* analyzing phenomena, we can discover its workings' (p. 20); T30 observes that 'all science is based on the assumption that the natural world behaves in a consistent and predictable manner' (p. 16); T31 elaborates a little further:

> A discussion of the scientific process is not complete without an admission that scientists do have to make certain assumptions. They have to believe, for example, that nature is real and understandable and knowable by observing it; that nature is orderly and uniform; that measurements yield knowledge of the thing measured; and that natural laws are not affected by time. (p. I-12)

I found no discussion in these texts of the notorious weakness of inductive generalizations as *explanations*. Also, most texts apparently viewed induction as an inference that proceeds from observations directly to theoretical hypotheses, that is, without an antecedent grounding in explanatory theory. And yet, as we shall see, most texts do take hypotheses themselves to be inherently explanatory.

In general, therefore, discussions of deductive and inductive reasoning in these textbooks exhibit the same rather unsophisticated empiricism found in discussions of other issues, such as the demarcation between science and non-science and the historical origins of natural science.

6 Hypotheses, Theories, Laws, Models

6.1 Description and explanation

As one would expect, virtually every account of scientific method in these texts focuses on the testing of hypotheses, and many of these accounts include contrasts among hypotheses, theories, and laws as well. These contrasts are presented fairly uniformly. T28 is typical:

> After a number of experiments [that test a hypothesis], the scientist may be able to summarize the results in a **natural law**, which describes how nature behaves but does not explain why nature behaves in that particular way. [...] Finally, the scientist may be able to formulate a **theory**. The theory explains why nature behaves in the way described by the natural law. (p. 8)

Among the texts that adopt this 'law-as-descriptive' and 'theory-as-explanatory' division, there is some divergence: T7, for example, suggests that a well-supported hypothesis 'becomes a scientific theory or law [...] that explains the [observed] behavior [...]' (p. 16); and T38 observes that 'we do *experiments* that help *explain* behavior' (p. 12). But these are rare exceptions.

T57, reflecting the prevailing view, notes that *laws* are 'generalizations' (p. 33); but T36 speaks of *hypotheses* as 'generalizations' (p. 8), as does T27 (p. 7).

One might infer from most of these texts that theories are superior to laws in their explanatory function, while laws are superior to theories in their empirical warrant, that is, that laws are more confirmed than theories. This may explain the tendency in seven texts to take laws as 'higher' than theories. T64, for example, notes that 'occasionally, but hesitatingly, a theory advances to a law' (p. 10), and T62 affirms that, after 'a group of related theories' has been extensively confirmed, 'we generally elevate it to a new status. We call it a *law of nature*—an overarching statement of how the universe works' (p. 6). T56 equates a law with a 'principle' (p. 20). While it's possible to attribute the 'elevation' of a theory to a law to its greater empirical warrant, the examples provided of such laws include Newton's law of gravitation and the second law of thermodynamics, rather than less impressive empirical generalizations such as Coulomb's Law, which is hardly an 'overarching statement of how the universe works'. There seems to be a genuine ambiguity in the way this term is used in these texts.

6.2 Models

Eight of the texts explored the role of *models*, which typically were taken as simplified and more 'visualizable' hypotheses. Thus T26: '**Scientific models** use tangible items or pictures to represent invisible processes' (p. 5). A model is a 'simplified representation' (T39, p. 54) used 'to test a hypothesis' (T34, p. 15). T23 elaborates: 'Although a model may be a physical object, a **model** in science refers to any representation intended to convey information about another object or event' (p. 6). That is, a model strikes an analogy between disparate contexts.

T16 treats hypotheses and models as equivalent (p. 22), while T70 equates models with theories (p. 12). For T18, models—like hypotheses and theories—are essentially explanatory and may be mathematical in nature (p. 18).

Unlike the treatment of hypotheses and theories, most of these discussions of models are perfunctory—without, for example, the sorts of analyses one might find in philosophical and psychological inquiries into the status of models as analog forms of representation (such as in my [1997]; [1998], Chapters 6, 10, and 11).

6.3 'Only a theory'

Most of the texts that describe the nature of a *theory* are also careful to distinguish this sense of the term from the more common use: 'The statement, "Oh, that's just a theory," made in everyday conversation, implies a lack of knowledge and careful testing—the opposite of the scientific meaning of the term' (T34, p. 15). It is not surprising, given recent debate on the advisability of teaching intelligent design alongside evolution, that 'evolutionary theory' should be singled out for special defense. T48 observes: 'To suggest, as many critics outside of science do, that evolution is "just a theory" is misleading. The hypothesis that evolution has occurred is an accepted scientific fact; it is supported by overwhelming evidence' (p. 8). T64 asserts that 'the theory that life has evolved and is ever changing is founded on as much evidence as the broadly accepted generalization that the earth is round' (p. 10). (If only evolution were as transparent as the shape of the earth from an orbiting satellite).

6.4 Simplicity

While philosophers of science are uniformly committed to an empirical basis for any scientific method, they would also probably agree that such an account is insufficient. Most of these textbook treatments, however, give just this impression of sufficiency, for there is hardly any strong defense of the role of *theoretical* components in scientific reasoning. While hypotheses and theories, in contrast to laws, are viewed as explanatory and not simply as descriptive, there is little or no attempt to explain what 'explanatory' may mean. This might have been provided, at least in part, for example, by a discussion of *simplicity*.

I found only two texts that acknowledged this important characteristic of scientific explanation. T46 observes: 'When differing or conflicting theories are proposed, the one that is most successful in its predictions is generally chosen. Also, the theory that involves the smallest number of assumptions—the simplest theory—is preferred' (p. 5). Even better was T24: '[Scientists] change their minds, however, only when confronted with solid experimental evidence to the contrary, or when a conceptually simpler hypothesis forces them to a new point of view' (p. 3). This latter observation is taken from a physics text; perhaps physicists are more sensitive to the power of theoretical simplicity than their colleagues whose disciplines focus more predominantly on empirical investigation.

It is significant that T24 raised the question of simplicity in the context of considering multiple or alternative theories, for in such cases, it often happens that the facts are not sufficient to decide the question, and novel prediction may not be feasible, at least for a time. That is, a simple empiricist methodology may fall short. I will return to this issue in Part III.

These two texts were the exceptions. Most neglected simplicity entirely, even when the discussion called for it most. T56, for example, notes that scientists must 'establish priority among the hypotheses to decide which to test first'. This text offers the following three guidelines to assist in making this decision: A good hypothesis (a) 'is reasonably consistent with well-established facts'; (b) 'is capable of being tested [...]'; and (c) 'is falsifiable [...]' (p. 18). Yet even a non-simple patchwork hypothesis might meet these criteria. This would also have been an ideal place to discuss the role of *ad hoc* hypotheses in scientific explanation. I will discuss their use below, as we come to consider the more dynamic components of 'the scientific method'.

Part III: Dynamics

7 The Generation of Hypotheses

Philosophers of science (but not authors of science texts) have taken seriously the issue of demarcating the *context of discovery* from the *context of justification*. The former involves the procedures whereby hypotheses are first generated, whereas the latter is confined to those procedures by which these already-generated hypotheses are tested ('justified', 'confirmed', 'falsified', 'corroborated'). The neglect of this issue in science texts is puzzling, for one of the philosophers whose theory of science is most tied to this distinction is Karl Popper, who famously declared that there can be no logical method of coming up with new ideas (no 'logic of discovery'). It is puzzling because *falsifiability*—Popper's central idea (which is connected with his rejection of a logic of discovery)—has figured very prominently in these same texts.

Rather, what we find in these texts is an almost casual treatment of how it is we 'come up with' our initial hypotheses. Not a single text mentioned either the discovery/justification distinction or the problem of a logic of discovery.

These texts treat the generation of hypotheses in two ways (singly or in combination):

(a) Hypotheses are generated inductively (an idea that Popper rejected). Yet, there was little discussion in these texts of either the logical problem of induction *per se*, or of the failure of inductive generalizations as *explanations*.

