Science & Education **12:** 315–329, 2003.

Lawson's Shoehorn, or Should the Philosophy of Science Be Rated 'X'?

DOUGLAS ALLCHIN
Minnesota Center for the Philosophy of Science and Program in the History of Science and Technology, University of Minnesota, Minneapolis MN 55455, USA
E-mail: allchin@pclink.com

Abstract. Lawson's (2002) interpretations of Galileo's discovery of the moons of Jupiter and other cases exhibit several historical errors, addressed here both specifically and generally. They illustrate how philosophical preconceptions can distort history and thus lessons about the nature of science.

1. Introduction

In his Pulitzer-Prize-winning book *Wonderful Life*, Stephen Jay Gould (1989) coined the expression 'Walcott's shoehorn' to label a significant bias in the work of paleontologist Charles Walcott (pp. 108, 244). Early last century, Walcott had characterized and classified organisms from the Burgess Shale, a trove of fossils of some of the earliest life on Earth. The Burgess organisms (by today's reckoning) reflect a striking diversity of life forms at a very fundamental level. But Walcott had been extraordinarily conservative. He classified them all within known groups (phyla), representing organisms today. The effect, Gould notes, was a misleading impression of progress, of an increasing diversity within established groups (the cone of diversity), rather than of the 'diversification and decimation' of many types of early life forms (p. 46). The apparently modest but pervasive flaw thus had substantial consequences. For Gould, shoehorning data into preconceived categories where they did not fit was a 'cardinal error' (p. 244) in science. Preconceptions were allowed to speak louder than the facts.

Here, I would like to discuss a parallel error in science education: shoehorning history into a particular view of scientific methodology. In this case, preconceptions of philosophy of science are allowed to speak louder than historical facts. The error, as a general type, is fundamental – and not uncommon. Yet its implications in science education are also substantial. Understanding the potential error is thus critical for anyone intending to teach 'nature of science' effectively. Here, I examine several cases, showing how a series of errors reflects a general syndrome (§§2–3). I then comment on the general error type and a prospective teaching strategy for addressing it (§4).

2. Did Galileo Discover Jupiter's Moons by the Hypothetico-Deductive Method?

Earlier in the pages of this journal, Lawson (2002) claimed that 'many, if not all, scientific discoveries are hypothetico-deductive in nature' (p. 21).[1] This included 'the insightful use of hypothetico- deductive reasoning employed by Galileo' in discovering the moons of Jupiter (p. 19, also pp. 9, 20), presented in some detail. These *historical* claims were offered to support strong *educational* recommendations about how to 'help students learn how to do science' (p. 2, also p. 21). Of course, if the history is flawed, then the validity of the educational claims may be, as well. In public comments following criticisms of this paper at the 6th International History, Philosophy and Science Teaching Conference (Denver, Nov., 2001), the author invited others (perhaps more expert than he) to correct any possible misinterpretations of history he might have made. I take up that challenge here.

Lawson's analysis of Galileo's celebrated discovery draws exclusively on his published account in *Sidereal Messenger*. Lawson casts it as a 'report' which 'chronologically reveals many of the steps in his discovery process' (p. 1). Direct reference to historical texts is, of course, usually preferred over reliance on secondary sources. But here, Galileo's discovery narrative is taken at face value. The published text is conflated with an actual account of the history. Nobel Prize-winner Peter Medawar (1964) famously profiled how modern scientific papers misrepresent scientific thought. They are persuasive texts, not narratives. They follow numerous rhetorical conventions (e.g., Bazerman 1988; Booth 1993). Just so for earlier texts in science. Galileo's rhetorical skills, in particular, are widely documented (e.g., Moss 1993; Biagioli 1993; Finnochiaro 1997; Shea 1998). Hence, one would be ill advised at the outset to regard Galileo's published "testimony" as accurately recounting his experience or original thinking. Rather, one should view it as an artful construction designed to persuade. For example, a hypothetico-deductive framework is strongly persuasive (retrospectively), even when it has not served for generating the discovery (prospectively) (Knorr-Cetina 1984, pp. 94–135). Galileo seems to have used a tactic that he used widely elsewhere. That is, he dramatized and amplified his claims by first considering other contrasting explanations and showing how each was not only plausible, but also ostensibly supported by evidence. Only then did Galileo introduce contradictory evidence, developing a strong sense of irony that supported his own claim rhetorically (Biagioli 1993, Chap. 3). His account of investigating the moons of Jupiter in *Sidereal Messenger* follows this pattern well. Thus, an analysis of Galileo's original thinking process based uncritically on Galileo's published text is suspect.

