

Teaching the Nature of Science through Scientific Errors

Douglas Allchin
University of Minnesota

ABSTRACT. Error in science is a prime occasion to teach the nature of science, especially the central feature of tentativeness. Here, I survey philosophical, historical and educational literature and classroom strategies and resources.

Error & the Nature of Science

If the goal is to teach “how science works,” then it seems equally important to teach, on some occasions, how science *doesn't* work. Imagine teaching law and law enforcement without crime. Or medicine without disease. One needs to understand how health can fail or how laws can be broken if one is to understand how the relevant systems function properly. Biologists and engineers are already quite familiar with cases of loss of function as vehicles for research and for teaching structure and function (Bechtel & Richardson 1993, Petroski 1994). So, too, for the nature of science and the basis for the reliability of scientific claims.

For over half a century, one feature has remained central in virtually every recommendation to teach the nature of science: that science is “tentative” (Lederman, Wade & Bell 1998, HIPST 2008). The label varies. Some say that science is fallible, provisional, or contingent. Others say it is developmental, changeable, or subject to revision — but with a sense that new concepts replace old ones, not merely that knowledge grows or progresses cumulatively (Kuhn 1970). Ultimately, scientists can err. Scientific claims can later prove to be mistaken. The corresponding challenge, then, is to teach fully about error in science.

The lessons of error are especially critical for scientific literacy. Consumer-citizens need to interpret, say, revised mammogram recommendations or the retraction of a study linking autism to the measles vaccine. One can teach from historical cases, such as the episodes that recast beriberi and pellagra as nutrient deficiencies rather than infections, or that discredited fraudulent studies linking race and IQ. If students are ever to learn that science is “tentative” they must, of course, encounter real examples of scientific error or failure. One may also couple these with examples of scientific change and the discovery and remedy of error. Just alluding to the potential for failure cannot dislodge the potent cultural image of science as an amalgam of fact, certainty and incontrovertible evidence. A mere mention that “science is tentative” seems to function culturally like an escape clause, excusing science any time it does not meet the ideal. The ideal persists. The spectrum of error types must be explained.

Philosophers and sociologists of science have long appreciated that methodological norms complement particular error types. For example, blind clinical trials are adaptations to the placebo effect among patients. Double blind studies accommodate the observer effect on the part of researchers. Controlled experiments are tools to expose the potential causal role of coincident variables. Statistical analysis became a way to describe and quantify sampling error (Porter 1986, Hacking 1990). Calibration ensures that an instrument's readings do not stray from their designed purpose (Franklin 1998). Peer review helps reduce errors in communicating results. All these methods have a history. They were not obvious at first. Scientists learned them through reflection and often after encounters with error. Even the notion of experimental control has a history (Boring 1954, Lilienfeld 1982). The earliest examples of control, one finds,

nearly always introduced a second experiment mindful of potential criticism or ensuring that a target cause was not mistaken. Only in the mid to late nineteenth century does one find the method articulated and the term ‘control’ introduced (Mill 1874, pp. 278-310; OED 1971, I, p. 927). The methods of science are, in many ways, the hard-won wisdom of experience. Similarly, students must learn the role of each method. History can be a teacher’s roadmap — perhaps guiding lessons without any of the costly negative consequences.

Although scientists may succumb to errors, as in any human endeavor, they surely aim to prevent, mitigate or accommodate errors. Ideally, researchers actively review their work for possible errors. This posture of critical questioning is often described as a skeptical attitude and hailed as a hallmark of science. Yet skepticism can easily become simply aimless disbelief. Vague skepticism differs significantly from a focused and targeted search for *particular* sources of error. Teachers should thus avoid the rhetoric of skepticism and teach instead the role of concrete error analysis.

My aim here is to survey some of the philosophical and educational literature on scientific error and to profile a number of fruitful teaching strategies and classroom resources — especially including the wealth of sources that seemed to have escaped the recent notice of Kipnis (2010).

The Heritage of Error in Philosophy of Science & Science Education

Concerns about portraying scientific error in science education extend back over half a century. For example, Leonard Nash (1951), co-founder and co-editor of the *Harvard Case Histories in Experimental Science*, commented on the motivations for using history, linking them to profiling error and the nature of science. He was concerned about portraying the misconstrual of experimental data, whether it was the historical case of weight gain of calcined metals or a contemporary case of the radioactive dating of meteorites. He also cited Dalton’s prejudices in the development of atomic theory:

Dalton's theory was derived by a wild and almost completely erroneous line of argument based on physical phenomena that were largely irrelevant to a chemical atomic theory. (p.151)

Rejection of Avogadro’s hypothesis was among Dalton’s (and others’) errors. Thus, Nash advised a role for history:

We believe that in calling attention to the difficulties, delays, and failures of science, as well as to its triumphs, we present a truer, better-rounded, and much more useful picture of it as it was in the beginning and as it is today. (p. 151)

Nash felt that the informed citizen needed to understand how scientists could err.

High school chemistry teacher Herb Bassow (1991) also noted the significance of conveying the importance of errors in science. He illustrated this with a case from the 1920s where a graduate student, E. Bächlin, encountered problems measuring the wavelength of X-rays. His research led him through a thicket of prior error: from Millikan publishing an incorrect value of the charge of the electron (having used an incorrect value for the viscosity of air in the Stokes equation) through calculations of Avogadro’s number (based on Millikan’s e) to calculations of atomic distances in salt crystals (based on N) and calculations of the X-ray wavelength based on the X-ray crystallography of NaCl (using a now erroneous d in Bragg’s equation). Errors by 5 Nobel Prize winners, in all. How better to convey the interdependency of science?, Bassow argued.

Earlier I, too, profiled “the virtues of teaching the wrong ideas from history” (Author, 1995). History helps us understand, first, how “wrong” ideas could once be considered “right.” It also shows how such views changed. History thus shows how science can err *and* how it can remedy those errors. Addressing historical errors contributes to understanding both the nature of scientific justification and its limits. Philosopher of science John Losee (2005) followed this approach in profiling “theories on the scrap heap”: namely, why scientists abandoned theories they once held as true. Learning how science could once reach unreliable conclusions is just as important as how it might reach reliable conclusions — at least if the modern student is to learn how to recognize so called Intelligent Design advocates or climate change naysayers as modern day counterparts to the followers of Velikovsky.

Ironically, we may not wish to banish astrology, alchemy, phrenology, craniology, mesmerism or other pseudoscience from the science classroom (Author, 1995). Historically, of course, each of these practices was once considered science — in some cases, exemplary science. Students are no different. They need to work through such examples to understand what makes modern science “science.” History of pseudoscience offers an ideal occasion for doing so. Every generation must re-learn what is scientific, what pseudoscientific (Author 2004a, pp. 189-191).

As noted by Wandersee (1986), historical error may also be important for understanding student misconceptions and identifying what may lead them to deeper understanding. Historical errors may also be fruitful vehicles for teaching content effectively.

Errors are also valuable for counteracting the distortions of scientific “myth-conceptions,” the idealized and monumentalized images of science typically found in the media and textbooks. To trigger reflection and initiate learning on the nature of science, I observed, one must use discrepant events or anomalies to the entrenched assumptions about School Science (Author, 2003a). Cases of reasonable error or fallibility, especially among scientific heroes, may function as a corrective to the mythic caricatures. Teachers should show that sometimes following the “right” methods can lead one to the “wrong” conclusion (at least in the short term). Even idealized science and heroic scientists are not perfect. Accordingly, we need to differentiate between idealized role models and *real* models (Author 2008, p. 504). The former may be so idealized that students view them as unattainable (that is, as *negative* role models). The latter are more human in scale and more accessible to a wider variety of students, especially women and minorities. Portraying error in science reveals its human dimension and may have important implications for recruitment into scientific careers.