(b) Hypotheses are generated 'creatively'. While this echoes Popper more closely, most texts were not clear whether such 'creative' or 'intuitive' invention is subject to logical analysis, or is extra-logical ('aesthetic') in some way. Many texts report that initial hypotheses emerge as 'guesses'; and two note that inductive generalization entails detecting 'patterns' and regularities (T46, p. 5; T62, p. 4), but without suggesting any logic to such detection. A number of other texts embrace 'creative intuition': T54 notes simply that hypotheses are 'derived from actual observation or from a "spark of intuition"' (p. 12). Besides using both deductive and inductive reasoning, T36 observes that hypotheses may also be generated in other ways, namely, by '(1) intuition or imagination, (2) aesthetic preferences, (3) religious and philosophical ideas, (4) comparison and analogy with other processes, and (5) serendipity [...]' (p. 6). That is, 'hypotheses are formed by all sorts of logical and extralogical processes' (*ibid.*). But this text does not comment on whether it matters, for example, that *only* extra-logical processes might be used. Similarly, T59 notes that 'there is considerable creativity within the boundaries of these investigative processes. Insights can result from accident, from sudden intuition, or from methodical research' (p. 21). T30 quotes with approval the sentiments of Rutherford and Ahlgren: 'Inventing hypotheses or theories to imagine how the world works and then figuring out how they can be put to the test of reality is as creative as writing poetry, composing music, or designing skyscrapers' (p. 17). It is interesting that whereas many texts embrace the idea that scientific inquiry involves creative elements, few if any take seriously the idea that areas of inquiry traditionally associated with creative imagination, such as the arts and other humanities, may adopt methods that are at least analogous to scientific thinking.

T25 notes that 'there is no single answer' to how a researcher generates a hypothesis, although this process favors the prepared mind, that is, one that is a keen observer that is already armed with a broader theory (p. 15). T19 observes that 'when new observations do require explanation and cannot be described within the conventional wisdom, it often requires the radical thinking of a genius to establish a new system of order' (pp. 2–3) (no indication is given whether such 'genius' can be logically analyzed). T56 seems to adopt the interesting proposition that it is the 'creative' component of inductive reasoning that may be responsible for its sometime logical unreliability: 'The generalizations in inductive conclusions come from the creative insight of the human mind, and creativity, however admirable, is not infallible' (p. 17).

It is not only authors of science texts but philosophers of science themselves who have for the most part abandoned any hope of developing a logic of discovery (of generating hypotheses). In contrast, I have argued ([1987], [1998]) that the logical generation of a new hypothesis may be taken as the intelligent correction of an antecedent hypothesis in response to the specific nature of the errors encountered by predictions from that hypothesis. I will discuss the applicability of this idea to the treatment of method in these science texts once we come to the final stages of the dynamics of this method—where responses to falsification are examined.

8 The Testing of Hypotheses

8.1 Proof/verification/confirmation

Most philosophers of science nowadays avoid the term 'verification'. The term 'confirmation' is preferred, not only because of its weaker sense of 'strengthening' the claim being tested, but also because even unequivocal empirical support for a hypothesis will not be decisive, given the possibility that alternative hypotheses may receive the same support. I will consider this 'alternative hypothesis' issue below, but let us first see how science texts describe the positive results of testing.

At the strong end of this terminology, a very few texts have hypotheses being 'proven' true. T6 speaks of a 'proven idea' as a *fact*; theories, however, are not proven (p. 6). T16 suggests we can either 'prove' or 'disprove' a hypothesis (p. 23). T45 notes that we can 'prove a hypothesis valid' (p. 8). Language such as this should probably not be taken strictly; 'proof' may mean no more than 'empirically supported', and a 'valid' hypothesis, despite the logician's fundamental distinction between truth and validity, may mean the same thing. And when other texts note that we cannot 'prove' but we can 'disprove' hypotheses (T29, p. 1.10; T52, p. 3), they really apparently mean 'verify' and 'falsify', respectively.

There were four texts that spoke exclusively of 'verification' (T34, p. 15; T61, p. 5; T63, pp. 14–5; T68, p. 15). Yet it is apparent that the use of 'verification' in these cases (like 'proof' above) does not really communicate anything stronger than what philosophers of science mean by 'confirmation'; in one text, what is 'verified' may even be 'overturned' (T34, p. 15). There is a single instance of using 'verification' to mean 'replication' of a positive result by other researchers (T45, pp. 8–9).

There were three texts that apparently equated 'verification' and 'confirmation' (T1, p. 10; T36, p. 5; T41, p. 3). Three texts (T7, p. 16; T31, p. I-12; T40, p. 33) avoided both 'verification' and 'confirmation', preferring, simply, 'support'. One text (T17, p. 5) spoke of a hypothesis that 'passes many tests'. I found only a single text (T2, p. 331) that spoke exclusively of 'confirmation' and 'disconfirmation'. In addition to the softer senses of 'verification' and 'confirmation' that we find in these texts, there are also many instances of texts affirming that, while the scientific method may aid in developing support for hypotheses, this support never rises to the level of certainty.

8.2 Why is confirmation inconclusive?

It is instructive to see how some of these texts characterize what it is that can go wrong with confirmation, and then consider why they don't have the same opinion of falsification.

8.2.1 Inductive generalization

While a number of texts, as we have seen, tend to see 'inductive' as simply 'empirical', many take seriously the traditional view of induction as logical generalization.

Confirmation seems at first to be dealt a death blow by the fact that it is practically impossible to examine every instance of a phenomenon whose nature we are trying to characterize by means of a hypothesis we are entertaining. But this outcome follows mainly from the idea that the instances we are examining are treated as independent of one another. They are treated this way because the generalization in question tends to take the connection between this set of instances and the dependent variables of the hypothesis more as correlational than as causal. *Explanations*, as distinct from generalizations, on the other hand, do focus on causal connections. As N. R. Hanson ([1958]) pointed out, we do not explain 'a bevelled mirror's showing a spectrum in the sunlight [...] by saying that all bevelled mirrors do this' (p. 70). Post-inductivist philosophy of science has appreciated this.

Once the causal connection is understood, the instances from which one has 'generalized' are no longer independent, but become tokens of a more general *type*. The hypothesis embodies an insight into a connection between an explaining principle and a *type* of phenomenon (composed of multiple instances, to be sure). It is this high-level connection, which may involve complex pattern recognition as well, that provides the basis for Norbert Wiener's famous low estimate of the number of instances needed for a full-blown generalization: 'Two instances would be nice, but one is enough'. That instances in an inductive generalization may be more than what is represented in 'singular statements' is a consideration that may derive from the same principle (examined earlier) Popper invoked to deny that seemingly singular statements were not falsifiable. That is, what may appear as isolated instances may really be bound together by theoretical principles by means of which these instances are treated *en masse*. While such reasoning still requires the assumption that we can treat different instances as being of the same type, this is a somewhat more modest version of the 'principle of uniformity of nature' than that required for inductive generalization as such.

8.2.2 Alternative hypotheses and the hypothetico-deductive method

Only three texts refer specifically to the *hypothetico-deductive* method (T10, p. 16; T56, p. 17; T64, p. 9), perhaps influenced by the authors' acquaintance with this idea in philosophy of science, but none discusses very extensively any of the logical problems that have beset this approach. None pointed out, for example, that if isolated from other considerations, the skeletal structure of hypothetico-deductive reasoning—where successful predictions deduced from a hypothesis (and initial conditions) are taken to 'support' the hypothesis—matches fairly well the logical fallacy of affirming the consequent. That is, successful predictions from a hypotheses will not especially serve to establish that hypothesis if there are other hypotheses available from which the same successful predictions can be made.

This is a missed opportunity for these science texts. The resemblance between the stereotypical characterization of hypothetico-deductive methodology and this logical fallacy is easy even for beginning science students to recognize. Highlighting this resemblance and the possible problems it poses for characterizing a 'scientific method' could prove a very fruitful way to encourage these students to develop a more sophisticated view of this method.

