Error Repertoires

Douglas Allchin
Minnesota Center for the Philosophy of Science

Scientists rely on knowledge of error both in designing experiments and in interpreting results. I profile here the epistemic role of informal catalogs of past mistakes, or *error repertoires* (Mayo 1996). How do scientists develop and use such repertoires and install them in the infrastructure of science? Memory of error, I claim, is critical to progress in science.

The road to wisdom?—Well, it's plain
and simple to express:
Err
and err
and err again
but less
and less
and less.
—Piet Hein

You can tell a good mechanic by what he knows—not about how a car runs, but about how it doesn't run.

1. Introduction. Trial and error is a common metaphor for scientific discovery. The image implies, unfortunately, that scientists grope blindly, only encountering truths haphazardly, learning minimally from their mistakes. For most philosophical commentators on error, error marks failure and burdens science. Taton (1962) expressed this posture well:

> . . . in the very large majority of cases errors of observation, of calculation or of interpretation are *harmful* to scientific research. Mistaken conclusions can often

> be put right only after long and *unproductive* verifications. Furthermore, there are some errors which, having been misunderstood for a long time, *impair or retard the development* of very large fields of science. (p. 92, emphasis added)

Recent concerns about "pathological science" (Langmuir 1989; Rousseau 1992, Dolby 1996) and scientific "blunders" (Youngston 1998) echo this view. For others, pervasive error, construed as the "clumsy antics" of a "lumbering fool," deflates the credibility of science (Collins and Pinch 1993, 2, 151). Wimsatt (1987, 2000) leads others, however, in more productive perspectives— for example, seeing "false models as a means to truer theories." Still, many tend to cast error as a negative, irrelevant product.

By contrast, I contend, discovering and documenting error represents healthy science. While errors may prove personally discouraging, they nonetheless constitute valuable information. Errors expose where otherwise reasonable assumptions or expectations are, in fact, unwarranted or misleading. Such knowledge can accumulate. Researchers are generally keenly aware of their own past error and sometimes those of their colleagues. Moreover, this body of negative knowledge can deepen reliability and effectively guide further research (Allchin 1999b, 2000b).

I begin with two cases studies to demonstrate the pervasiveness and importance of ascertaining error (Sections 2-3 below). Central to this account is Mayo's concept of an *error repertoire*, an ensemble of historical errors:

> a list of mistakes that we would either work to avoid (before-trial planning) or check if committed (after-trial checking). (Mayo 1996, 5)

While one may characterize error repertoires as "negative" knowledge, they nevertheless contribute to the growth and justification of "positive" knowledge. Moreover, I concur with

Mayo (1996) that many methodological rules are techniques for circumventing or managing the errors documented in such repertoires. Error repertoires thereby contribute to escalating standards of evidence and interpretation. Memory of error is critical to progress (Section 4). Finally, I consider the concrete challenges of developing error repertoires and communicating them, especially across successive generations of scientists (Section 5).

2. The Tale of "The Scientist's Apprentice". In Walt Disney's rendering of "The Sorcerer's Apprentice," Mickey Mouse learns the nightmarish consequences of failing to understand potential mistakes. As a novice he suffers from epistemological hubris: assuming that by knowing something important about a topic, he knows all that is important. The master, by contrast, anticipates possible error and knows how to control it. Mickey's ill-fated adventure holds a lesson for scientists seeking reliability in science. In the tale that follows, a scientist's apprentice finds that it is all too easy to make mistakes, although for him, the result is far less catastrophic.

Once upon a time there was a junior physicist named Greg. We find him having just earned his Ph.D. and serving a post-doc at UCSD's prestigious Institute for Pure and Applied Physical Science. Greg is studying high-temperature superconductivity. This recent discovery has his field all abuzz. In particular, his advisor has him investigating superconducting in several three-part compounds: zirconium–, hafnium– and titanium–rubidium phosphide (ZrRuP, HfRuP, TiRuP). The study is relatively routine. Greg varies the temperature of his samples and measures magnetic susceptibility, as an indirect indicator of their superconductivity. All seems to go well and he and the research team report the promising "high" critical temperatures for the zirconium and hafnium compounds, a balmy –260-263EC

(Barz et al 1980, 3133).