Mermelstein and Young (1995) also profiled a role for error in the process of science and learning. They emphasized that individual “mistakes are vehicles for discovery,” both among scientists and students. They followed, in particular, the lead of William Beveridge in *The Art of Scientific Investigation* (1957):

Perhaps the most striking examples of empirical discoveries are to be found in chemotherapy where nearly all the great discoveries have been made by following a false hypothesis or a so called chance observation. (p.43)

That is, student errors are ideally treated as “springboards for inquiry” (p. 773). This psychological process at the individual level is then integrated with the reduction of error at the level of the collective, or scientific community. Scientific knowledge grows by trial and error, of course, but not blindly. Science — and science education — needs appropriate management of those errors.

Carmen Guinta (2001), manager of the ChemTeam website for historical papers in

chemistry, has also turned to the role of errors in “using history of chemistry to teach scientific method.” He observes that errors help mitigate two major pitfalls in historical lessons: hero worship, along with its “unwarranted impression of infallibility”; and impressions of scientific method as algorithmic, rather than guided by “judgment and imagination” (p. 623).

Acknowledging the false steps of great scientists avoids putting those scientists on a pedestal and illustrates that there is no automatic or mechanical scientific method. (p. 626)

Guinta, too, finds error essential in rendering the persistent nature of science theme of tentativeness. He provided several historical examples in chemistry.

One particularly prominent error in science and elsewhere is confirmation bias (Gilovich 1991, Sutherland 1992, Nickerson 1998). That is: “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” (Gilovich 1991, p. 50). Earlier I noted the need to profile this cognitive tendency in science (as well as in ideas *about* science) and to foster practical skills in cognitive checks and balances. (Author, 2003b). One solution to this problem is a variant of Karl Popper’s severe tests:

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. [Deborah] Mayo thus adds an important principle for regulating error: *error probes* ([1996] pp. 64, 445). That is, to deepen reliability, one must actively and aggressively search for possible mistakes. (Author, 2003b, p. 325)

Confirmation alone can be epistemically treacherous. An awareness of errors in science can highlight importantly how the canonical “scientific method” — simple hypothetico-deductive thinking, without robustness or consideration of alternatives or sources of error — can be flawed.

Kipnis (2010) recently claimed that “So far, incorporating the subject of error into science education apparently has been limited to errors of measurement ... and ethical issues...” (Section 1.1).¹ This seems in conflict with the important precedents and the rich heritage of the idea of addressing error in the sources noted above (also see Author, 2004c, for an online essay for high school teachers).

Error in science has been a focus of many popular books: Nichols (1984), Kohn (1988), Youngson (1998), Gratzner (2000) and Grant (2006). That includes one (rather delightful) children’s book: Kelly & Parker (1996). As one might expect, they largely entertain more than inform. They generally portray error in science as a blend of embarrassment and bemusement, as reflected in their titles: *Scientific Blunders*, *Shocking Science*, *False Prophets*, *The Undergrowth of Science*, and *Discarded Science*. There is little engagement with error as a normal feature of science. None of these volumes reflect the consensus in the current nature-of-science educational literature and policy documents for rendering science as tentative.

¹For example, Zachos, Pruzek & Hick (2003) also opened a “place for the issue of scientific error in the secondary science curriculum” (p. 947). They framed a vast scope for error — for example, drawing on Bacon’s characterization of the many “idols,” or psychological sources of bias in reasoning — yet remained focused rather narrowly on experimental errors in measurement, drawing on the conventional distinction between measurement uncertainty (a form of randomness misleadingly called “error”) and systematic bias.

Still, there have also been many scholarly treatments of error. Even some of these authors, however (typically veteran scientists), seem to revel in portraying error in science as *pathological* (Langmuir 1989, Rousseau 1992, Dolby 1996). Their retrospective accounts betray efforts to bolster the pristine authority of scientists, and tend to suffer from historiographic problems. By disregarding historical context, they oversimplify and idealize science in a way that is only possible with the privilege of hindsight (Butterfield 1959, Chang 2009). Error in science is seldom about delusion, self-deception or human frailty. That posture exhibits a poor appreciation of error types, the sources of error, and the history of scientific error.

A number of volumes focus on more general cognitive errors, which are equally relevant to science. While aimed for a broader audience, they generally summarize research in a balanced and informative way: Gilovich (1991), Sutherland (1992), Shermer (2002), Hallinan (2009) and Schulz (2010). A few philosophers have begun to sketch the implications of such cognitive limitations for scientific practice and scientific discovery: Bechtel & Richardson (1992) and Wimsatt (2007). Another two volumes focus, more journalistically but still responsibly, on a socially based error — fraud: Broad & Wade (1982), Judson (2004) (for more on conceptualizing fraud as a form of error, see below).

Four scholarly volumes highlighting the philosophical dimensions of error in science are especially valuable. The first is Deborah Mayo's *Error and the Growth of Experimental Knowledge* (1996). Error statistics form a core concept, from which Mayo explores many traditional philosophical problems, such as underdetermination (the Duhem-Quine problem), Popper's severe tests, hunting and snooping, and an analysis of Peircian error correction and induction. In *Theories on the Scrap Heap*, John Losee (2005) takes a clever twist on the usual philosophical question of why theories are *accepted* and asks why they are sometimes *rejected*. He examines numerous historical cases and such issues as the role of falsification, prediction (vs. accommodation) and personal thematic style. Next is *Errors: Consequences of Big Mistakes in the Natural and Social Sciences*, a special issue of *Social Research* edited by Gerald Holton and Arien Mack (2005). As Holton notes in his introduction:

At first blush, “science” and “error” seem to be polar opposites— the one a heroic pursuit of provable and widely sharable truths, the other a miserable exemplar of human frailty. (p. vii)

Holton summarizes the spirit of the volume that any heroism in science is largely due to the dogged struggle with errors. There are insights, to be sure,

But on the way to those rare eureka moments, practitioners of science know well that the path is strewn with hurdles and pitfalls, costly detours, with minor and major blunders and gremlins in the experimental equipment or in the theoretical presuppositions. (p. viii)

Finally, one may note another collection, based on a symposium largely limited to just the realm of experiment, but casting a wide net in understanding error more broadly as “going amiss”: Hon, Schickore & Steinle (2009).

There are also a handful of philosophical papers that are particularly significant. Fred Suppe (1988) analyzed the structure of arguments in a large sample of scientific papers. His surprising result was that they do not reflect the canonical hypothetico-deductive method, often portrayed as “the” scientific method (Hempel 1966), and promulgated especially in an educational context by Anton Lawson (2002, 2003, 2004, 2009). Such an argument structure discounts the importance of the Methods section, for example. A fine scaled parsing of each step of the typical argument shows that the scientific paper is structured around the many possible

sources of error. Attention is given to each possible alternative interpretation of the evidence. Namely, the paper argues its conclusion stepwise, addressing each potential flaw in procedure, observation and reasoning. This focus on error in argument is especially relevant given recent views on the role of scientific argument and justification in teaching nature of science (for example, Osborne 2010). Two other philosophical articles on the classification of errors in science (Hon 1989, Author 2001) are addressed below.