A few texts do explicitly note that confirmation cannot be conclusive if alternative hypotheses are available. T16 observes that 'the advocates of a new model may have gathered an impressive array of evidence for their hypothesis, but there may also be substantial evidence in support of another hypothesis' (p. 23). In such a situation, neither hypothesis is ruled out 'until the evidence unequivocally contradicts it' (*ibid.*). T20 insists that 'testing is crucial for weeding out unproductive hypotheses; without tests, rival theories may compete endlessly for acceptance' (p. 4). T48 similarly stresses elimination: 'A successful experiment is one in which one or more of the alternative hypotheses is demonstrated to be inconsistent with the results and is thus rejected' (p. 7). T3 points out that it is not enough to have confirming evidence: 'To be scientifically valid, [...] the experiment must also rule out other possible variables that might have caused the phenomenon' (p. 5). This and other texts often include a description of what constitutes a controlled experiment. I will return to this idea below.

Some texts also note that *novel prediction* is an important factor is accepting well-confirmed hypotheses. At a first glance, T30 seems to support this idea:

> It is not enough for scientific theories to fit only the data that are already at hand. Theories must also fit additional observations that were not used to formulate them in the first place. Put another way, theories should have predictive power. (p. 17)

As it stands, however, this description could be taken only as an endorsement of confirmation, since the 'additional observations' are apparently being used to confirm the hypothesis *initially*. (This text also speaks of observations that were 'used to formulate [the theories] in the first place'. This is not, needless to say, a component of the hypothetico-deductive view, which rejects the idea of a logic of 'formulating' hypotheses).

T41 provides clearer support for novel prediction, noting that ' [...] we use this theory to predict characteristics of materials we have not yet made and outcomes of experiments not yet performed' (p. 6). T63 observes that a theory 'not only helps organize current knowledge, but should also predict how new facts may fit into the established pattern' (p. 15). These two texts thus hint at a dual role for successful prediction: It functions to *confirm* a given hypothesis; but it is also *projective*, helping to anticipate the discovery of important new data. Yet the texts do not explicitly connect the logical function of the latter to the concept of confirmation itself.

These texts suggest that the problem of multiple hypotheses is resolved by eliminating (falsifying) all but one; and perhaps by invoking a projective use of novel prediction as well. In general, therefore, while they do notice that support for a hypothesis depends on the disposition of its rivals, this 'disposition' tends to be entirely empirical in nature. There is no indication that hypotheses may be eliminated for purely theoretical reasons—if they are burdened by *ad hoc* exceptions, for example, or if they lack other qualities of plausibility and simplicity.

These last texts tried to characterize the problem in *general* terms; a number of others preferred to pose the question in the context of specific, concrete examples. T21 considers the hypothesis that pigeons use the sun's position to navigate. This hypothesis appears to be supported by the observation that whereas all of the 25 pigeons released on a sunny day returned home, fewer than half returned on a cloudy day. But 'the conclusion may not be correct. Perhaps a storm [an alternate hypothesis] kept many of the pigeons from returning on a cloudy day' (p. 7). Similarly, T35 examines the hypothesis that leaves change their color in the fall because of lower temperature. Suppose a potted tree is placed in a growth chamber in July and, after a lowered temperature of the sort typical of the fall, the leaves do change their color. This would not mean, this text continues, that the hypothesis was confirmed, for the result may have been due to the artificial light of the growth chamber or the nature of the pot the tree was planted in (p. 8).

8.3 Disproof/falsification

In Part I of this discussion, I examined one part of the emphasis on falsifiability we find in these texts. We must now consider another feature of this idea that is usually embraced in these same texts: that whereas hypotheses may not be conclusively confirmed, they *can* be conclusively falsified. This is a principle associated with Popper's theory of inquiry.

Seventeen texts emphasize falsifiability (not just testability) as a hallmark of the scientific method. Eleven of these provide a formula similar to that found in T52: 'It is in the nature of physical theory that we can *disprove* a theory by finding a phenomenon that is inconsistent with it, but we can never prove that a theory is always correct' (p. 3). T24 cites Einstein to the same effect (p. 4), while T2 invokes Richard Feynman (p. 331).

There are some interesting variations in the ways these texts interpret this principle: T56 notes that 'in theory (but not necessarily in practice) a well-stated hypothesis can be proven false' (p. 16) (it is not made clear exactly what would 'in practice' hinder falsification). T63 observes that, 'if a hypothesis cannot be falsified, it is accepted as a **theory**' (p. 15). Perhaps the authors meant '*has not* been falsified' and misspoke.

While the asymmetry between 'can't prove' but 'can disprove' is traceable to treating hypotheses as inductive generalizations (as T59 suggests; p. 22), the majority of those texts that admit the inconclusiveness of confirmation do so mostly because of the difficulty of excluding alternative hypotheses, not because of the logical problem of induction. Non-inductive methodologies (such as hypothetico-deductive approaches) are equally concerned about the alternative hypothesis problem.

8.4 Why is falsification inconclusive?

There is little recognition in these texts that falsification can always be circumvented and that 'disproof' is as elusive as 'proof' (a point closely connected to the Duhem–Quine thesis). This is a serious methodological defect in these treatments.

There are two principal ways in which falsification may be avoided:

8.4.1 Saving a hypothesis through *ad hoc* exceptions

If an observed phenomenon seems to falsify an established hypothesis, a scientist always has the liberty of saving the theory by adding an *ad hoc* supplementary hypothesis to account for the (apparently falsifying) result.

Thus, while a portion of the precession of Mercury's orbit is today taken as an important empirical confirmation of Einstein's General Theory of Relativity (and simultaneously a falsification of Newton's gravitational theory), it was not only logically permissible for Newtonians in the eighteenth century to save their theory by invoking another factor to explain the precession (an undetected planet between Mercury and the Sun, for example), but scientifically wise as well—even though it turned out to be false. Logically, conclusive falsification would require that we isolate the main hypothesis from all of the possible *ad hoc* exceptions that might save it; that is, we would have to confirm that 'all things are equal' and that no undetected factor is responsible for the negative result. Such a global confirmation would prove impossible to effect. Yet this is the aim of controlled experiments where, *as far as they are able*, scientists seek to eliminate such interference.

Many scientists worry about abuse of this logical escape from the threat of falsification. Too often, theories are protected because their proponents will not bend even in the face of the facts. Avoiding falsification in this way would destroy the important demarcation between science and non-science that is the very point of these accounts of the 'scientific method'.

Yet the use of *ad hoc* exceptions for the purpose of saving hypotheses is still regulated by *scientific* constraints—even if these constraints are not empirical *per se*. Too many tacked-on exceptions (each an independent explanation) will inevitably overburden a hypothesis and make it suspect. Such hypotheses could be discarded *prior* to testing for theoretical reasons: They will have become far-fetched, lacking simplicity and plausibility. It is true that this criterion is less clear-cut and harder to characterize than empirical adequacy, and it is perhaps the reluctance of scientists to acknowledge such non-empirical, theoretical components of scientific method that leads them to overestimate the decisiveness of falsification. While they might try to avoid this question by tying their method to novel prediction—comparing the future fruitfulness of the saved hypothesis (and its *ad hoc* rider) with that of a rival hypothesis (without such an exception)—this would not in itself be a reason *not to try* saving the former hypothesis in this way.

A number of texts do note what it is that may 'go wrong' with falsification, but none of them addresses this formal point: that it is always logically possible and sometimes scientifically advisable to resist falsification. This is another missed opportunity for these textbook authors, for even beginning students would be able to appreciate this simple and important aspect of scientific inquiry.