Second, Lawson's account of the discovery of Jupiter's moons discounts the role of the very instrument that enabled Galileo to observe them. Galileo certainly paraded the foundation of his revolutionary observations, his newly fashioned 'perspective eyepiece', later dubbed the telescope:

At length, by sparing neither labor or expense, I succeeded in constructing for myself an instrument so superior that objects seen through it appear magnified nearly a thousand times. (cited in Lawson 2002, p. 2)

I had prepared for myself a very excellent instrument. (p.2)

I noticed a circumstance which I had never been able to notice before, owing to want of power in my other telescope. (p. 5)

Without his new telescope, Galileo would not have discovered the moons of Jupiter. It also allowed others importantly to confirm his discovery once announced. Yet in Lawson's account, Galileo's development of the critical scientific instrument is peripheral. It was merely (passively) 'at his disposal' (p. 2). Not long ago philosophers of science regarded elements of experimental practice as peripheral to or irrelevant to discovery proper. However, recent studies have profiled how they are constitutive of the process (Hacking 1984; Franklin 1986, 1991; Galison 1987). Moreover, there is substantial epistemic work in learning how to use such new instruments effectively and in validating observations using them (Hacking 1984; Rothbart & Slayden 1994; Franklin 1997). As Galileo himself noted, constructing his instrument involved much labor and expense. Science is not practiced in a world of ideas only, but also relies on material resources and human time and effort. Galileo's case exemplifies how technological development factors into discovery (Pitt 2000, pp. 92–96). The lessons are essential if teachers want to convey the 'nature of science' authentically. Certainly, from a common sense perspective, one tends to credit the development of the telescope more than Galileo's thinking process as critical to his scientific discovery on this occasion. But if one already equates 'how to *do* science' with just 'how to *think* scientifically' (pp. 2, emphasis added; 4–9), one may be blinkered to the non-intellectual dimensions of history. One will miss significant elements of what Galileo's discovery of Jupiter's moons can tell us about the process of scientific discovery.

Third, Lawson focuses exclusively on reasoning and hypothetico-deductive elements. The resulting sense of 'discovery' and/or of science is very limited and counterintuitive. For example, the process apparently began only after Galileo had turned his telescope to Jupiter and only after he had seen the new 'stars' in its vicinity (pp. 5, 17). Oddly, *observing the night sky* with the new telescope for potentially new phenomena (without a hypothesis) does not count as part of the discovery process or formalized science. Nor does Galileo *noticing* the new celestial bodies as significant or worth further consideration. (One might well argue that when these were complete, the bulk of the discovery was past, not that it was just beginning.) Further, the process apparently ended every time Galileo reached a conclusion about each separate, successive 'hypothesis' explaining the new 'stars' (p. 17, Table II, Step #9). *Revising concepts* based on results or *generating new hypotheses* are also excluded. If the remaining events mark a 'scientific method', then the process is highly abbreviated and intermittent and does not describe fully Galileo's relevant behavior. Moreover, note that Galileo did not use hypothetico-deductive reasoning to *predict* Jupiter's moons. Nor did Galileo seem to use deductive reasoning to *generate his hypotheses*. In what sense, then, did the hypothetico-deductive method guide Galileo in making a *discovery* in any meaningful sense? The emphasis in this account is on successive testing, or the context of justification. There is no context of discovery. The process of science depicted thus omits much from both historical and philosophical perspectives.

The eclipse of hypothesis generation as part of Galileo's science is especially significant. Lawson does include speculation on how Galileo's ideas originated. In this account, the alternative hypotheses are based on background knowledge. They are applied through analogical reasoning or analogical transfer (pp. 5–6). Here, analogy is an 'if-then' form of inference based on similarity, or non-enumerative induction. Yet this reasoning pattern, ostensibly critical in the history, is not considered in the later analysis of scientific method (pp. 15-21).[2] There, all methods give way to hypothetico- deductive reasoning. Now, according to that method (as defined there) the would-be investigator is instructed to 'search the literature and your own knowledge base for as many possible answers as possible' (p. 17, Table II, Step #3). So analogy would seem to have an implicit role. But that role is suppressed. Hypotheses are portrayed as resulting from 'search', not creative imagination. There is a conundrum, of course. If Galileo's use of analogy was not creative, then his "discovery" introduced no novel idea. It only extended what was already known previously. On the other hand, if the analogy was creative, then it was relevant to the process of science, which cannot be limited to hypothetico-deductive reasoning. Ultimately, in Lawson's account, whatever process Galileo might have used to generate his concepts or ideas, it seems not to contribute to the process *scientifically*.

Lawson's account is, by his own admission, *entirely speculative*[3] Sometimes Lawson's claims are hedged, suggesting that he is only reconstructing Galileo's imagined thinking process *hypothetically*. Lawson introduces it as 'an attempt to fill in the gaps with what Galileo *may* have been thinking' (p. 2, emphasis added, see also pp. 4, 5, 9). 'Of course we can not know if this is what Galileo was really thinking' (p. 7). But the initial tentativeness wanes as the paper progresses (pp. 8, 9, 12, 15, 20, 21). No new historical evidence is introduced. Eventually, one finds such unqualified and unwarranted references as 'the insightful use of hypothetico-deductive reasoning employed by Galileo' (p. 19) and similar assertions in the abstract (p. 1). *Nothing substantive historically* is established from *imagining* that Galileo used the hypothetico-deductive method, *speculating* that he *could* have, or even reconstructing how he *might* have. Historical documentation matters. This is especially true if one treats the historical case as a 'planned test' of the metascientific hypothesis that 'many, if not all, scientific discoveries are hypothetico-deductive in nature' (p. 21). One needs conclusive historical evidence, not merely *possible* scenarios (and especially not ones based on shaky assumptions). Elsewise, preconceptions reign freely.