Sociological studies of error in science are few, despite the vast literature on scientific controversies. A notable contribution is Star & Gerson's (1986) anthropological analysis of error. They described error *behaviorally*: observable as an interruption of the work flow. Their categories of error, nonetheless, tend to match a few philosophical analyses (Author 2001). Gilbert and Mulkey (1984) discuss the attribution of error, in particular showing that interpreting error is just as theory-laden as interpreting evidence.

There are also numerous historical treatments — especially relevant because they frame the history in terms of its more philosophical “lessons”. Felix Franks (1981) book on *Polywater* is now virtually a classic. As is Stephen Jay Gould's (1981) *Mismeasure of Man*. Lindley Darden (1998) has composed a concise review of major errors by four famous scientists “who did good work but also made honest mistakes.” The work was motivated in part by hearings in Congress on proposed legislation to protect scientists from frivolous lawsuits about their published mistakes (Steinbach 1998). Douglas Allchin (2008, 2009) has followed with two similar essays: one briefly profiling errors by 11 Nobel Prize winners in biology, and a second “celebrating Darwin's errors” on the occasion of his bicentenary. Another collection, *Right for the Wrong Reasons*, explores a series of historical cases, seeking themes about error and method across time and disciplines (Buchwald & Franklin 2005). These papers and volumes certainly do not exhaust the treatment of errors by historians. There are, of course, numerous accounts of individual scientists and specific errors, such as Kottler (1974) on the miscounting of human chromosomes, Nye (1980) on N-rays, or Gould (1985) on Buckland's work on Noah's flood — to name just a few samples. The major treatments cited above, however, are more oriented to the general problem of nature of science as addressed in an educational context.

All these resources may — and perhaps should — inform a complete consideration of teaching about error in science.

Conceptualizing Error: Definition or Structured Inventory?

Kipnis (2010) recently contended that to address error, science educators must begin with a definition. It is a common impulse. One might equally imagine that one cannot do science until it is defined. Of course, people were investigating nature and differentiating between reliable and unreliable conclusions long before there was a formal concept of science. Definitions emerge only late in the process. The rejections of Ptolemaic astronomy, catastrophic worldwide floods and polywater did not hinge on any subtle definition of error.

Error is, ironically, notoriously difficult to *define* unambiguously. One reason is that the concept of error is most often diachronic. If one knew that one was making an error, it would no longer be an error. Such judgments are made in retrospect (or from one “informed” perspective with respect to another, “ill-informed” perspective). In clumsy hands, a definition of error merely reduces to regarding the perspective of the speaker at hand as legitimate: *I* am right and *you (they)* are wrong. Another common tendency (reflected in Kipnis's own definition effort, Section 2) is to refer to the ultimate truth or falsity of a claim — trying to transcend the very

epistemological problem that error presents. With no independent access to reality, however, all knowledge claims must negotiate the challenges of evidence and demonstration through such epistemic strategies as intervention or robustness of data (Hacking 1984; Wimsatt 2007, Chapter 4). An error claim is ultimately a form of *negative* knowledge and requires evidence. An error once ascertained is, paradoxically, a new fact on its own (Author 1999).

In practice, awareness of an error typically precedes any clear articulation of what precisely the error is. Errors first “appear” as a *discordance* of different experimental results or as an *anomaly* between observations and theoretical expectations (Kuhn 1970, Darden 1991, Franklin 2002). The next step is to “isolate” or identify the error. Once the error is confidently articulated, one has, paradoxically perhaps, *discovered* something new. That new conception replaces the former one. Thinking about error fully and in detail can be puzzling indeed.

One could also choose to fuss over what counts as a *scientific* error (for example, Kipnis 2010, Section 2). This is akin to debating the scope of science. Fraud, for example, is often characterized as external to science *proper*, even though it is perpetrated by scientists, is found in the scientific literature, and can affect day to day research activities. It is all too easy, of course, to define science *normatively* rather than *descriptively*. One can try to stipulate that science *is* what scientists ideally *ought* to do. That fits a *political* goal of safeguarding scientific authority. However, the ideal does not always match how scientists reach reliable conclusions *in practice*. An idealized approach inevitably leads to a view that science is error-free and that all scientific error is pathological (see above). While one can adopt this posture, it does not contribute to solving the challenge of interpreting how errors emerge and with what consequences. When scientific conclusions go awry, including in public settings, it is important to understand why, whether one labels it “scientific” or not.

Another tendency is to associate error with blame or accountability (whether political or moral). Error is viewed in terms of the scientist and an implicit standard of appropriate actions or decisions. That is, one might characterize errors in terms of adherence to particular scientific norms (Kipnis 2010, Sections 6-7). This often leads to harsh personal judgments of “impeding” scientific progress (Langmuir 1989). But scientific norms are not absolute. They are heuristics, or guides to epistemic productivity, not inviolable rules (Wimsatt 2007). Ultimately, what matters is the reliability of the conclusions, not some imagined means for ensuring such conclusions. A judgment of *fault* or *responsibility* is independent of (and typically based on) a prior assessment of the reliability of a scientific *claim*. Scientific error relevant to scientific literacy and education is about scientific *knowledge* or *conclusions*, not about *scientists*.

What matters for teaching the nature of science, ultimately, is how scientific claims can fail. One needs, more than any formal or abstract definition, an awareness of the various *error types*. A science teacher needs an *inventory* of possible sources of error, an appreciation of their *variety*, and some scheme to *organize* them. Of course, sources of error parallel methods for avoiding them and for ensuring reliable claims. A scheme of errors is also, indirectly, a scheme for validating scientific knowledge (Author 2001; 2011a, Table 2).

In interpreting error, it is fruitful to conceptualize theories as selective representations, models, or *maps* (Ziman 1978; van Frassen 1980; Turnbull 1989; Giere 1998, 2006). That is, a theory, concept or other claim is a “mapping” of some aspect of the physical world (see Hacking 1984, pp. 208-209, on the complex transformations yielding microscope images; and Latour 1987, pp. 195–257, on the transmission of claims extended through social networks). In the widely used map metaphor, the phenomena are the territory. Scientists develop maps and then maps *of* maps in successive layers. All may be traced to a collection of initial observational

benchmarks (including measurements and other data collection). Each stage in combining or transforming those original observations is a remapping and requires justification. Each is also subject to error. (Again, the structure of justification and the structure of potential error mirror one another.) An error may thus be seen as a faulty mapping that does not preserve the structure of the world as intended (within the context of the map's limitations). That is, when researchers uncover an error, they have identified an alternative mapping (and justification structure) that they deem more deeply warranted by the available evidence.

The concept of mapping can help organize an inventory of error types. Where along the continuum of transformations do they occur? What is the nature of the transformation that failed? The concept of mapping provides a simple organization: from local data to global theories and cultural contexts, from test tubes to YouTube, from the lab bench to the judicial bench. Errors may further be classified or categorized. At the most general level, error types or possible sources of error may be sorted into experimental (material and observational), conceptual and social (see Table 1)(see Author 2001 for fuller discussion and Hon [1989] for an similar framework oriented more narrowly to experimental practice). The experimental errors (such as an uncontrolled experiment or mistaken protocol) and the conceptual errors (such as mistaking correlation for causation, hasty generalization, or confirmation bias) are largely familiar territory for educators. The value of a scheme like this is in indicating how the social system of checks and balances or of communication are also seamlessly integrated into the process of science, a dimension too often peripheralized in science education (Author 2004b). When the system of credibility fails, as in cases of fraud or gender bias, it is as much a part of science as a contaminated DNA sample, an uncalibrated instrument or mistaken assumption.² Scientific justification is built on complex, converging chains of reasoning from the evidence. Ultimately, it is the spectrum of error types and their organization or classification, not definitions, that is important in thinking about and analyzing error.