How do these texts characterize a failed falsification? Here again, we have a variety of accounts. T62 simply notes that 'tests do not necessarily [...] disprove a theory' (p. 5). T31 observes that a scientist might 'retest' a hypothesis that has already been falsified, but without saying why (p. I-12). Seven texts provide specific circumstances in which an apparent falsification may need to be reviewed, but none of these rises above the level of concrete examples to a more reflective recognition of principle. T30, for example, notes that a negative experimental result may derive from errors in the experimental process (p. 16). T52 uses a classic historical example: While a feather and a cannonball do not fall at the same rate, this result does not falsify Galileo's law; the law is saved by attributing the negative result to air resistance (p. 3). T58 examines a case where a falsifying instance is simply excluded from the range of the hypothesis. The hypothesis is that all lemons are sour. On encountering the non-sour Meyer lemon, the hypothesis is modified to 'any lemon *except* the Meyer lemon' will be sour (p. 12); but there is no discussion of the theoretical suitability of such an artificial exclusion.

There is one more important point to consider. As we have just seen, even falsification may not be conclusive, if one is able to appeal to *ad hoc* exceptions. In fact, much of what it is that can go wrong with confirmation derives from the same principle underlying what it is that can go wrong with falsification. Each form of testing ends with a result, positive in one case and negative in the other. Neither of these results would be conclusive, however, if one can find some other reason for their occurrence. In the case of confirmation, this other reason takes the form of an alternative hypothesis; in the case of falsification, the other reason takes the form of an *ad hoc* exception. The former is incompatible with the main hypothesis being tested, while the latter is quite compatible, and is added to the main hypothesis. Put another way, an inconclusive confirmation rests on a *disjunction* (posing an alternative to the main hypothesis), while an inconclusive falsification rests on a *conjunction* (adding an *ad hoc* rider to the main hypothesis). None of the examined texts notices this logical symmetry between the two forms of testing.

8.4.2 Revising/correcting a hypothesis

Another way of saving a hypothesis subsequent to an apparent falsification is by *revising* or *correcting* it in the light of the errors encountered— thereby retaining at least a version of the original. This is an idea that is acknowledged today more widely by science textbook authors than by philosophers of science. But these authors do not examine the methodological problem that underlies the related problem of a 'logic of discovery'.

I will not review in detail why it is that, among philosophers of science, both those that espouse a hypothetico-deductive approach and those that favor a more Popperian approach have tended to support the idea that, as Popper has said, ‘there is no logical method for coming up with new ideas’, that is, no logic for generating new hypotheses. For Popper, this rule is a simple application of the Darwinian principle. Novelties arise from blind (unplanned, non-rational) processes; unfit variants are eliminated by a hostile environment, leaving a ‘match’ between fit variants and the environment—a match that might erroneously suggest just such a planned design. Hypotheses arise blindly (through ‘creative intuition’ or the proverbial ‘monkeys at typewriters’, but with many pre-established constraints as well); ‘unfit’ hypotheses are eliminated/falsified by the empirical data to which they are unable to adapt. ‘Fit’ hypotheses are not ‘true’, but only relatively ‘strong’ or ‘well corroborated’; they may be eliminated tomorrow. A hypothesis can no more be ‘verified’ than an animal species can be rendered immortal. (This, incidentally, would prove an interesting way for authors of science texts to present the productive yet inconclusive nature of scientific inquiry—while providing a little lesson on Darwin as well).

My own work ([1987], [1989], [1995], [1996], [1998]) has concentrated on this problem. I have argued that a logic of discovery (that is, of the generation of new hypotheses) is possible. I hold that the rational discovery of a new hypothesis can be taken as the intelligent correction of an antecedent hypothesis in response to the precise nature of the discrepancies encountered between the predictions generated from that antecedent hypothesis and the facts. Even Popper violated his Darwinian principle by allowing intelligent correction into his accounts of ‘trial and error’. A true trial-and-error process would be purely eliminative; errors could not be used to intelligently improve (‘correct’) subsequent trials, as I have argued elsewhere ([1995]; [1998], Chapter 8).

The problem for authors of science texts is that, while many of them recognize the importance and relative prevalence of ‘revision’ and ‘correction’ of hypotheses, they are not able to characterize this process with any logical or methodological precision. In fact, none of the examined texts scrutinizes revision/correction very closely at all. Eleven of them simply refer to the fact that hypotheses may be ‘modified’, ‘revised’, or ‘refined’ subsequent to an unsuccessful attempt at confirmation. Twenty-three other texts, however, embed this idea in a flow chart.

These flow charts accomplish two purposes.

First, they provide an easy-to-read presentation of the sequence of steps that comprise the ‘scientific method’. Sometimes this is done very simply, with only two or three steps; in other cases, it becomes considerably more complex, with 10 or more steps or branches. Providing these steps in this form also lends to the discussion a logical air, encouraging students to appreciate the fact that it is indeed a ‘method’ that is being discussed.

Second, for almost all of these 23 texts, flow charts allow the authors to represent in a graphic way a ‘feedback’ relation (rendered as arrows that return from the testing stage back to previous stages of hypothesis formation).

It is in these feedback loops that revision and correction of hypotheses is represented.

T10 provides a good example: The flow chart proceeds from 'Observations' to a 'Question' to a 'Hypothesis' to a 'Prediction' to a 'Test' (p. 17). It then branches, depending on whether the results of the test were positive or negative: If positive, the flow feeds back to 'Prediction', where additional evidence is sought to support the hypothesis; if negative, the flow feeds back to 'Hypothesis', where the old hypothesis is either revised or rejected entirely.

Four of these 23 texts provide a very simplified version in which the flow is represented as a circle, and the method as a cycle; that is, feedback is almost the sole feature of the chart. T65, for example, provides a circuit that includes four components: developing hypotheses, testing them, analyzing the results, and reformulating the theories with which the initial hypotheses were consistent, all in 'continuous feedback' (p. 26).

Interestingly, there were four texts with flow charts but with no feedback in the flow at all, and four others with feedback, but no revision/correction. T34 has a peculiar example of the former, for the last step is 'Revise hypothesis if necessary', yet no arrow feeds back from this to the earlier 'Formulate hypothesis' step (p. 15). T48 has the most consistent of these no-feedback flow charts (p. 7). It provides a quite elaborate representation where multiple hypotheses undergo successive elimination processes. There is no feedback here because the authors view method basically as a pure process of elimination (Popper is acknowledged), with no revision or correction stage. (Nor do they discuss the legitimate use of *ad hoc* exceptions in protecting a hypothesis from falsification).

As for the four texts that have feedback components but do not mention revision/correction: Some of these refer somewhat nebulously to 'evaluation' of tests or 'drawing conclusions', but without specifying what this entails, while others apparently see the process simply returning to the beginning following testing—again, without discussing the possibility of revision/correction. T31 mentions 'retesting' or testing a 'related hypothesis' (p. I–12).

As I mentioned at the beginning of this discussion, Reiff *et al.* ([2001]) also paid special attention to the feedback loops that often appear in these flow chart schemes. They hold that 'these loops represent some effort by the textbook authors to be less linear (stepwise) in their presentation, a key criticism in the literature' (p. 4). They also assert, however, that such feedback is much more prevalent in actual scientific inquiry than is depicted in these diagrams and that these simplified models 'perpetuate the misconceptions that scientific achievements occur through following a predetermined path [...]' (p. 6). The gist of this criticism focuses on the failure of such idealized schemata to represent the richness of inquiry in a natural setting where, indeed, a 'predetermined path' is hardly to be seen. But I view such a feedback loop from a more methodological perspective, signifying (if the revision is significant enough) the generation of a new hypothesis and the basis for a 'logic of discovery'. Many textbooks that have simple linear patterns of the sort criticized by Reiff *et al.* in fact acknowledge that actual scientific inquiry does not proceed in such a 'predetermined' manner. This is not a contradiction, but only a recognition of the difference between a formal characterization and actual practice—a difference that is evident in all methodological analyses, even in purely deductive reasoning.

9 Experimental Controls and the 'Broken Lamp'

Science texts are at their best when they provide concrete examples of controlled experiments, for it is in this context that scientists are most at home. My concern, however, is with the methodological adequacy of these accounts—whether they succeed in identifying what may be unique in scientific inquiry and distinguishing it from what they take to be non-scientific.