Challenges confront anyone trying to interpret Galileo's methodology, as illustrated in the history of such efforts. For example, Nicholas Jardine (1994), reviewing several books by prominent scholars, noted how Galileo seemed adaptable to each. For pragmatist Joseph Pitt, Galileo relied essentially on common sense experience. For the flamboyant Mario Biagioli, Galileo was a savvy politician and showman. For William Wallace, O.P., Galileo exhibited Aristotelian method, notably of a Thomist flavor. For rhetorician Jean Dietz Moss, Galileo was an expert in persuasion. To these four, Jardine adds 'Mach's Machian phenomemalist, Koyré's Koyrean metaphysician, and Feyerabend's Feyerabendian anarchist' (p. 280). All these interpretations seem to support Alistair Crombie's observation that 'philosophers looking for a historical precedent for some interpretation or reform of science which they are themselves advocating have all, however much they may have differed from one another, been able to find in Galileo their heart's desire' (p. 280). Galileo, it seems, exemplifies many methods. He is a Protean scientific hero. Ultimately, some histories tell us more about the writer than about what happened in the past (White 1987; Sapp 1990). One may thus approach Lawson's analysis, with its emphasis on hypothetico-deductive reasoning, with appropriate caution.

Multiple perspectives can enrich history. Each may highlight a different dimension of a complex phenomenon. Problems arise, however, if someone presents partial or incomplete evidence from one perspective as exhaustive. The eclipse of other perspectives in such cases is illusory. Indeed, one benefit of contrasting perspectives is that each may help hold the others accountable to a range of demonstrable facts. Disregarding alternative perspectives can thus lead to error.

Lawson's title question, 'What does Galileo's discovery of Jupiter's moons tell us about the process of scientific discovery?', ostensibly suggests an open inquiry into history. The impression is that history yields its answer transparently. Yet Lawson's history is selective and incomplete. Relevant facts are missing. Assumptions, and the limits they imply, are ultimately suppressed. A perspective seems to masquerade as fact. Indeed, one might contend that the historical facts are irrelevant to Lawson's conclusions (pp. 9–21). But the history is not idle. It functions rhetorically. 'The intent is . . . , more importantly, to reveal and model some of the key elements of scientific discovery in general' (p. 2). History should thereby inform educators 'how scientists are thinking while engaged in scientific discovery' (p. 2). The history here is not just a casual example or illustration. It fits a strategy of drawing on the renown of Galileo's discovery to legitimize science as hypothetico-deductive in nature (p. 21). For this reason, it is all the more critical that any interpretation of history be sound and responsible to criticism.

3. Did Other Scientists Use the Hypothetico-Deductive Method?

Lawson (2002) further claims that 'the hypothetico-deductive pattern of thinking seen in Galileo's discovery of Jupiter's moons [sic] can also be found in the discoveries of other scientists' (p. 15). He provides a table summarizing several such examples (p. 16). One might well be tempted to imagine that even if Galileo's case was not secure, that these other cases, and perhaps many more besides, readily support the thesis that ‘many, if not all, scientific discoveries are hypothetico-deductive in nature’ (p. 21). I will address two of the cases.

3.1. MARCELLO MALPIGHI

First, consider whether Marcello Malpighi used the hypothetico-deductive method to discover capillaries (p. 16; also Lawson 2000, p. 484). Lawson (2000, pp. 483–484) follows Lewis (1988) in identifying the prediction of capillaries as one of William Harvey’s landmark achievements, integral to his theory of the circulation of the blood. Yet Harvey did not predict capillaries (Young 1929, p. 1; Elkana & Goodfield 1968). He did not see arteries and veins as connected by blood vessels. Rather, he said, the blood ‘permeates the pores’ of the flesh (Harvey, 1628, Chaps. 10, 14). It ‘percolates’ through the ‘porous structure’ of the lungs and is ‘drawn’ from the lungs as though from a compressed sponge (Chap. 7). The blood is ‘absorbed’ and ‘imbibed from every part’ by the veins (*A Second Disquisition to John Riolan*, 1648). Many organisms that Harvey dissected – ‘crabs, shrimps, snails and shell-fish’ (Chap. 2) – have hearts but no blood vessels. Hence, it seemed perfectly reasonable based on direct observation not to postulate the need for tiny vessels to complete the circular flow. Indeed, the acceptance of Harvey’s concept by contemporaries did not depend on this supposed promissory note, any more than the acceptance of the Copernican solar system depended on observing stellar parallax. The notion of capillaries was not (*contra* Lawson 2002, p. 16) ‘crucial’ to Harvey’s circulation theory. However, it does fit neatly the preconception of hypothetico-deductive method as essential to science. But if Harvey did not postulate capillaries, how should one interpret Malpighi’s subsequent investigation?