Teaching Strategies & Resources

Teaching about the role of error and developing skills in error analysis is facilitated by historical cases studies. One can, of course, introduce error analysis in the students' own inquiries. But the scope of error will be limited. It will be hard to profile confirmation bias or the limits of one's own cognitive perspective. In addition, there is a strong affective component to "being wrong," placing extraordinary demands on the teacher in managing student emotions, as well as the intellectual lessons. Consider, for example, one case where emotion eclipsed an opportune lesson in experimental practice. The occasion was a lab for non-majors where students measured the cumulative activity of an enzyme. The student's data reflected a steady increase in reaction product (in a set of parallel samples), except for one time interval, where the level appeared to be zero. Had the product accumulated, then disappeared, then resumed its increase? Obviously not. Reasoning by context, the data point was in error. With the other data mapping a clear trend, one could safely exclude the one "measurement" with no harm to interpreting the results. Here

²The error in fraud is not the original lie, but the mistaken belief in the unannounced lie, along with the implicit trust that ultimately proves unwarranted. Misconduct is different from error: it is a sanctioned behavior (based on a moral judgment), not an unreliable claim (based on an epistemic assessment).

(I imagined) was an exceptional teachable moment for learning about why scientists might throw out bad data (for example, as Millikan did with some of his oil-drop runs; Franklin 1981). Yet the student insisted that the measurement was correct. I offered a plausible explanation: one of the reagents for *measuring* the enzyme had likely not been added to that particular tube — easy enough to imagine with the crowded lab, the rushed time schedule, and the multitude of tubes to manage. This happens to real scientists, I noted: no fault, no blame. The trick was to notice the slip. Here, one could simply drop the unneeded extra data point. To my dismay, the student defended that he had *not* made a “mistake”. Yet the student was unable (upon invitation) to explain the anomalous graph, other than to declare that this was what was “observed”. I marveled, of course, at the complexity of the human mind. Teaching about error in students’ own inquiries is not necessarily easy. Errors of historical figures, by contrast, can be addressed with a sense of emotional distance, while still engaging the student in the intellectual challenge of finding flaws in experimental design or reasoning, and in imagining alternative explanations or tests.

Using history may seem to invite conventional lecture or passive storytelling. However, respecting ideals in active learning, one should rather engage students in historical case studies in a guided inquiry mode (Hagen, Allchin & Singer 1996; Author 2011b). As in all nature of science lessons, students should reflect explicitly on their experience (Akerson, Abd-El-Khalick, & Lederman 2000; Craven 2002; Khishfe & Abd-El-Khalick 2002; Scharmann et al. 2005; Seker & Welsh 2005). Error is an integral part of recreating science-in-the-making (Latour 1987). Classroom narratives can trace a notable scientist down a path of error, with the students following along, perhaps reflecting on the status of the reasoning along the way. Students can *experience* error, shifting from an initially reasonable conclusion to an unexpected revised conclusion. The reasonableness of errors must be felt and internalized. But here the vicarious experience is emotionally less threatening.

By contrast, treating errors as merely foolish or naive, or resulting from credulity or pathological self-delusion, risks alienating students. Admonitions to “be objective!” or “avoid bias!” are thus ultimately ineffective (Evans 2002). They do not plumb the cognitive or social roots of error. Alas, this strategy is all too common in treatments of pseudoscience, fraud and marginal science (for example, Friedlander 1998, Feder 1999, Park 2000, Fritze 2009). Such approaches assume the “right” answer and impose a rationally reconstructed justification, rather than trace a conceptual path from the perspective of an errant believer (Author 1995). The teacher aiming to develop a deep sense of the nature of science, then, will profile and contextualize the concrete sources of error.

Some historical cases that exemplify this approach are already available. For example, one can appreciate how Leonard Nash (1957) integrated error as a matter of course into his history of “Plants and the Atmosphere” (although his text did not engage students in explicit reflection). Errors are featured more prominently in several inquiry cases from the 1996 collection, *Doing Biology* (Hagen, Allchin & Singer 1996). For example, one case study focuses on Christian Eijkman, who shared a Nobel Prize for the discovery of vitamins, but originally claimed that beriberi (a vitamin deficiency) was caused by a bacterial toxin (pp. 116-127). Another addresses Hans Selye’s studies on stress and his proposal of a general adaptation syndrome, which was accepted at first, then abandoned (pp. 104-115). Yet another profiles George Gaylord Simpson, a paleontologist who rejected the role of continental drift, interpreting biogeographical patterns as due to migrations and “sweepstakes routes” (pp. 174-184). These cases all follow the scientist’s reasoning and the evidence that was appealed to in making the

initial errors.

To these one may add several cases from the SHiPS Resource Center. Novak (2008) follows the rejection and later revival of Amodeo Avogadro's hypothesis on gas volumes and weight. The notion was initially overshadowed by the deference afforded to Jacob Berzelius, whose electrochemical theories did not allow diatomic molecules, and by Avogadro's untenable theory of heat. Almost a half-century later, his ideas were revived by Stanislao Cannizzaro in a different context. Leland (2007) invites students into various historical contexts to interpret Native American claims about herbal remedies. Some cures were real, some ineffective, some dangerous, and some (patent medicines) fraudulent. Students must negotiate their way through the tensions between testimony and evidence and reflect on the role of trust across different cultures. Moran (2009) focuses on Richard Lower, who initially accepted William Harvey's claim that the heart was the source of the vital power of the blood (and thus of its bright red color). Lower works through several anomalous findings — and further investigation — to “discover” Harvey's error. Strandemo (2005) addresses the Prout-Berthollet debate on the law of definite proportions. Each had an unyielding view on how elements combine: for one, in any ratio; for the other only in whole-number ratios — implying atomism. Students must discern who is right and who is wrong, and what evidence is needed to make that assessment. Montgomery (2010) provides an expansive view of the debate over glacial theory from 1800 to 1870. Originally, what we now know as glacial erratics and scouring were originally interpreted as evidence of a Noachian flood. Exercises guide students through challenges of historical reasoning, interpreting field observations and the power of personality in accepting Agassiz's theory that glaciers once covered Europe. Finally, Haselberger (forthcoming) addresses Joseph Weber's claims to have detected gravity waves. While the reports provoked substantial research, they could not be replicated. Further investigation helped isolate the sources of error, which have since helped guide subsequent efforts. These cases all concretely illustrate the “tentativeness” of science, showing not only that error occurs, even among highly regarded scientists, but also rendering just how error occurs, and how it is remedied. Again, these are not just stories. They are inquiry cases that engage students in reasoning through the errors and reflecting on how they are part of the nature of science.

Learning Progression on Error in Science

An important approach to teaching science and the nature of science emerging in recent years is the notion of learning progressions — arranging lessons serially to accommodate growing sophistication of understanding. Here, I sketch one prospective sequence of progressively deeper lessons about error and the nature of science.