Twenty-four of the examined texts discuss some of the details of experimental controls. Eighteen specifically describe what it is that constitutes such a design, while six others only give examples. Eleven of the texts provide examples drawn from the specific sciences with which they are concerned, while 10 texts provide examples drawn from 'everyday life'.

A number of the everyday examples involve some sort of failed lighting or 'broken lamp': T65 looks at a broken desk lamp (p. 27); T15 (p. 12), T34 (p. 16), and T48 (p. 7) examine house lights, bedroom lights, and inadequate room light, respectively; T66 considers a flashlight (p. 5). T3 (p. 5) and T42 (p. 5) both look at a car that won't start, T15 examines a squeaky door (p. 13), T20 a broken coffee pot (p. 5), T43 a broken calculator (p. 6), T66 a broken VCR (p. 5) (in addition to a flashlight), and T54 discusses a music system with noisy audio (p. 13).

What is evident from these accounts (including those that use specifically scientific examples) is that these controls function to exclude as far as possible any extra factors that might be blamed or credited for a given result: blamed if the result is an apparent falsification and the extra factor accounts for that negative result; and credited if the result is an apparent confirmation and the extra factor accounts for the positive result in place of the hypothesis under test.

Unfortunately, however, these controls are described in a way that has the confirmation of a hypothesis follow merely on the elimination of its rivals, with no element of revision/correction. T3 provides an excellent example. You start with the *observation*: Your car will not start; you formulate the *hypothesis* that the battery is dead; you *experiment* by replacing the battery and find that the car will start:

> But Wait. You haven't provided controls for several variables. Perhaps your battery was fine all along, and you just needed to try to start your car again. Or perhaps the battery cable was loose and simply needed to be tightened. Realizing the need for a good *control*, you replace your old battery, making sure the cables are secured tightly, and attempt to restart the car. If your car repeatedly refuses to start with the old battery but starts immediately when you put in your roommate's new battery, you have isolated a single *variable*, the battery. And [...] you can now safely draw the *conclusion* that your old battery is dead. (p. 5)

This text concludes by noting that 'scientists cannot be sure that they have controlled *all* the variables other than the one they are trying to study. Therefore, scientific conclusions remain tentative and are subject to revision [...]' (*ibid.*). This acknowledges that alternative hypotheses may not all have been eliminated, and that therefore the confirmation is provisional. The same could be said of an apparent falsification: We can't be sure that experimental controls have eliminated all possible *ad hoc* exceptions, and therefore cannot conclusively falsify the hypothesis.

Consider, however, the limited logic of these examples. If a flashlight fails to light, we judge that the cause was a bad bulb, a bad battery, or a bad switch (or some combination). That is, this *disjunction* is a *necessary* condition for the failure of the light. Put another way, the *conjunction* of a good bulb, battery, and switch is a *sufficient* condition for light.

But there is an additional requirement for this type of problem: If we replace the bulb, and the light comes on, we make the judgment that the battery and switch are also working properly. That is, we would not in addition test the battery and switch once the light comes on. The reason is that a conjunction of these three conditions (good bulb, battery, and switch) is also taken as a *necessary* condition of light (if the light works, then all three components are working). Thus, the conjunction is taken as both a necessary and sufficient condition for light.

Notice that, without a conjunction of *multiple* conditions, the 'testing' would have the appearance of affirming the consequent: We observe the absence of light. We replace the bulb and find that this succeeds in producing the light. That is, the reasoning seems to be: If I put in a good bulb, then the light will be produced; the light is produced; therefore it was the bulb. What of the 'alternate hypotheses' (the battery and the switch)? These are not alternatives because we've already established that all three conditions are necessary for the light. When the light appears, we automatically know the battery and switch are working. By examining a conjunction of causes (rather than a single cause) as a necessary and sufficient condition for a given effect, therefore, these everyday examples appear to involve multiple alternative hypotheses, but can avoid the appearance of affirming the consequent that besets a simpler form of confirmation.

The example of the car that won't start also points to the consequence of failing to distinguish all the possible causes beforehand. If the problem were really a loose cable, we would only erroneously conclude it was the battery if we did not at first realize that replacing the battery *ipso facto* affected the cable. That is, we would not have initially distinguished and isolated all the possible causes from one another.

The more unfortunate side of such examples has to do with the fact that, once we have determined these necessary and sufficient conditions, our testing consists in a quite simple process of elimination until we come to a confirmation: I try replacing the bulb first; if that doesn't work, I replace the battery; if that doesn't work, I examine the switch. In no case does the failure of a trial provide any positive information as to which of the remaining alternatives is more likely.

I am not claiming that processes of elimination do not occur or aren't important for understanding the logic of scientific inquiry, but I do hold that the intelligence with which scientists pursue their testing has a great deal to do with the fact that a negative result can be positively informative. When Kepler tested his hypothesis that the orbit of Mars was circular and then encountered a significant error, he didn't simply abandon this 'falsified' hypothesis; he observed the nature, degree, and direction of the error (the true orbit should be narrower at the quadrants) and then intelligently used this information to correct the hypothesis. Such a correction, I have argued elsewhere ([1987], [1998]), requires a mapping of sets of hypotheses onto sets of predictions and positioning the data among these possible predictions.

It is therefore misleading when some science text authors use the term 'revise' when they mean a process that is non-corrective and eliminative. In its everyday example of fixing a squeaky doorknob, T15 notes that '[...] once you know that the doorknob did not cause the squeak, you can revise your hypothesis to see if oiling the door hinges stops the noise' (p. 14). This is not corrective 'revision' but elimination of one hypothesis and proceeding to the next without new guidance. T66, in listing the possible causes for a VCR producing a poor quality picture, describes as 'corrective action' the 'experiments' designed to eliminate one-by-one the possible causes for this effect (p. 5). Elimination is not correction.

10 Conclusion

10.1 Different sciences, different concerns

Although it is not an easy matter to quantify the differences between specifically different sciences in their approach to an explanation of scientific method (for many aspects of these treatments conceptually overlap and are difficult to isolate), there are a few important areas where they are at least noticeable.

I want to stress again, however, that because these areas were discussed in only a small minority of the 70 texts examined, a much larger sample size would be needed to establish the few patterns that seem to emerge.

In the table that follows, I single out a number of cases where textbooks within a given science discuss a specific issue either significantly more or less than its proportion of the total number of texts. If, for example, chemistry texts account for 30% of the texts that discuss a specific topic, and also account for 30% of the total of 70 texts, then chemistry has an 'occurrence ratio' of 1.0 for this topic. A ratio above 1.0 indicates a given science's positive preference for this topic relative to other sciences, and a ratio below 1.0 indicates a science's negative preference.

Issue (#texts addressing this issue)	Occurrence Ratio
Models (8)	Chemistry 1.67
'Only a theory' (9)	Biology 2.92
'Hypothetico-deductive' (3)	Biology 4.35
Deduction as 'if-then' reasoning (10)	Biology 2.17
Historical origins of science (7)	Biology 1.87
" "	Chemistry 1.43
No 'science' before 1600 (7)	Biology 1.87
" "	Chemistry 1.43
Science in everyday activity (5)	Biology 1.74
" "	Chemistry 1.33
" "	Physics 0.00
Science *not* in everyday activity (1)	Psychology 7.70
Simplicity as a criterion (2)	Physics 3.85
Induction relying on assumptions (3)	Biology 0.00
" "	Chemistry 0.00

Eight texts specifically discuss the role of *models* in scientific method. Half of these were chemistry texts. For this topic, therefore, chemistry has an occurrence ratio of 1.67 (i.e., 50%/30%). If models are discussed in the context of visualizable configurations of some kind (as they usually are), then one might surmise that chemistry may favor including this idea in its discussion of method because molecular structures can be readily represented in this way.

While 16 texts discussed the scientific meaning of the term 'theory', nine of them also warned about speaking of a scientific theory as 'only' a theory. Two-thirds of these were biology texts (a 2.92 ratio), perhaps reflecting the special concern biologists have in protecting the 'theory' of evolution. It may also be for this reason that biology is more defensive with regard to the integrity of the 'scientific method' generally: As can be seen from the table, biology texts also feature (more than other sciences) discussions of hypothetico-deductive (and deductive) reasoning as well as arguments why science, properly speaking, didn't exist before 1600. Chemistry texts share some of this latter concern: They, perhaps, want to be sure to demarcate 'chemical science' from alchemical pre-science.