Should we believe, then, that Malpighi made the prediction in Harvey’s stead and organized a ‘planned test’ with the microscope, anticipating that he would observe capillaries as an ‘expected result’ (p. 16; see also Lawson 2000, p. 484)? No. No historical evidence indicates that Malpighi entered his investigation with any intent to validate Harvey’s theory. Nor is there any evidence that he planned or expected to observe small blood vessels connecting arteries and veins. Malpighi made his now landmark observations while focusing (instead) on the fine structure of the lungs, as expressed in the title of his publication, *De pulmonaris*. Here, one may indeed profit from turning to original texts (as noted above). In a letter to his mentor Alfonso Borelli, Malpighi noted his initial beliefs based on the limits of unaided observation ([1661] 1929, p. 8):

> . . . the blood, much divided, puts off its red color, and, carried round in a winding way, is poured out on all sides till at length it may reach the walls, the angles, and the absorbing branches of the veins.
>
> The power of the eye could not be extended further in the opened living animal, hence I had believed that this body of the blood breaks into the empty space, and is collected again by a gaping vessel and by the structure of the walls.

Echoing Harvey perhaps, Malpighi revealed his preliminary belief in the ‘empty space’ between the observable blood vessels, where blood ‘poured out’ and was 'collected again by a gaping vessel'. Using his microscope, Malpighi did indeed observe 'that the blood flows away through the tortuous vessels, that it is not poured into spaces but always works through tubules ... ' (p. 8). But Malpighi did not claim that his observations vindicated Harvey's theory of circulation. Indeed, Malpighi did not refer to Harvey at all. The observation of capillaries was not an 'expected result'. Rather, it seemed quite unexpected. Nor was it part of any 'planned test' about circulation. Hypothetico-deductive method did not guide this discovery.

One might well imagine an alternative universe where someone predicted capillaries and then observed them in a planned test. It could happen. But this is not *our* history. Substituting an idealized or fantasized history for an authentic one tends to discredit the original discovery – and discounts whatever process led to it. If the goal is to characterize the process of scientific discovery, however, one might well profit from focusing on how this and other such discoveries actually occur.

What factors seem important in this case? First, as in the case of Galileo's telescope, the instrument and the technique of using it are essential. The microscope itself has a history of development (e.g., Wilson 1995; Ruestow 1996). This is essential to the discovery, as underscored in Malpighi's own testimony. In addition, Malpighi comments several times about the methods of preparing a specimen and the methods of lighting it (e.g., 'by the microscope of one lens against the horizontal sun', or 'on a crystal plate illuminated below through a tube by a lighted candle'; [1661] 1929, p. 9). These observational skills are garnered from experience: from tinkering and from exploratory trial and error (not from deep conceptual hypothesizing or prediction). Here, again, elements of experimental practice are constitutive of the process (§2). Moreover, Malpighi was fortunate to have observed a frog. The connections would not have been visible in mammals with the type of microscope he was using. He also used a tortoise (p. 8), where the blood vessels were more clearly visible. Here, Malpighi had been guided by a principle of comparative anatomy, or analogy between types of organisms:

> For Nature is accustomed to rehearse with certain large, perhaps baser, and all classes of wild (animals), and to place in the imperfect the rudiments of the perfect animals. ([1661] 1929, p. 7)

Such comparisons do not generate hypotheses nor do they constitute a planned test of an expected result. Rather, they orient exploratory search. Malpighi's particular choice of a frog may also have involved a degree of luck or coincidence, which can also contribute to scientific discovery (Judson 1981; Roberts 1989). Ultimately, transferring conclusions from other animals to humans without confirmation is another example of analogy, or non-enumerative induction (inference based on similarity). Thus, an account of the discovery of capillaries that excludes the development of the microscope and its use, observational heuristics, analogy and, possibly, chance is at best incomplete and misleading, at worst decidedly incorrect. Such omissions may occur, however, if preconceptions of scientific method limit too strongly what a historical episode may 'tell us about the process of scientific discovery'.

3.2. GREGOR MENDEL

Next, consider whether Gregor Mendel used a hypothesis of independent assortment in genetics to discover independent assortment (Lawson 2002, p. 16). My locution here may seem odd. It seems to beg the question of discovery, or the origin of the hypothesis, as noted earlier in Galileo's case (§2). It also highlights a potential problem with characterizing the hypothetico-deductive method. If the method describes a general process of discovery, then we should expect it to generate hypotheses or important novelties. If the method merely describes how hypotheses are tested, then the "reasoning" amounts to little more than an expression of the principle of empirical relevance in science. Everyone might agree to that. But it would hardly reflect a *method* of science. Even less a method *of discovery*.

Mendel constitutes an extraordinarily interesting case of discovery. Historians continue to debate just how much he understood before he began his quantitative studies in earnest. Yet certain, sometimes surprising, aspects of his investigations are now clear. For example, Mendel did not state explicitly any 'Law of Independent Assortment', although his data illustrate it (Olby 1979). Mendel certainly showed that he understood independence, in the mathematical sense of the relationship of two probabilities. But he seems not to have distinguished what we now call segregation and independent assortment. Later, William Bateson, Mendel's chief champion in England, also initially conflated the two (Olby 1985). Not until Bateson observed odd phenotypic ratios (e.g., $12:1:1:3$), where genes segregated but were *not* independently assorting, did he formulate the distinction clearly. Thus, a historian would not say that Mendel formulated a law (or hypothesis) of independent assortment.