Not that Greg and his advisor are unaware of the limitations of their technique. For example, Greg's measurement is susceptible to shielding. One can misinterpret certain surface behaviors of the sample as a "signal" from the bulk core. But no explicit anomalies invite doubts and they proceed confidently. They next enlist a collaborator at Los Alamos to test the same compounds in more sophisticated ways. This should reveal further important properties, while confirming the critical temperatures (T_cs) that Greg's already determined but now via a different method (disturbances in specific heat). Alas, the collaborator reports a much lower T_c for HfRuP: discordant results, a consilience anomaly. Who erred? How? Normally, Greg would now check for shielding: return to his original sample, grind it to a powder and remeasure (finding the error, of course, entails work). But not enough of the specially prepared compound remains. Greg and his advisor review the original data. The more experienced advisor notices a pattern in the x-ray diffraction spectrum that likely indicates impurity phases. Was the hafnium not pure? Had they done a second test with a powdered sample, results there could have also signaled an impurity. Greg commiserates with colleagues in the lab. They kindly cue him to check the manufacturer's label. And lo, he finds that his "99.9% pure" supply of hafnium "may contain 2-3% zirconium." Well, the zirconium would certainly explain the elevated temperature! Now Greg has more work. He orders purer hafnium and makes new samples. Now he finds a T_c more in accord with his collaborator's, who must have started with purer hafnium. Greg, appropriately humbled, acknowledges his modest error as an aside at a conference the following year, retracting the original HfRuP temperature report (Stewart, Meisner and Ku 1982, 332). The error has been "fixed."

But our tale has not yet ended happily. Not before considering the "ever after." Didn't

Greg's error tarnish his credibility and ruin his career? Indeed, it did not. Scientists make such mistakes all the time. No sweat. So long as somebody finds them. Greg's unintended error was "forgiven" in the wake of more reliable information. But this is not the important sequel. Notably, Greg has learned an important class of errors in his field of investigation. He can now anticipate the error. And he knows how, henceforth, to circumvent it. Thus, when Greg first publishes "solo," on a new set of superconducting compounds, he has already checked for the error and even mentions this in his report (Meisner 1981, 763). He has graduated from apprentice to master. His colleagues have heeded the lesson, too (Meisner, Ku and Barz 1983, 984). Everyone has *learned* the error. *That* knowledge (not just finally knowing the correct critical temperature for HfRuP) makes for "happily ever after."

The brief tale of "The Scientist's Apprentice" holds several potential epistemic morals. First, mistakes are common in science. They need not be monumental. And typically, scientists find the errors and remedy the situation, as our hero Greg did. But science is not automatically "self"-correcting. Some errors persist for years: the concepts of heat as a substance and electricity as a fluid, a faulty count of the number of human chromosomes (Kottler 1974), attributing ulcers to stress (Thagard 1999), etc. Greg was lucky, perhaps, that further studies helped him catch his mistake. He also had to spend extra time in the lab ascertaining the nature of the error. Finding and ascertaining error involves epistemic *work* (Allchin 2000a, 2000b, Star and Gerson 1986).

Second, scientists can learn from errors. No one intends to err, surely. Still, when one errs unexpectedly, one gains knowledge that can be used later. Scientists learn to recognize potential error in similar cases and know to avoid or accommodate them. For example, Greg understands the consequences of impure samples. He has not succumbed to this error again.

It's on his mental checklist. It's part of his growing *repertoire* of what errors may occur and how each appears. Hereafter, it will guide his experimental design and practice. Though Greg abandoned one early T_c measurement (he did not care about a compound mixing 97% hafnium and 3% zirconium), he did not discard the knowledge of the error. Indeed, it became composted into an important epistemic tool, or standard, for subsequent research (also see Section 4). The value of this type of experience was evident in his advisor's ability to "read" traces of impurities in the x-ray diffraction spectrum, as well as Greg's colleagues' familiarity with their supplier's materials. Without this knowledge, Greg would likely have invested considerably more effort isolating and identifying the error. This modest episode thus underscores the epistemic importance of scientists building and applying an *error repertoire*.

As a further example, consider the various types of aberrations that threaten the reliability of images from light microscopes (Hacking 1984, 193-94). Any researcher using a microscope needs to be aware of such pitfalls, lest they succumb to an errorenous interpretation. Knowledge of these aberrations—as error repertoire—serves as a check for any result. Is the observed feature due to chromatic aberration? —No. Spherical aberration? —No. Etc. Reliable claims rely on negative responses to all these questions. Knowledge of these aberrations constitutes an error repertoire—and a familiar foundation for reliability among microscopists.