How, then, are we to take the seemingly contrary tendency in biology and chemistry texts to find the scientific method in 'everyday' activity? I suspect that at least part of the motivation here is to represent scientific thinking as the culmination of all that is best in thinking itself; in doing so, they sharpen the demarcation between 'good thinking' and the 'wishful thinking' associated with more speculative disciplines (the humanities, of course, usually representing this type). But a more pedagogical concern may also be operating here, reflecting the desire of biology and chemistry text writers to make their disciplines approachable for beginning students. Although the sample size is obviously too small to support a general conclusion, it is also interesting that no physics or psychology text makes so strong a point: In fact, it is only a psychology text that goes out of its way to *deny* that everyday thinking can be scientific.

It is, as I shall explain more completely below, the simple empiricism that drives biology and chemistry texts that may prevent them from considering more seriously the importance of simplicity in determining the acceptability of explanatory hypotheses and that leads them to neglect assumptions such as the uniformity of nature that lie behind their inductivism.

I have not in this review discussed all of the special issues regarding scientific method found in these texts. Many of these issues are specific to the science in question: for example, measurement, statistical design, the role of replicability/reproducibility, the socio-political context of research or the treatment of various experimental cases.

I have concentrated, rather, on the more general methodological issues common to a broad spectrum of sciences and of particular interest to philosophers of science. Although my evaluation and the summary I offer below is mainly critical, I hasten to add that many of these accounts are surprisingly detailed, and some quite extensive; further, contrary to my initial expectations, quite a number of them refrained from presenting scientific method as foolproof. In fact, many hold up the provisional nature of scientific claims as a hallmark of scientific inquiry (despite their frequent reliance on falsifiability and 'disproof'). Yet it is evident that many science text authors are torn: If the criteria for conclusive demonstration (either confirmation or falsification) are softened, what is to prevent the barbarians from storming the walls of science and erasing any real difference between its methodology and unconstrained speculation?

There are three basic areas into which it is possible to pull together the various evaluations and criticisms I have offered of these textbook accounts: (a) the simple empiricist view of science that tends to dominate these descriptions;

(b) the demarcation question, that is, how it is that scientific inquiry differs from inquiry in other areas; and (c) the avoidance of controversy.

10.2 Simple empiricism

At one time, philosophers of science embraced more thoroughly inductivist views of scientific inference and reasoning than they do now—a change that followed many decades of investigation and repeated unsuccessful attempts to make such an inductivism work. Subsequently, both hypothetico-deductive approaches and Karl Popper's novel approach tried to preserve a reworked conception of an empirical foundation of science, but without trying to account for the generation or confirmation of hypotheses through inductive inference.

Yet we still find in many basic science textbooks an uneasy and awkward marriage of older inductivist views with these newer non-inductivist approaches. This sometimes results in inconsistent characterizations of basic ideas, such as the inconclusiveness of confirmation (is it mainly the result of the defect in inductive generalization? Or is it the result of the alternative-hypothesis problem, which *per se* has nothing to do with induction?). Further, the 'problem of induction' that led philosophers of science to consider these new approaches was not based simply on the fact that induction remained inconclusive; rather, it was difficult to see how inductive methods could, logically speaking, provide any warrant at all for a given claim, and beyond this, how generalizations could ever function as adequate *explanations*. Most authors of science textbooks give no evidence that they are aware of this development, even if they sometimes invoke Popperian and hypothetico-deductive methodologies—often offering them side-by-side with inductivist views that these methodologies were designed to replace.

While these later non-inductivist developments remain empiricist in nature, a lesson learned from the failure of inductivism is that empirical warrant, while a necessary condition for scientific advance, is not sufficient. That is, these newer empiricist views have a more complete account of the role of non-empirical (theoretical) constraints on scientific method.

We find very little awareness in these texts of the possibility that theoretical simplicity and plausibility (components of good explanations) may function as co-criteria (along with empirical adequacy) of hypothesis evaluation and theory choice. This is apparent not only in those texts that portray even theory-heavy scientific breakthroughs as empirical or even inductive (Copernican theory, for example), but also in those that describe how scientists treat the alternative hypothesis problem or the use of *ad hoc* hypotheses in resisting falsification. The solution always seems to be: Eliminate these extra hypotheses through further empirical testing—with little recognition of the purely theoretical grounds there may be for disposing of them before testing.

This simple empiricism is also apparently the main cause of the devotion to falsifiability as an essential condition for scientific inquiry. The commitment to the 'can't prove but can disprove' principle seems based, first of all, on an initially inductivist conception of confirmation. Yet the very meager treatment afforded the use of *ad hoc* hypotheses (as a means of circumventing falsification) reveals, I believe, a fear that letting go of falsifiability even to the smallest degree may affect the demarcation between science and non-science. It is in this context that we can also appreciate important differences in the ways that textbooks in different sciences characterize scientific method in general—as I have outlined in the preceding section.

10.3 The demarcation question

Maintaining a strict demarcation between science and non-science is probably the principal motivation for espousing the simple empiricism just described. Unfortunately, this aim has led a number of authors of these texts to stereotype the research activities we find in non-scientific areas. These stereotypes work to undermine this aim, for professional historians, philosophers, and even theologians and artists would not recognize themselves or their work in most of the descriptions we finds in these texts. It would not take much to assist scientists in developing a more sophisticated view of the methodologies employed in these non-scientific areas. Better communication, obviously, would be needed to do so.

One also senses some resistance in the way these texts treat the historical beginnings of science—resistance to the possibility that Hellenistic technology and inquiry, for example, could be 'scientific' (the treatment of Eratosthenes), and resistance to more filled-out characterizations of some classic early figures (Newton's alchemy and religious fundamentalism; Kepler's sometimes mystical motivations in his astronomy; or even Copernicus' reliance on theoretical simplicity or Galileo's minimizing of the importance of confirmational testing).

We also find an unresolved ambiguity in the way many texts treat 'everyday' science. Either everyday problem-solving embodies the scientific method or it does not. If it does, then the sorts of problems that historians struggle with—or philosophers or poets—should also qualify. My own preference is to affirm that 'scientific method' is not the exclusive property of the natural sciences and therefore is not distinctively 'scientific'. Rather, it is shared by all those who push at the limits of existing categories in a systematic and empirical way. That its use in science is more precise and hence more successful might be explained in part by the quantifiability of its subject matter. Quantitative precision, however, is not a sufficient demarcation for a qualitatively different (or superior) method of thinking. Naturally, such a claim of a 'common' method must be explained and defended. I address this issue squarely in my other work.

In not extending the problem-solving methods of 'everyday' science to non-scientific disciplines, many authors of science texts give the impression that researchers in these disciplines are victims of subjective bias, with little check on their speculations. These authors need to discuss these matters more with their colleagues in other disciplines in order to discover what these checks are (including empirical checks) and how constraining they can be.

Finally, if scientists and science educators are still serious about demarcation (and the recent turmoil triggered by the 'intelligent design' debate indicates that this concern has never been greater), then they will need the assistance of philosophers who can provide methodological analyses beyond naturalistic descriptions of problem-solving.

10.4 Textbook standardization and the avoidance of controversy

It is the fact that these accounts of scientific method form part of the introduction to *textbooks* that may be the cause of much of what is inadequate in them. Textbooks are not the best place to look for controversies in scientific research. Consequently, there may be a tendency on the part of these authors to provide a relatively quick account of scientific method that is as standardized as the accounts provided, say, of optics, or photosynthesis, or basic mechanics that follow in the main body of the text.

The problem is that scientific inquiry is in its own way a less understood phenomenon than those examined in these settled studies. Even 'textbook' introductions to the philosophy of science intended for philosophy students must be careful not to play favorites or to gloss over still controversial views. Minority positions (regarding the possibility of a logic of discovery, for example) should also have at least some representation in such discussions.