What, then, did Mendel think he was investigating? Most historians now agree (*contra* Lawson 2002, p. 16) that Mendel was not seeking abstract laws of inheritance (e.g., Olby 1974, Monaghan & Corcos 1990). Rather, he was investigating the nature of hybridization, as expressed in his title. He was possibly trying to understand how (or when) hybrids might breed true and/or how hybridization relates to the nature of species and Darwinian evolution. Mendel's main conclusion appears to be expressed in six nearly identical statements in his now landmark paper (Hartl & Orel 1992). Each describes a $\frac{1}{2}:\frac{1}{4}:\frac{1}{4}$ distribution of offspring, which Mendel seems to label a 'law of combination of differing traits according to which hybrid development proceeds'. Mendel's 'law of hybridization' thus seems to follow in a tradition of framing basic laws: mathematical regularities of nature, such as Snel's Law of Refraction or Boyle's Law of Gases. One may note that these laws originate as empirical generalizations. No theory or concept - -or specific prediction or expected result – guided their discovery. Rather, they exemplify plain arithmetic analysis and enumerative induction (summarily dismissed in Lawson's analysis of possible methods in science [2002, pp. 15–18]).

Some question may remain whether Mendel anticipated his results. Modern biologists, for example, may be impressed that Mendel's seven characters exhibited independence rather than linkage. Such apparent coincidence seems to hint of deliberate choice, guided by intimate knowledge of inheritance patterns. However, as Di Trocchio (1991) has noted, Mendel reported experimenting with *twenty-two* varieties, not just seven (p. 495). Mendel's strategy was apparently to search for patterns somewhat blindly among a vast number of controlled crosses. He would then have focused on just the meaningful results (i.e., those he could interpret). The reported results would be self-selected: hence, the seven traits so renowned today, only four of them in dihybrid crosses. Other results would have been discarded or disregarded as too confusing (due, in modern terms, to linkage or absence of dichotomous dominant/recessive hybrids). Several details support this interpretation of Mendel's work: the limited availability of true-breeding varieties, time factors and Mendel's comments on traits other than the canonical seven. It is consistent, too, with practices about manipulating and reporting data during the period (prior to statistical norms). Mendel seems to have had no 'planned test' of any specific hypothesis. He had no 'expected result'. (Yet he discovered something significant just the same!) No overarching hypothetico-deductive method seems to have guided Mendel's discovery of his law of dihybrid development. Rather, it was a combination of blind search and selection, and limited induction across a number of cases with similar arithmetic patterns.

Interpreting Mendel can be as challenging as interpreting Galileo. For example, Jan Sapp (1990) has noted how many biologists present their claims as commensurate with Mendel, even though such claims contradict one another. Mendel-the-historic-hero is a powerful ally rhetorically, and biologists try to shape Mendel to their ends. As a result, Mendel can be as Protean as Galileo. In characterizing Mendel, therefore, caution and reliance on documented historical facts are essential. In this case, his "method" seems to have been far more complex – and far more interesting – than in Lawson's summary capsule (2002, p. 16). Historical facts can keep preconceptions in check, to our benefit.

In summary, these two historical cases (Lawson 2002, p. 16) do not exhibit hypothetico-deductive reasoning. The historical evidence does not match the philosophical hypothesis that 'many, if not all, scientific discoveries are hypothetico-deductive in nature' (p. 21). Someone who advocated the hypothetico-deductive method should therefore, it seems – strictly by their own standards – reject it as a universal method (Donovan et al. 1988). Hypothetico-deductive reasoning may well have a role in science, but these examples do not illustrate it, much less prove that it is universal or nearly so.

4. Lawson's Shoehorn

Pinpointing specific historical errors (§§2– 3) to forestall their propagation in science classrooms or among science educators is important. So, too, is correcting specific philosophical errors based on, or presented as justified by, such errors.[4] However, a fruitful analysis will probe even deeper. That is, one may consider how a series of specific errors may reflect a more general, or systematic, error. In this case, the errors above follow a syndrome. All the errors may thus be reexpressed as a single, core error. The fundamental error is *trying to fit the history of science into one particular philosophical conception of science*. Readily available historical information on these cases indicates other patterns and methods. Lawson's history shows evidence everywhere of being adapted. Relevant facts are omitted, false or imaginary details are added, emphases are misplaced, and qualifications are abandoned – all to accord with a model of hypothetico-deductive reasoning. One can hardly portray the practice of science authentically with such distorted history. By analogy with Gould's analysis of Walcott (§1), I call this dominance of philosophical preconception in science education *Lawson's shoehorn*.

Charles Walcott was probably not aware of his taxonomic shoehorning (Gould 1989, pp. 244–277). Nor need one suppose that Lawson's work reflects deliberate distortion or conscious manipulation of historical facts (more below). Lawson's goal is clearly stated: 'to help students learn how to do science' (2002, p. 2). Thus, educators may support Lawson's aim by finding ways to avert the errors of Lawson's historical shoehorn.

Brush (1974) considered whether the history of science should be rated 'X' for students. Based on the potential for distortions noted here, one might wonder whether such an 'X'-rating should also apply to *science educators*. Of course, 'X'-ratings are not outright bans. Rather, they seek to limit exposure to those who have gained the intellectual resources to address the material responsibly (Allchin 1995). But here, the problem is deeper than the history itself. Rather, it is a philosophically narrow approach to history. The problem emerges from how philosophy of science is used. Well, then, should the *philosophy* of science be rated 'X'?