First, students might learn that scientists, even good scientists, can make mistakes. For example, Galileo is justly celebrated for his *Dialog on the Chief Two World Systems* in advocating a Copernican worldview. Yet the book's central thesis (betrayed in its original title, later censored) was that the tides are caused by the combined daily and annual motions of the Earth — and are physical evidence thereof. But tides are generated by the gravity of the moon and sun. Galileo was wrong. Galileo also developed a law for pendular motion. Yet this presumed universal law is a lie. It approximates results only for small angles. It neglects friction and possible unequal distribution of mass. No wonder that students cannot get their data in the lab to match the expected “universal law”. An idealization is inherently an error, even if a potentially fruitful one (Wimsatt 2007). Galileo's achievements were many. But he was not free

from error. Culture surrounds students with mythic depictions of scientists. Often, the ideal displaces the real. Teachers need to help temper the tendency to romanticize scientists by introducing a healthy dose of respect for error, even among famous scientists. Even at an early age, celebrations of scientists can be coupled with the acknowledgment of flaws. Greatness need not imply perfection. As noted above, teachers need to differentiate between *role models* and *real models*. Conveying the inevitable and acceptable *fact* of error in science is a first step.

Students may then be prepared to learn, second, the “how” of error. Errors are not just random events or lapses in personal judgment. They have identifiable sources. For example, Eijkman’s erroneous view about the cause of beriberi was rooted in prior expectations (confirmation bias). Those perspectives, in turn, were shaped by the recent emergence of germ theory and excitement about its potential to revolutionize medicine, as well as patterns that indicated local contagion. A student can appreciate that anyone in Eijkman’s position would tend to reach the same conclusion. Bächlin’s error in X-ray wavelength (above) had discernible sources in a cascade of earlier measurement errors, originating in someone inadvertently using an incorrect value in a formula. The explosion of work on polywater was ultimately traced to dirty glassware. (Yes, those teachers who insist that students clean up before rushing off to their next class are honoring important methodological norms!) Such cases prime an appreciation of the need to analyze claims for specific errors — not just to adopt a blanket skepticism. This lays a foundation for skills in analyzing claims in public and personal decision making, in a context of scientific literacy. This is where an inventory of error types becomes an important guide for educators.

In the absence of a clear analysis of sources, error is typically attributed — even among scientists — to psychological dispositions or undue social influence. A team of sociologists documented how one group of scientists from a contentious controversy interpreted each other’s “errors.” They were allegedly due, variously, to succumbing to charisma, a rhetorical “aura of fact,” personal rivalry, dislike, and an “ostrich approach” of willfully disregarding the facts. They cited “intellectual inertia” and confrontation with “unorthodox” views. Some researchers were deemed “dogmatic,” others apparently only “tenacious.” Others saw error as due to “prejudice, pig-headedness, strong personality, subjective bias, emotional involvement, naivety, sheer stupidity, thinking in a woolly fashion, fear of losing grants, threats to status and so on.” For one participant, the whole generation was simply “unequal to the task” (Gilbert and Mulkay 1984, pp. 49, 65, 66, 71, 79, 81, 93, 96).³ “Thinking in a woolly fashion”? For the most part, these are not sources of error. They are convenient rationalizations for one’s own position in the face of disagreement or criticism: they portray one’s own view as rational and alternative views as irrational. One might well learn about the inherent cognitive tendency exhibited in such

³ Lest anyone think this list is unrepresentative, one can consider how another author characterized those who believe various ‘pseudohistorical’ claims. They may be: naive, biased, prejudiced, cynical, gullible, indiscriminating, unscrupulous, undisciplined, unorthodox, irrational, spiritual, flawed, fallacious, sensationalistic, amusing, quirky, eccentric, crazy, bizarre and embarrassing, pathetic, off-beat, audacious (or ‘almost unimaginably audacious’), outrageous, rhetorically clever, wild, extremist, over-eager, obsessive, manic, nefarious, reprehensible, and contemptible, not to mention communist and obfuscating (Fritze 2009, pp. 8, 9, 10, 14, 16, 61, 66, 70, 83, 100, 136, 160, 164, 165, 167, 168, 169, 180, 183, 188, 202, 218, 219, 220, 253, 254).

judgments, but some good examples of reasonable error may be needed first as leverage. A deeper appreciation of the concrete sources of error and error types is needed to get beyond such simplistic dichotomies and to engage error in context.

Next, students are ready to consider how errors are found and remedied. In popular conceptions, especially common in the lore of practicing scientists, “science is self-correcting.” Yet this is an aggrandizing myth. Errors do not announce themselves, else they would hardly ever occur. Some errors can persist for decades. The racial and gendered errors of craniology lingered for decades, only to be transformed into errors about IQ testing (Gould 1981). Kelvin’s thermodynamic calculations of the age of the Earth held sway for decades, until the discovery of radioactivity as a source of heat exposed his erroneous assumptions (Hallam 1989, Darden 1998). Avogadro’s hypotheses, too, lay abandoned for half a century (Guinta 2001, Novak 2008). Errors in science do not disappear merely with the passage of time. The belief that “the truth will out” is vague and unacceptable as an explanation (Gilbert & Mulkay 1984, pp. 91-111). Finding errors requires scientific *work*.

Consider, for example, Joseph Priestley’s experiments on plants and the “goodness of air.” Priestley originally found that plants immersed in water yield a gas that helps restore the air for breathing. When he returned to those experiments later, he found, like others, that the effect was not consistent. Eventually, he noticed that light was required. So he pursued simple samples of water exposed to light (with no plants) and found that they, too, yielded the “purer,” more respirable air. He concluded that his earlier claim was mistaken, and that the process was related to light, not plants. Ironically, it was the newly revised conclusion that was in error. Priestley also noticed that his vessels produced a green scum. He erred further in interpreting that as a by-product of the enriched atmosphere. Jan Ingenhousz and others, however, saw a connection between the green scum and green plants. With further microscopic analysis they realized that the scum was living algae. It was the green living matter that transformed the air — but only in the presence of light. They coupled their knowledge of plants as food and fuel to the sun’s light, realizing that there was a further connection. Priestley had mistaken correlation and causation in two ways. To his credit, Priestley ultimately acknowledged his error, once the new explanation had been clearly demonstrated. But the alternative perspective of his peers had been essential to finding and remedying the error. Once again, identifying an error was coupled to a significant discovery, and it required work (Nash 1957).

Errors are encountered and identified in various ways. Often, other scientists try to build on and develop earlier findings, only to find that they cannot do so. That leads to further work “isolating” the error. On other occasions, anomalies emerge by chance in an unrelated project — when someone notices the connection and pursues it. Other times, as in the case of Priestley, errors are profiled through alternative conceptual perspectives. Contrasting views help highlight deficits in the evidence or expose conceptual blind spots. Alternative perspectives — from various disciplines, biographical backgrounds, cultures, social classes, genders, etc. — enhance collective awareness. Ultimately, testing scientific claims against the evidence alone is not sufficient. Responsible claims include engaging potential critics. “Critical thinking” may rely less on being “critical” than on listening well to criticism. Error can be exposed and weeded out. But it frequently involves a system of checks and balances. Through social interaction, scientific knowledge edges forward. In terms of reliability of claims, then, the social dimension of science is just as important as the experimental (Longino 1990, Harding 1991, Solomon 2001). That is an important perspective for educators who would dismiss the social dimension of science because they see it as a source of error rather than as a fundamental mechanism for remedying it

(Finkel 1992, Author 2004b).