This naturally produces some impatience among scientists, for whom the existence of this method is, for the most part, unquestioned. T65 observed that 'philosophers of science debate the relative merits of these approaches [deductive and inductive reasoning], but in reality they are well-integrated parts of the scientific process' (pp. 26–7). This sounds as if this 'debate' is gratuitous or irrelevant, when in reality what is being debated is whether in fact these *are* 'well-integrated parts of the scientific process'. I found only four texts that mentioned specific philosophies of science: T36 has David Hull and Thomas Kuhn (pp. 14, 19); T48 has Popper (p. 8); T65 has Karl Hempel and Popper (pp. 26–8); and T68 mentions logical positivism (p. 16).

There is also no doubt some tendency to view philosophical analyses as untrustworthy because of the unsettled nature of continuous philosophical 'debate'. Yet there are plenty of areas even within science whose progress has been slow because of the nature of the subject matter: the functioning of the human brain, for example, or the nature of animal behavior, or hominid paleontology. One or two science texts did note that the obstacle to the scientific study of such phenomena was not the 'subjectivity' of the study, but the great number of variables that had to be dealt with. Yet this may not be an obstacle in principle, for even complex organization may yield to analysis in time—perhaps by means of principles that are as basic as, yet fundamentally different from, the foundational laws of physics (see Steven Wolfram [2002], for example, and Steven Weinberg's [2002] review of Wolfram's work).

Scientists rightfully value what is nowadays referred to as 'critical thinking' skills. These skills are gradually developed by students of a particular science as they make their way to more challenging areas of their subject matter. Philosophers of science are trained specifically to 'think critically' about the nature of science. Students in this discipline too must make their way into more and more challenging areas, until they reach the frontier, where many issues remain unresolved. It would not be an unhappy result if science students fascinated by puzzling natural phenomena were to come to appreciate how puzzling the phenomenon of science itself is.

Department of Philosophy
Loyola University Chicago
6525 N. Sheridan Road, Chicago, IL 60626
jblacho@luc.edu

References

Blachowicz, J. [1987]: 'Discovery as Correction', *Synthese*, **71**, pp. 235–321.

Blachowicz, J. [1989]: 'Discovery and Ampliative Inference', *Philosophy of Science*, **56**, pp. 438–62.

Blachowicz, J. [1995]: 'Elimination, Correction and Popper's Evolutionary Epistemology', *International Studies in the Philosophy of Science*, **9**, pp. 5–17.

Blachowicz, J. [1996]: 'Ampliative Abduction', *International Studies in the Philosophy of Science*, **10**, pp. 141–57.

Blachowicz, J. [1997]: 'Analog Representation beyond Mental Imagery', *Journal of Philosophy*, **XCIV**, pp. 55–84.

Blachowicz, J. [1998]: *Of Two Minds: The Nature of Inquiry*, New York: State University of New York Press.

Ennis, R. [1979]: 'Research in Philosophy of Science Bearing on Science Education', in P. Asquith and H. Kyburg (*eds*), *Current Research in Philosophy of Science*, East Lansing, MI: Philosophy of Science Association, pp. 138–70.

Hanson, N. R. [1958]: *Patterns of Discovery*, Cambridge, MA: Cambridge University Press.

Hodson, D. [1991]: 'Philosophy of Science and Science Education', in M. Matthews (*ed.*), *History, Philosophy, and Science Teaching*, Toronto, ON: OISE Press, pp. 19–32.

Martin, M. [1976]: 'The Relevance of Philosophy of Science for Science Education', in R. S. Cohen *et al.* (*eds*), *PSA 1974*, Dordrecht: D. Reidel, pp. 293–300.

Marx, M. H. [1963]: 'The General Nature of Theory Construction', in M. H. Marx (*ed.*), *Theory in Contemporary Psychology*, New York: Macmillan, pp. 4–46.

Matthews, M. [1994]: *Science Teaching: The Role of History and Philosophy of Science*, New York: Routledge.

Nickles, T. [1987]: 'From Natural Philosophy to Metaphilosophy of Science', in R. Kargon and P. Achinstein (*eds*), *Kelvin's Baltimore Lectures and Modern Theoretical Physics*, Cambridge, MA: MIT Press, pp. 507–41.

Nickles, T. [1990]: 'Discovery Logics', *Philosophica*, **45**, pp. 7–32.

Popper, K. [1959]: *The Logic of Scientific Discovery*, English edition translated by K. Popper, New York: Basic Books.

Reiff, R., Harwood, W. S. and Phillipson, T. [2001]: 'A Scientific Method Based upon Research Scientists' Conceptions of Scientific Inquiry', (retrieved online March 6, 2007 from Education Resources Information Center, Eric No. ED 465618).

Rudolph, J. [2005]: 'Epistemology for the Masses: The Origins of the "Scientific Method" in American Schools', *History of Education Quarterly*, **45**, pp. 341–76.

Siegel, H. [1991]: 'The Rationality of Science, Critical Thinking, and Science Education', in M. Matthews (*ed.*), *History, Philosophy, and Science Teaching*, Toronto, ON: OISE Press, pp. 45–62.

Weinberg, S. [2002]: 'Is the Universe a Computer?', *The New York Review of Books*, October 24, 2002, pp. 43–8.

Wolfram, S. [2002]: *A New Kind of Science*, Champaign, IL: Wolfram Media.

Science Textbooks

(1) Arny, T. T. [2002]: *Explorations: An Introduction to Astronomy*, 3rd edition, New York: McGraw-Hill.

(2) Aronson, E. [2004]: *The Social Animal*, 9th edition, New York: Worth Publishers.

(3) Audesirk, T., Audersik, G. and Byers, B. [2003]: *Biology: Life on Earth*, 7th edition, Upper Saddle River, NJ: Pearson Education.

(4) Becker, W., Kleinsmith, L. and Hardin, J. [2003]: *The World of the Cell*, San Francisco, CA: Benjamin Cummings.

(5) Bettleheim, F. A. and March, J. [1988]: *Introduction to General, Organic and Biochemistry*, 2nd edition, New York: Saunders Publishing Co.

(6) Bledsoe, L. J. [1994]: *Biology*, 2nd edition, Paramus, NJ: Globe Fearon Educational Pub.

(7) Bloomfield, M. M. [1992]: *Chemistry and the Living Organism*, 5th edition, New York: John Wiley.

(8) Brady, J. E. and Humiston, G. E. [1986]: *General Chemistry: Principles and Structure*, 4th edition, New York: John Wiley.

(9) Brown, T., Lemay, E., Jr., Bursten, B. E. and Burdge, J. R. [2003]: *Chemistry: The Central Science*, 9th edition, Upper Saddle, NJ: Prentice Hall Pearson Education.

(10) Campbell, N. A. and Reece, J. B. [2002]: *Biology*, 6th edition, San Francisco, CA: Benjamin Cummings.

(11) Chang, R. [1988]: *Chemistry*, 3rd edition, New York: Random House.

(12) Cozby, P. C. [2004]: *Methods in Behavioral Research*, 8th edition, New York: McGraw-Hill.

(13) Davis, R. E., Metcalfe, H. C., Williams, J. E. and Castka, J. F. [1999]: *Modern Chemistry*, Austin, TX: Holt, Rinehart and Winston.

(14) Davis, S. F. and Palladino, J. J. [2000]: *Psychology*, 3rd edition, Upper Saddle River, NJ: Prentice Hall.

(15) Dobson, K., Holman, J. and Roberts, M. [2001]: *Science Spectrum*, Austin, TX: Holt, Reinhart and Winston.

(16) Dolgoff, A. and Falcon, M. [1998]: *Essentials of Physical Geology*, Boston, MA: Houghton Mifflin.

(17) Ebbing, D. D., and Gammon, S. D. [1999]: *General Chemistry*, Boston, MA: Houghton Mifflin.

(18) Feldcamp, S. [2002]: *Modern Biology*, Austin, TX: Holt, Rinehart and Winston.