While I am fascinated by the image of a corps of whistle-blowing philosophy-of-science police "busting" dangerously wayward teachers, such playful reflection flirts unnecessarily with academic freedom, among many other problematic issues. Educators must come to terms with the errors of shoehorning by another route. Indeed, conceptual shoehorning may reflect how human brains function unchecked. Sunderland (1992), for example, underscores the pervasiveness of the 'availability error', the tendency to be biased by first impressions and use them to filter further thinking by noticing confirming examples and disregarding exceptions. Gilovich (1991), similarly, stresses how we often see what we expect to see, especially by evaluating ambiguous, incomplete or inconsistent data according to preconceptions. 'Information that is consistent with our pre-existing beliefs is often accepted at face value, whereas evidence that contradicts them is critically scrutinized and discounted' (p. 50). Such was Walcott's 'cardinal error', according to Gould. Gould claims that under ordinary circumstances, data should keep a healthy check on preconceptions in science, limiting their potential to mislead. We may transfer that standard to historical facts in histories of science in the classroom. To minimize error in either context, facts that challenge preconceptions must be acknowledged and addressed. Ultimately, regulation of thinking is important – but not by external policing or legislating pre- approved ideas. Rather, one may consider "counter-shoehorning" an individual skill or habit. Of course, it must be learned – and, hence, taught.

Science education – and science teacher education, as well – need to foster an understanding of and practical skills in cognitive checks and balances. One may first consider Karl Popper, widely known for the concept of falsifiability (that science is demarcated by ideas that can potentially be contravened by empirical evidence). The pitfalls of simple falsification in actual scientific practice are well known to philosophers, even among Popper's staunchist supporters (e.g., Lakatos 1970). However, Popper advocated more generally the importance of criticism (especially in his political writings). In science, he introduced the allied notion of a *severe test*. For Popper, tests based on easy confirmation counted little. An ideal test was framed critically, almost in an unforgiving sense. A severe test thus offered ample opportunity for "failure". The more severe the test (the greater the potential for being falsified), the better. Deborah Mayo (1996) has recently revived Popper's notion and developed it further in statistical contexts. Whereas Popper focused almost exclusively on falsification, Mayo now highlights the corresponding positive role of *passing* severe tests (for example, using a low p-value to reject a null hypothesis). For Mayo, reliability hinges on a dual process of confirmation *and* ruling out error (pp. 4-7, 184–185, 315). It is not enough, for example, merely to advance a hypothesis, deduce some of its implications, and then confirm them through testing. The hypothetico-deductive method leaves too much open to error. One must pursue *severe* tests. Mayo thus adds an important principle for regulating error: *error probes* (pp. 64, 445). That is, to deepen reliability, one must actively and aggressively search for possible mistakes. For example, one might probe whether hitherto uncontrolled variables are relevant. In the spirit of a medical diagnostic probe, an effort to falsify can be a constructive tool for arguing (conversely) about reliable fact (p. 183). Error probes thus potentially serve as a methodological corrective to conceptual shoehorning.

Error probes seem to echo many characterizations of nature of science in K-12 teaching. These standards (for example, Rutherford & Ahlgren 1990, National Research Council 1996) typically profile a 'skeptical attitude' or 'organized skepticism' as part of healthy science. But skepticism, as mere doubt (in the Cartesian tradition, for example), is undirected and not based on particular reasons. It is a general reminder to demand evidence: not enough to stop a shoehorn. Likewise, 'critical thinking' often reduces to unfocused criticism or license to denounce opposing views (Lawson 2002, pp. 15–21). Error probes go further conceptually. They involve actively reviewing evidence for potential errors (Allchin 2001). They shift the primary focus from 'Does this claim seem warranted, based on the evidence provided?' to 'How might claim this be wrong, even with the evidence provided?' They are also reflexive. Not: 'Why is this right?' Rather: 'In what ways might I be misleading myself?' Error probes articulate the limits of claims and thus the specific ways further evidence may help expose hidden assumptions or resolve residual uncertainties. They highlight the area eclipsed by shoehorning.

Someone schooled in error probes will be equipped to approach histories of science as well. Considering Lawson's interpretation of Galileo's discovery of the moons of Jupiter, one can see immediately that although it fits the evidence, it is all supposition, and thus rife with potential error. The accounts of other important scientific discoveries are not referenced to any peer reviewed historical accounts. Without evidence available for inspection, they are suspect, too. What other hypotheses or tests were also entertained? What happened prior to each test? What led to framing each hypothesis? What creative or technological work contributed to getting results? Who may have criticized the interpretation of results and for what reasons? The remedy to Lawson's shoehorn in science education, as for Walcott's shoehorn in science, may be active deployment of error probes.

No one will dispute, I think, Lawson's claims (p. 21) about the importance of teaching hypothetical reasoning to students, as one among many modes of thinking – and one which students often find difficult. But no one needs history to support this claim. Nor would anyone likely deny that *some scientists* on *some occasions* have profitted from hypothetical reasoning, especially in testing tentative ideas. Or that science is often *reconstructed* in a deductive format for reporting and persuasive purposes. Even less is anyone likely to criticize the development of thinking skills, such as generating alternative hypotheses, designing appropriate tests, or evaluating results in the light of possible explanations. But these piecemeal positions do not amount to a conclusion about one exclusive, monumental, algorithmic method of science (Lawson 2002, pp. 19–21).