The final two “stages” of the prospective learning progression on scientific error involve more subtle and complex lessons, and hence may likely be reserved for more mature students. Both involve blurring the sharp dichotomy of right and wrong, discovery and error, fruitful methods and poor practice. The first of these is to recognize that errors may be productive or fruitful, even while being “wrong” in a sense⁴. The concept of phlogiston, for example, is often denigrated and even ridiculed as one of the most foolish concepts in the history of science. Yet using the concept of phlogiston, chemists predicted that hydrogen (inflammable air) could reduce metals. The concept of phlogiston facilitated discovering the role of light in the chemistry of plant growth. It opened the possibility of one metal reducing another (rather than depending on charcoal). It also stimulated investigation of the chemical properties of electricity, including its ability to reduce metals and acids, and the potential of the galvanic apparatus (including electrolysis). Phlogiston was not so wrong-headed, after all (Partington 1962, v. III, pp. 268-270; Siegfried 1964; Sudduth 1978; Carrier 1991, pp.29-30; Author 1992; Kim 2008; Chang 2009). Similarly, the notion that heat is a substance — caloric — led to the founding of calorimetry and the principle that heated gases all expand at the same rate. A caloric interpretation of the adiabatic expansion of gases was apparently a major part of the reasoning of the eight scientists who can lay claim to discovering the principle of the conservation of energy (Kuhn 1958; Fox 1971, pp. 69-79; Holmes 1985, pp. 160-183; Carrier 1991, pp. 30-31; Levere 2001, pp. 75-77). These are pretty serious examples given conventional School Science. They indicate that the essence of “the scientific method” as it is often taught in schools — predict, test, decide — is an impoverished model. Many theories that have now been abandoned were once empirically successful. Some even made significant novel predictions that were indeed confirmed (Losee 2005). Here may be a lesson for science teachers as much as science students.

The cases of phlogiston and caloric, along with others, may lead one to fuller discussion of models and scientific theories as selective representations or maps (see above). It is paradoxical, indeed, that the fluid model of electricity may prove correct *in a certain domain of phenomena*, even if scientists no longer regard electricity as a fluid. Models may prove useful and ostensibly true (within prescribed limitations), even if they are strictly “wrong” from another perspective. One might even find a place for phlogiston in a standard chemistry curriculum, both to help explain a certain set of phenomena and to open discussion about the nature of scientific theories and error (Scott 1958; Mamlok-Naaman, Ben-Zvi, Hofstein, et al. 2005). For a classroom case study, see Author (1997).

The second deeper lesson on error is that, ironically, in many cases errors can be traced to the very same source that led to great discoveries. What leads to insight on one occasion is a blind spot on another. The potential for discovery and error seem coupled. For example, Darwin relied on Lyell’s principle of uniformitarianism to develop his theory of the origin of coral atolls. It was accepted and helped launch his career. Not long after, Darwin applied the same kind of large-scale gradualist thinking to the parallel roads of Glen Roy. Here, Darwin was wrong. He

⁴Here I refer to “false” ideas that were widely accepted by a scientific community as “true.” One could also comment on the potential of “false models as a means to truer theories” — either when a researcher acknowledges that s/he is making false assumptions as a heuristic or when one incomplete model is modified in yielding the next, rather than jettisoned entirely (Wimsatt 2007, esp. Chapter 4).

later acknowledged his “great blunder.” In another instance, Darwin used a hierarchy of races to map out a theory on the evolution of morality. He saw the natives of Tierra del Fuego as mentally intermediate between orangutans and his peers in the British upper class. That enabled him to envision a transition, and present human evolution as applying to mental powers as well as anatomy. Alfred Russel Wallace, by contrast, was from the middle class and interacted with and depended on the natives in the Malay Archipelago. His positive view of the natives emphasized the differences between humans and other primates — and he never accepted that the human mind might be a product of natural selection (Author 2009). As illustrated in these two cases from Darwin’s life, sometimes the very same concept or way of thinking that leads to discovery also leads to error. Unique viewpoints can foster insight in one context, while blinding one to alternatives in others. Paradoxically, if insights and blind spots are indeed two sides of the same coin — the expression of unique perspectives — then we cannot expect to prevent error in science. Nor would we want to, if we value new ideas. The cost of innovation seems to be the risk of failure. That is a potentially profound lesson, but one that might be reached only after a long series of reflections about error in science.

Conclusion

Pluto is no longer a planet. DDT, cyclamates and Vioxx, all once deemed safe based on scientific study, are now viewed as unsafe, based on subsequent scientific study. Acupuncture, once dismissed as “quackupuncture,” by contrast, is now regarded as effective for pain relief. In all these cases, science initially erred. Yet the revised conclusions are important for the scientifically informed consumer and citizen. If educators want students to understand and accept such changes in scientific consensus, rather than dismiss science itself as arbitrary or unreliable, they must teach the limits of science and the foundations of scientific error. Indeed, the ability to interpret such cases may well be a way to assess knowledge of the nature of science. Such knowledge of the nature of science seems equally important in recognizing that other purported scientific claims — about serial dilution of homeopathic remedies, “worlds in collision,” so called Intelligent Design, or the dismissal of anthropogenic climate change — exhibit critical sources of error, and may be rejected as not meeting appropriate standards of evidence and credible expert testimony. Educators need to acknowledge that vague hand waving about skepticism, objective thinking, or science as “tentative” or “self-correcting” are inadequate for developing functional scientific literacy. Understanding error and potential sources of error (Table 1) is integral to complete science education.

References

- Akerson, V. L., Abd-El-Khalick, F. S., & Lederman, N. G. (2000). The influence of a reflective activity-based approach on elementary teachers' conceptions of the nature of science. *Journal of Research in Science Teaching*, 37, 295–317.
- Allchin, D. (2008). Nobel ideals and noble errors. *American Biology Teacher*, 70, 389-392.
- Allchin, D. (2009). Celebrating Darwin's errors. *American Biology Teacher*, 71, 116-119.
- Author. 1992. Phlogiston after oxygen. *Ambix* 39, 110-116.
- Author. 1995. How *not* to teach history in science. In F. Finley, D. Allchin, D. Rhees and S. Fifield (eds.), *Proceedings, Third International History, Philosophy and Science Teaching Conference* (Minneapolis, MN, 1995), vol. 1, 13-22. Reprinted in *Journal of College Science Teaching* 30(2000), 33-37. Reprinted in *The Pantaneto Forum* 7 (July, 2002).
- Author. 1997. Rekindling phlogiston: From Classroom case study to interdisciplinary relationships. *Science & Education* 6, 473-509.
- Author. 1999. Negative results as positive knowledge, and zeroing in on significant problems. *Marine Ecology Progress Series* 191, 303-305.
- Author. 2001. Error types. *Perspectives on Science* 9, 38-59.
- Author. 2003a. Scientific myth-conceptions. *Science Education* 87, 329-351.
- Author. 2003b. Lawson's shoehorn, or should the philosophy of science be rated 'X'? *Science & Education* 12, 315-329.
- Author. 2004a. Pseudohistory and pseudoscience. *Science & Education* 13, 179-195.
- Author. 2004b. Should the sociology of science be rated X? *Science Education* 88, 934-946.
- Author. 2004c. Error and the nature of science. *ActionBioscience*.
- Author. 2011a. Evaluating knowledge of the nature of (whole) science. *Science Education* 95, 918-942.
- Author. 2011b. The Minnesota Case Study Collection: New historical inquiry cases for nature of science education. *Science & Education*. doi:10.1007/s11191-011-9368-x.
- Bechtel, W. & Richardson, R.C. 1993. *Discovering Complexity: Decomposition and Localization as Strategies in Scientific Research*. Princeton, NJ: Princeton University Press.
- Beveridge, W.I.B. 1957. *The Art of Scientific Investigation*. New York, NY: Vintage Books.
- Boring, E. G. 1954. The nature and history of experimental control. *American Journal of Psychology* 67, 573-589.
- Broad, W. & Wade, N. 1982. *Betrayers of the Truth*. New York: Simon and Schuster.
- Buchwald, J.Z. & Franklin, A. 2005. *Wrong for the Right Reasons*. Dordrecht: Springer.
- Butterfield, H. H. 1959. *The Whig Interpretation of History*. London: G. Bell and Sons.
- Carrier, M. 1991. What is wrong with the miracle argument? *Studies in History and Philosophy of Science* 22, 23-36.
- Chang, H. 2009. We have never been Whiggish (about phlogiston). *Centaurus* 51, 239-264.
- Craven, J.A. 2002. Assessing explicit and tacit conceptions of the nature of science among preservice elementary teachers. *International Journal of Science Education*, 24, 785–802.
- Darden, L. 1991. *Theory Change in Science*. Cambridge, UK: Cambridge University Press.
- Darden, L. 1998. The nature of scientific inquiry. URL: www.philosophy.umd.edu/Faculty/LDarden/sciinq/index.html.