(19) Grant, I. S. and Phillips, W. R. [2001]: *The Elements of Physics*, New York: Oxford University Press.

(20) Griffith, W. T. [1998]: *The Physics of Everyday Phenomena*, 2nd edition, Boston, MA: WCB/McGraw-Hill.

(21) Heimler, C. H. and Price, J. [1987]: *Focus on Physical Science*, Columbus, OH: Merrill Publishing.

(22) Hein, M. and Arena, S. [2000]: *Foundations of College Chemistry*, 10th edition, Pacific Grove, CA: Brooks/Cole Publishing Co.

(23) Herron, J. D., Sarquis, J. L. and Schrader, C. L. [1996]: *Health Chemistry*, Lexington, MA: D. C. Heath & Co.

(24) Hewitt, P. G. [1997]: *Conceptual Physics*, 3rd edition, Menlo Park, CA: Addison-Wesley.

(25) Hilgard, E., Atkinson, R. L., Smith, E. E., Bem, D. J. and Noel-Hoeksema, S. [2000]: *Hilgard's Introduction to Psychology*, 13th edition, Orlando, FL: Harcourt College Publishers.

(26) Hill, J. W. and Kolb, D. K. [2001]: *Chemistry for Changing Times*, 9th edition, Upper Saddle River, NJ: Prentice Hall.

(27) Kroschwitz, J., Winokur, M., and Lees, A. B. [1995]: *Chemistry*, 3rd edition, Dubuque, IA: Wm. C. Brown, Publishers.

(28) Lemay, H. E., Robblee, K. M., Beall, H., and Brower, D. C. [2002]: *Chemistry: Connections to Our Changing World*, Upper Saddle River, NJ: Prentice Hall.

(29) Lewis, R., Gaffin, D., Hoefnagels, M., and Parker, B. [2004]: *Life*, New York: McGraw-Hill.

(30) Lutgens, F., and Tarbuck, E. [2001]: *The Atmosphere*, 8th edition, Upper Saddle River, NJ: Prentice Hall.

(31) Mader, S. [2001]: *Biology*, 7th edition, New York: McGraw-Hill.

(32) McGuigan, F. J. [1990]: *Experimental Psychology*, 5th edition, Englewood Cliffs, NJ: Prentice Hall.

(33) McMurry, J. and Fay, R. [2001]: *Chemistry*, 3rd edition, Upper Saddle River, NJ: Prentice Hall.

(34) Miller, G. T., Jr. [2004]: *Sustaining the Earth: An Integrated Approach*, 6th edition, Pacific Grove, CA: Brooks/Cole.

(35) Miller, K. and Levine, J. [1998]: *Biology: The Living Science*, Upper Saddle, NJ: Prentice Hall.

(36) Minkoff, E. C. and Baker, P. J. [2001]: *Biology Today: An Issues Approach*, New York: Garland Publishing.

(37) Muller, W. H. [1979]: *Botany: A Functional Approach*, New York: Macmillan.

(38) Myers, D. G. [2002]: *Exploring Psychology*, New York: Worth Publishers.

(39) Myers, R. T. and Oldham, K. B. [2000]: *Chemistry: Visualizing Matter*, Austin, TX: Holt, Rinehart and Winston.

(40) Nairne, J. S. [2003]: *Psychology: The Adaptive Mind*, 3rd edition, Belmont, CA: Wadsworth; Thomson Learning.

(41) Nebel, B. J. [1987]: *Environmental Science: The Way the World Works*, 2nd edition, Englewood Cliffs, NJ: Prentice Hall.

(42) Ostdiek, V. J. and Bord, D. J. [2000]: *Inquiry into Physics*, 4th edition, Pacific Grove, CA: Brooks/Cole Thomson Learning.

(43) Ouellette, R. J. [1988]: *Introduction to General, Organic, and Biological Chemistry*, New York: Macmillan Publishing Co.

(44) Perry, A. G. and Potter, P. A. [2001]: *Fundamentals of Nursing*, St. Louis, MO: Mosby.

(45) Perry, J. A. and Perry, E. K. [1997]: *Contemporary Society: An Introduction to Social Science*, Menlo Park, CA: Addison-Wesley Longman.

(46) Petrucci, R. H. and Harwood, W. S. [1997]: *General Chemistry Principles and Modern Applications*, 7th edition, Upper Saddle River, NJ: Prentice-Hall.

(47) Plummer, C. and Mcgeary, D. [1996]: *Physical Geology*, Dubuque, IA: Wm. C. Brown.

(48) Raven, P. H. and Johnson, G. B. [2002]: *Biology*, 6th edition, New York: McGraw-Hill.

(49) Russo, S. and Silver, M. [2000]: *Introductory Chemistry: A Conceptual Focus*, San Francisco, CA: Addison-Wesley Longman.

(50) Santrock, J. W. [2003]: *Adolescence*, 9th edition, New York: McGraw-Hill.

(51) Schad, J. [1996]: *Physical Science: A Unified Approach*, Pacific Grove, CA: Brooks/Cole.

(52) Sears, F. W., Zemansky, M. and Young, H. [1991]: *College Physics*, 7th edition, Reading, MA: Addison-Wesley.

(53) Shaughnessy, J. J. and Zechmeister, E. B. [1985]: *Research Methods in Psychology*, New York: Alfred A. Knopf.

(54) Silverberg, M. S. [2000]: *Chemistry: The Molecular Nature of Matter and Change*, New York: McGraw-Hill.

(55) Smith, R. A. and Davis, S. F. [1997]: *The Psychologist as Detective*, Upper Saddle River, NJ: Prentice Hall.

(56) Solomon, E. P., Berg, L. R. and Martin, D. W. [2002]: *Biology*, 6th edition, Florence, KY: Brooks/Cole Thomson Learning.

(57) Spaulding, N. and Namowitz, S. N. [2003]: *Earth Science*, Evanston, IL: McDougall Littell.

(58) Starr, C. [2003]: *Biology*: *Concepts and Applications*, Belmont, CA: Brooks/Cole Thomson Learning.

(59) Starr, C. and Taggart, R. [1984]: *Biology: The Unity and Diversity of Life*, 3rd edition, Belmont, CA: Wadsworth.

(60) Suchocki, J. [2001]: *Conceptual Chemistry*, Menlo Park, CA: Addison-Wesley.

(61) Thibodeau, G. A. and Patton, K. T. [1999]: *Anthony's Textbook of Anatomy and Physiology*, 17th edition, St. Louis. MO: Mosby.

(62) Trefil, J. and Hazen, R. M. [1997]: *The Sciences: An Integrated Approach*, 2nd edition, New York: John Wiley.

(63) Turnbaugh, W. A., Jurmain, R., Kilgore, L. and Nelson, H. [2002]: *Understanding Physical Anthropology and Archaeology*, 8th edition, Belmont, CA: Wadsworth Group.

(64) Volpe, E. P. and Rosenbaum, P. A. [2000]: *Understanding Evolution*, New York: McGraw-Hill.

(65) Webster, D. L., Evans, S. and Sanders, W. [1993]: *Out of the Past: An Introduction to Archaeology*, Mountain View, CA: Mayfield Publishing Co.

(66) Wilbraham, A. C., Staley, D. D. and Matta, M. S. [1997]: *Chemistry*, 4th edition, Menlo Park, CA: Addison-Wesley.

(67) Wysession, M., Frank, D. and Yancopoulos, S. [2004]: *Physical Science: Concepts in Action*, Upper Saddle River, NJ: Prentice Hall.

(68) Yegidis, B. L. and Weinbach, R. W. [2002]: *Research Methods for Social Workers*, 4th edition, Boston, MA: Allyn and Bacon.

(69) Zumdahl, S. S. [1997]: *Chemistry*, 4th edition, Boston, MA: Houghton Mifflin.

(70) Zumdahl, S. S., Zumdahl, S. L. and Decoste, D. J. [2002]: *The World of Chemistry*, Boston, MA: Houghton Mifflin.