Ultimately, science education may thus be enriched by historical cases, such as Walcott's, showing the potential for error among well intentioned scientists. Likewise, science educators may be better informed by learning examples, like Lawson's, showing how philosophical preconceptions can strongly distort histories of science – and thus lessons about the nature of science.

Acknowledgement

The author acknowledges support from the Herbert P. and Alice W. Bailey Trust.

Notes

1 The hypothetico-deductive method (HD), according to Lawson (p. 17), is similar to the much critiqued 'Scientific Method' (e.g., Bauer 1992). As such, it differs from HD as characterized by Hempel (1966) and others who underscore the importance of general laws and the particular instances derived from them using numerous auxiliary hypotheses and boundary conditions. For example, here in the episode of Jupiter's moons, the hypothesis here is not the Copernican hypothesis, of which Jupiter's moons become a specific prediction. Lawson's informal 'if-then' formulation (pp. 6, 9) opens the way for *any* form of reasoning expressed in 'if-then' terms to count as HD, as exemplified in his analysis. Lawson further conflates HD with Chamberlain's method of multiple working hypotheses

(p. 15), which importantly adds to HD an important dimension of variation and selection. The role of alternative hypotheses in testing is common (e.g., experimental controls), but HD proper is silent on such matters, as its logical structure links only one hypothesis and its consequences. A more inclusive framework, such as eliminative induction or inference to the best explanation, requires a more sophisticated framework than conventional HD describes. Lawson's characterization also does not allow one to distinguish between genuine discovery (of novelties) and search within a known conceptual space (such as medical diagnosis). Despite all these variations with the extant literature on HD, I have tried to follow Lawson's usage.

[2] In a similar manner, scientific studies of long-term potentiation (LTP) (e.g., Johnston 1997; McGaugh 2000; Kandel 2001) are missing from Lawson's survey of cognitive and neurological models (pp. 9–15). LTP relies on repeated, or reinforced, stimuli to establish the new synaptic connections critical to learning. Activity-dependent plasticity is, essentially, induction on a cellular level. Lawson's treatment seems again to disregard models that challenge the exclusivity of hypothetico-deductive reasoning. Like the methodological focus on search rather than novel hypothesis generation, they also tend to focus on recognition or recall, rather than learning.

[3] Historians often endeavour to reconstruct plausible thought processes of historical figures, a professional skill known as the *historical imagination*. Historians are schooled to exercise this skill conservatively, lest they misrepresent history. The skill depends very much on familiarity with the context of the period and being steeped in relevant details. The numerous historical errors I document may speak to the depth of understanding of historical context here.

[4] While Lawson cites an apparently impressive array of literature to support his philosophical perspective, it is nonetheless selective, also. Five of the nine texts cited were published prior to 1970 and one may question whether they reflect 'contemporary' consensus (Lawson 2002, p. 15). Notably missing are works by Kuhn, Laudan, Lakatos and Hull, who all addressed the issue of philosophy informed by the history of science. The other 'accounts of science' include a critical thinking text, a reflection by a retired biologist, a biology textbook that actually endorses many methods, and an introductory philosophy of science text that is far more accommodating than the citation would suggest. The limitations of the hypothetico- deductive method have been widely critiqued and need not be echoed here. My focus, instead, is how the history, strongly shaped by a philosophical view, is used inappropriately to support the philosophical view.

References

Allchin, D.: 1995, 'How Not to Teach Historical Cases in Science', in F. Finley, D. Allchin, D. Rhees and S. Fifield (eds.), *Proceedings, Third International History, Philosophy and Science Teaching Conference*, University of Minnesota, Minneapolis, MN, Vol. 1, pp. 13–22.

Allchin, D.: 2001, 'Error Types', *Perspectives on Science* **10**, 38–58.

Bauer, H. H.: 1992, *Scientific Literacy and the Myth of the Scientific Method*, University of Illinois Press, Urbana, IL.

Bazerman, C.: 1988, *Shaping Written Knowledge: The Genre and Activity of the Experimental Article in Science*, University of Wisconsin Press, Madison, WI.

Biagioli, M.: 1993, *Galileo, Courtier*, University of Chicago Press, Chicago, IL.

Booth, V.: 1993, *Communicating Science*, 2nd ed., Cambridge University Press, Cambridge.

Di Trocchio, F.: 1991, 'Mendel's Experiments: A Reinterpretation', *Journal of the History of Biology* **24**, 485–519.

Donovan, A., Laudan L. & Laudan, R.: 1988, *Scrutinizing Science: Empirical Studies of Scientific Change*, Johns Hopkins University Press, Baltimore, MD.

Elkana, Y. & Goodfield, J.: 1968, 'Harvey and the Problem of the "Capillaries" ', *Isis* **59**, 61–73.

Finocchiaro, M. A.: 1997, 'Varieties of Rhetoric', in *Galileo on the World Systems*, University of California Press, Berkeley, CA, pp. 356–372.