- Dolby, R.G.A. 1996. A theory of pathological science. In *Uncertain Knowledge*, Cambridge: Cambridge University Press, pp. 227-44.
- Evans, J. St B.T. 2002. The influence of prior belief on scientific thinking. In *The Cognitive Basis of Science*, P. Carruthers, S. Stich & M. Siegal, eds. Cambridge, UK: Cambridge University Press, pp. 193-210.
- Feder, K.L. 1999. *Frauds, Myths, and Mysteries*. Mountain View, CA: Mayfield.
- Finkel, E. 1992. Examining school science: Using sociology of science in the study of the science classroom. In *History and Philosophy of Science in Science Education*, S. Hills, ed. Kingston, ON: Queen's University, pp.311-322.
- Fooke-Achterrath, M., Lickfield, K.G. , Reusch, V.M., Jr., Aebi, U. , Tschöpe U. & Menge, B. 1974. Close-to-life preservation of *Staphylococcus aureus* mesosomes for transmission electron microscopy. *Journal of Ultrastructure Research* 49, 270-85.
- Franklin, A. 1981. Millikan's published and unpublished data on oil drops. *Historical Studies in the Physical Sciences* 11, 185-201.
- Franklin, A. 1997. Calibration. *Perspectives on Science* 5, 31-80.
- Franklin, A. 2002. *Selectivity and Discord*. Pittsburgh, PA: Pittsburgh University Press.
- Franks, F. 1981. *Polywater*. Cambridge, MA: MIT Press.
- Friedlander, M.W. 1998. *At the Fringes of Science*. Boulder, CO: Fairview.
- Fritze, R.H. 2009. *Invented Knowledge: False History, Fake Science and Pseudo-religions*. London: Reaktion Books.
- Giere, R. 1998. *Understanding Scientific Reasoning*, 4th ed. New York: Harcourt Brace.
- Giere, R. 2006. *Scientific Perspectivism*. Chicago: University of Chicago Press.
- Gilbert, G.N. & Mulkay, M. 1984. *Opening Pandora's Box: A Sociological Analysis of Scientists' Discourse*. Cambridge, UK: Cambridge University Press.
- Gilovich, T. 1991. *How We Know What Isn't So*. New York, NY: Free Press.
- Gould, S.J. 1981. *The Mismeasure of Man*. New York, NY: W.W. Norton.
- Gould, S.J. 1985. The freezing of Noah. In *The Flamingo's Smile*. New York: W.W. Norton. pp. 114-25.
- Grant, J. 2006. *Discarded Science*. Wisley, UK: Facts, Figures & Fun/Artists' and Photographers' Press.
- Gratzer, W. 2000. *The Undergrowth of Knowledge: Delusion, Self-Deception and Human Frailty*. Oxford, UK: Oxford University Press.
- Guinta, C. 2001. Using history to teach scientific method: The role of errors. *Journal of Chemical Education* 78, 623-627.
- Hacking, I. 1984. *Representing and Intervening*. Cambridge: Cambridge University Press.
- Hacking, I. 1990. *The Taming of Chance*. Cambridge, UK: Cambridge University Press.
- Hagen, J., Allchin, D. & Singer, F. 1996. *Doing Biology*. Glenview, IL: Harper Collins. Reprinted online. URL: doingbiology.net (accessed 14 May 2011).
- Hallam, A. 1989. *Great Geological Controversies*, 2d ed. Oxford, UK: Oxford University Press.
- Hallinan, J.T. 2009. *Why We Make Mistakes*. New York: Broadway Books.
- Harding, S. 1991. *Whose Science? Whose Knowledge?: Thinking from Women's Lives*. Cornell University Press.
- Hasleberger, M. (forthcoming). A ripple in science: Joseph Weber & the detection of gravitational waves. Minneapolis, MN: SHiPS Resource Center.
- Hempel, C.G. 1966. *Philosophy of Natural Science*. Englewood Cliffs, NJ: Prentice Hall.

- HIPST [History and Philosophy in Science Teaching Consortium]. (2008). Quality standards for HIPST. http://hipst.eled.auth.gr/hipst_htm/quality_standards_for_hipst.htm. Accessed 3 January 2011.
- Holmes, F.L. 1985. *Lavoisier and the Chemistry of Life*. Madison, WI: University of Wisconsin Press.
- Holton, G. & Mack, A. (eds). 2005. *Errors: Consequences of Big Mistakes in the Natural and Social Sciences*. *Social Research* 72(#1, Spring).
- Hon, G. 1989. Towards a typology of experimental errors: An epistemological view. *Studies in History and Philosophy of Science* 20, 469-504.
- Hon, G., Schickore, J., & Steinle, F. 2009. *Going Amiss in Experimental Research*. *Boston Studies in the Philosophy of Science*, No. 267. Dordrecht: Springer.
- Huff, D. 1954/1965. *How to Lie with Statistics*. New York, NY: Norton.
- Judson, H.F. 2004. *The Great Betrayal: Fraud in Science*. Houghton Mifflin Harcourt.
- Kelly, J.R. & Parker, S. 1996. *Shocking Science: 5,000 Years of Mishaps and Misunderstandings*. Atlanta, GA: Turner Publishing.
- Khishfe, R. & Abd-El-Khalick, F. 2002. Influence of explicit and reflective versus implicit inquiry-oriented instruction on sixth graders' views of nature of science. *Journal of Research in Science Teaching* 39, 551-578.
- Kim, M.G. 2008. The 'instrumental' reality of phlogiston. *HYLE* 14, 27-51.
- Kipnis, N. 2011. *Science & Education*.
- Kohn, A. 1988. *False Prophets: Fraud and Error in Science and Medicine*. Blackwell.
- Kottler, M.J. 1974. From 48 to 46: Cytological technique, preconceptions, and the counting of human chromosomes. *Bulletin of the History of Medicine* 48, 465-502.
- Kuhn, T.S. 1958. The caloric theory of adiabatic compression. *Isis* 49, 132-140.
- Kuhn, T.S. 1970. *The Structure of Scientific Revolutions*, 2d. ed. Chicago: University of Chicago Press.
- Langmuir, I. 1989. Pathological science. *Physics Today* 42(10), 36-48.
- Lawson, A. 2002. What does Galileo's discovery of Jupiter's moons tell us about the process of scientific discovery? *Science & Education* 11, 1-24.
- Lawson, A. 2003. The nature and development of hypothetico-predictive argumentation with implications for science teaching. *International Journal of Science Education* 25, 1387-1408.
- Lawson, A. 2004. *T. rex*, the crater of doom, and the nature of scientific discovery. *Science & Education* 13, 155-177.
- Lawson, A. 2009.
- Lederman, Wade & Bell. 1998.
- Leland, T. (2007). Interpreting Native American herbal remedies. Minneapolis, MN: SHiPS Resource Center. <http://www1.umn.tc.edu/ships/modules/biol/native-herb.htm>. Accessed 25 April 2011.
- Levere, T.H. 2001. *Transforming Matter*. Baltimore: Johns Hopkins University Press.
- Lilienfeld, A. M. 1982. *Ceteris paribus*: the evolution of the clinical trial. *Bulletin of the History of Medicine* 56, 1-18.
- Longino, H. 1990. *Science as Social Knowledge: Values and Objectivity in Scientific Inquiry*. Princeton, NJ: Princeton University Press.
- Losee, J. 2005. *Theories on the Scrap Heap: Scientists and Philosophers on the Falsification, Rejection and Replacement of Theories*. Pittsburgh, PA: University of Pittsburgh Press.

- Mamlok-Naaman, R. Ben-Zvi, R., Hofstein, A., Menis, J. & Erduran, S. 2005. Learning science through a historical approach: Does it affect the attitudes of non-science-oriented students towards science? *International Journal of Science and Mathematics Education* 3, 485-507.
- Mayo, D. 1996. *Error and the Growth of Experimental Knowledge*. Chicago: University of Chicago Press.
- Mermelstein, E. & Young, K.-C. 1995. The nature of error in the process of science and its implications for the teaching of science. In *Proceedings, Third International History, Philosophy and Science Teaching Conference*, Finley et al. (eds.), pp.768-775. Minneapolis, MN: University of Minnesota.
- Mill, J.S. 1874. *System of Logic*, 8th ed. New York: Harper & Brothers.
- Monmonier, M. 1991. *How to Lie with Maps*. Chicago, IL: University of Chicago Press.
- Montgomery, K. 2010. *Debating glacial theory, 1800-1870*. Minneapolis, MN: SHiPS Resource Center. <http://glacialtheory.net>. Accessed 25 April 2011.
- Moran, E. 2009. Richard Lower and the 'life force' of the body. Minneapolis, MN: SHiPS Resource Center. <http://www1.umn.tc.edu/ships/modules/biol/lower.htm>. Accessed 25 April 2011.
- Nash, L.K. 1951. An historical approach to the teaching of science. *Journal of Chemical Education* , 146-151.
- Nash, L.K. 1957. Plants and the atmosphere. In *Harvard Case Histories in Experimental Science*, J.B. Conant and L.K. Nash, eds., II, pp. 325-436. Cambridge, MA: Harvard University Press.
- Nichols, H.L., Jr. 1984. *Scientific Blundering: An Outsider's View*. Greenwich, CT: North Castle Books.
- Nickerson, R.S. 1998. Confirmation bias: a ubiquitous phenomenon in many guises. *Rev. of General Psychology* 2, 175-220.
- Novak, L. (2008). Determining atomic weights: Amodeo Avogadro & his weight-volume hypothesis. Minneapolis, MN: SHiPS Resource Center. <http://www1.umn.tc.edu/ships/modules/chem/avogadro.htm>. Accessed 25 April 2011.
- Nye, M.J. 1980. N-rays: An episode in the history and psychology of science. *Historical Studies in the Physical Sciences* 11, 125-156.
- Oreskes, N. 1999. *The Rejection of Continental Drift*. New York, NY: Oxford University Press.
- Osborne, J. 2010. Arguing to learn in science: The role of collaborative, critical discourse. *Science* 328, 463-466.
- Oxford English Dictionary*. 1971. Oxford: Oxford University Press.
- Park, R. 2000. *Voodoo Science*. Oxford: Oxford University Press.
- Partington, J.R. 1962. *A History of Chemistry*. London: Macmillan.
- Petroski, H. 1994. *Design Paradigms: Case Histories of Error and Judgment in Engineering*. Cambridge: Cambridge University Press.
- Platt, J.R. 1964. Strong inference. *Science* 146, 347-353.
- Porter, T.M. 1986. *The Rise of Statistical Thinking*. Princeton, NJ: Princeton University Press.
- Rousseau, D. L. 1992. Case studies in pathological science. *American Scientist* 80, 54-63.
- Scharmman, LC., Smith, M.U., James, M.C., & Jensen, M. 2005. Explicit reflective nature of science instruction: evolution, intelligent design, and umbrellaology. *Journal of Science Teacher Education*, 16, 27-41.

- Schulz, K. 2010. *Being Wrong: Adventures in the Margin of Error*. New York Harper Collins.
- Scott, J.H. 1958. Qualitative adequacy of phlogiston. *Journal of Chemical Education* 29, 360-363.
- Seker, H., & Welsh, L. C. (2005). The comparison of explicit and implicit ways of using history of science for students understanding of the nature of science. Eighth International History, Philosophy, Sociology & Science Teaching Conference, Leeds, UK.
- Shermer, M. 2002. *Why People Believe Weird Things*, 2d. ed. New York, NY: W.H. Freeman/Henry Holt.
- Siegfried, R. 1964. The phlogistic conjectures of Humphry Davy. *Chymia* 9, 117–124.
- Solomon, M. 2001. *Social Empiricism*. Cambridge, MA: MIT Press.
- Star, S.L. & Gerson, E.M. 1986. The management and dynamics of anomalies in scientific work. *Sociological Quarterly* 28, 147-169.
- Steinbach, S.E. 1998. How frivolous litigation threatens good science. *Chronicle of Higher Education* (Dec. 4), A56.
- Strandemo, K. 2005. From Lavoisier to Dalton: Debating the law of definite proportions. Minneapolis, MN: SHiPS Resource Center.
<http://www1.umn.tc.edu/ships/modules/chem/proportions.htm>. Accessed 25 April 2011.
- Sudduth, W.M. 1978. Eighteenth-century identifications of electricity with phlogiston. *Ambix* 25, 131-47.
- Suppe, F. 1998. The structure of a scientific paper. *Philosophy of Science* 65, 381-405.
- Sutherland, S. 1992. *Irrationality: Why We Don't Think Straight*. New Brunswick, NJ: Rutgers University Press.
- Turnbull, D. 1989. *Maps are Territories: Science is an Atlas*. Chicago, IL: University of Chicago Press.
- Van Frassen, B. 1980. *The Scientific Image*. New York: Oxford University Press.
- Wandersee, J. H. 1986. Can the history of science help science educators anticipate students' misconceptions? *Journal of Research in Science Teaching* 23, 581-597.
- Wenner, A.M. & Wells, P.H. 1990. *Anatomy of a Controversy*. New York, NY: Columbia University Press.
- Youngson, R.M. 1998. *Scientific Blunders*. New York, NY: Carroll & Graf.
- Zachos, P., Pruzek, R. & Hick, T. 2003. Approaching error in scientific knowledge and science education. In *7th International History, Philosophy of Science and Science Teaching Conference Proceedings*, D. Metz (ed.), Winnipeg: University of Winnipeg, pp. 947-957.
- Ziman, J. 1978. *Reliable Knowledge*. Cambridge: Cambridge University Press.