Franklin, A.: 1986, *The Neglect of Experiment*, Cambridge University Press, Cambridge.
Franklin, A.: 1991, *Experiment, Right or Wrong*, Cambridge University Press, Cambridge.
Franklin, A.: 1997, 'Calibration', *Perspectives on Science* **5**, 31–80.
Galison, P.: 1987, *How Experiments End*, University of Chicago Press, Chicago, IL.
Gilovich, T.: 1991, *How We Know What Isn't So*, Free Press, New York, NY.
Gould, S. J.: 1989, *Wonderful Life*, W. W. Norton, New York, NY.
Hacking, I.: 1984, *Representing and Intervening*, Cambridge University Press, Cambridge.
Hartl, D. L. & Orel, V.: 1992, 'What did Gregor Mendel think he discovered?', *Genetics* **131**, 245–253.
Harvey, W.: [1628] 1952, *De Motu Cordis et Sanguinis in Animalibus* [On the Motion of the Heart and Blood in Animals], trans. by Robert Willis, Encyclopedia Britannica, Chicago, IL.
Hempel, C.: 1966, *Philosophy of Natural Science*, Prentice-Hall, Englewood Cliffs, NJ.
Jardine, N.: 1994, 'A Trial of Galileos', *Isis* **85**, 279–283.
Johnston, D.: 1997, 'A Missing Link? LTP and Learning', *Science* **278**, 401–402.
Judson, H. F.: 1981. 'Chance', in *The Search for Solutions*, Holt, Rinehart & Winston, New York, NY.
Kandel, E. R.: 2001, 'The Molecular Biology of Memory Storage: A Dialogue Bewteen Genes and Synapses', *Science* **294**, 1030–1038.
Knorr-Cetina, K.: 1984, *The Manufacture of Knowledge*, Pergamon Press, Oxford.
Kuhn, T. S.: 1970, *The Structure of Scientific Revolutions*, 2d ed., University of Chicago Press, Chicago, IL.
Lakatos, I.: 1970, *The Methodology of Scientific Research Programmes*, Cambridge University Press, Cambridge.
Lawson, A.: 2000. 'The Generality of the Hypothetico-Deductive Method: Making Scientific Thinking Explicit', *American Biology Teacher* **62**, 482–495.
Lawson, A.: 2002. 'What Does Galileo's Discovery of Jupiter's Moons Tell Us about the Process of Scientific Discovery?', *Science and Education* **11**, 1–24.
Lewis, R. W.: 1988, 'Biology: A Hypothetico-Deductive Science', *American Biology Teacher* **50**, 362–366.
Malpighi, M.: [1661] 1929, 'On the Lungs', trans. by J. Young, *Proceedings of the Royal Society of Medicine* **23**, 1–11.
Mayo, D.: 1996, *Error and the Growth of Knowledge*, University of Chicago Press, Chicago, IL.
McGaugh, J. L.: 2000, 'Memory – A Century of Consolidation', *Science* **287**, 248–251.
Medawar, P.: 1964, 'Is the Scientific Report Fraudulent? Yes: It Misrepresents Scientific Thought', *Saturday Review* 47(August 1, 1964), 42–43.
Monaghan, F. V. & Corcos, A.: 1990, 'The Real Objective of Mendel's Paper', *Biology and Philosophy* **5**, 267–292.
Moss, J. D.: 1993, *Novelties in the Heavens: Rhetoric and Science in the Copernican Controversy*, University of Chicago Press, Chicago, IL.
National Research Council: 1996, *National Science Education Standards*, National Academy Press, Washington, DC.
Olby, R. C.: 1979, 'Mendel No Mendelian?', *History of Science* **17**, 53–72.
Olby, R. C.: 1987, 'William Bateson's Introduction of Mendelism to England: A Reassessment', *British Journal for the History of Science* **20**, 399–420.
Pitt, J.: 2000, *Thinking about Technology*, Seven Bridges Press, New York, NY.
Roberts, R.: 1989, *Accidental Discoveries in Science*, John Wiley & Sons, New York, NY.
Rothbart, D. & Slayden, S. W.: 1994, 'The Epistemology of a Spectrometer', *Philosophy of Science* **61**, 25–38.
Ruestow, E. G.: 1996, *The Microscope in the Dutch Republic*, Cambridge University Press, Cambridge.

Rutherford, F. J. & Ahlgren, A.: 1990, *Science for All Americans*, Oxford University Press, New York.

Sapp, J.: 1990, 'The Nine Lives of Gregor Mendel', in H. LeGrand. (ed.), *Experimental Inquiries*, Kluwer Academic, Dordrecht, pp. 137–166. Also available at MendelWeb, URL: www.netspace.org/MendelWeb/.

Shea, W.: 1998, 'Galileo's Copernicanism: the Science and the Rhetoric', in P. Machamer (ed.), *The Cambridge Companion to Galileo*, Cambridge University Press, Cambridge, pp. 211–243.

Sunderland, S.: 1992, *Irrationality*, Rutgers University Press, New Brunswick, NJ.

White, H.: 1987, *The Content of the Form*, Johns Hopkins University Press, Baltimore, MD.

Wilson, C.: 1995, *The Invisible World: Early Modern Philosophy and the Invention of the Microscope*, Princeton University Press, Princeton.

Young, J.: 1929, 'Malpighi's "De Pulmonibus" [Introduction]', *Proceedings of the Royal Society of Medicine* 23, 1.