Third, while philosophers enjoy discussing knowledge as abstractly generated and permanent, Greg's experience shows how knowledge must be partly reinstantiated again in each new scientist. At some point, Greg had to learn about error for himself, whether by explicit instruction, guidance from more experienced colleagues, or personal experience. This case underscores, therefore, the role of including knowledge of error in learning science (and perhaps learning about the nature of science) (Section 5).

The error in this episode of "The Scientist's Apprentice" was hardly earth-shattering. But partly for this reason, it may be all the more telling of normal science. At the same time, an error from impure materials need not be minor. Sweat-contaminated water led to over a decade of research on "polywater," construed as a polymerized form of water (Rousseau 1992). Nor is error limited to apprenticing or emerging scientists. Even Nobel prize winners can err (Darden 1998, Allchin 1999a, 2000a). In all cases, the pattern of encountering errors, characterizing them and adding them to a growing error repertoire is the same. Through such knowledge of error, all scientists great and small deepen knowledge claims and navigate through the shoals of misinterpreting results.

3. Deploying Error Repertories: Jean Perrin, Master of Error. Error repertoires are not simply archived failure. Errors are not sequestered like obsolete theory in some remote corner to grow dusty. Rather, they help guide effective research. They indicate appropriate (or inappropriate) procedures and offer caveats to interpretation. Sometimes, researchers need to address specific mistakes noted in earlier studies. Other times, past errors serve as models or exemplars to inform a more general consideration of error (Allchin 2000c). To understand the active role of error repertoires and of anticipating error, consider Jean Perrin's classic work on Brownian motion (Mayo 1996, Chap. 7). Perrin's experimental study was fraught with epistemological peril (pp. 217-20). Yet Perrin skillfully negotiated his way through an interpretive minefield by anticipating and addressing each potential error.[1] Indeed, his handling

[1]Here, the image of the negotiation metaphor is very different from that used by many social constructivists. In their image, negotiation is mere politics. In the alternate image, negotiation is a

of these risks contributes to the perceived elegance of his study.

Perrin wanted to determine whether molecules in an ideal gas moved randomly, according to the molecular-kinetic theory of gases (from there, he could also assess a theoretically derived value for Avogadro's number). He explored this experimentally by simulating the gas with a fluid and observing the displacement (or Brownian motion) of particles suspended in it. The first locus of potential error, therefore, involved whether particles in a fluid mimic the relevant behavior of gas molecules: would he observe the "correct" (intended) phenomenon? To avoid error, Perrin needed particles of uniform radius. By choosing gamboge (a natural latex) as a material, he could attain spheres with very precise size using fractional centrifugation. This "simple" task to avoid one particular error took several months (pp. 236-37). The behavior he wished to monitor also needed to take place in an "unlimited fluid." No one could guarantee that the behavior of grains close the edge of a container would act appropriately; Perrin thus invalidated outright any series of observations where particles drifted into this questionable boundary zone (pp. 237-38). Perrin's experimental model thereby controlled the first and most basic error: not observing the right phenomenon.

One also needs to adopt an appropriate vantage point or method of observing to detect the desired dimensions of a phenomenon. Others before Perrin had tried to measure Brownian motion. But they had focused on the velocity of the particles. Because one can never ascertain the actual path-distance traveled in Brownian motion, their calculations included ineliminable

---

process of carefully finding an effective path to a reliable conclusion (*Oxford English Dictionary*, defs. 2, 4). Such negotiation may well be achieved in part through social interaction, or epistemic checks and balances at the communal level. The diversity of conceptual perspectives available in a scientific community, for example, increases the chances of detecting and isolating error (Harding 1991).

uncertainty. No one could even assess the scope of the potential error adequately. Perrin heeded the past error. Inspired by a recent theoretical development, he measured linear displacement, instead of velocity. One more source of error, encountered earlier, was avoided.

Perrin observed the motion of his gamboge grains, marking their position at regular time intervals. Still, as any experimentalist knows (and might well fear), some feature of the experimental setup may alter or distort the phenomenon being observed. Plenty of factors could concern Perrin, and he checked each of them. Would the viscosity of the fluid affect the motion? He used glycerine, urea and sugar solutions whose viscosity ranged from 1 to 125 (pp. 233-34). Would the size, or mass, of the grains matter (even if uniform)? Perrin tried six grain sizes, from 0.367 to 5.5 microns (p. 234). Would variations in temperature alter the results? What about evaporation of the fluid, vibrations, the composition of the grains, or impurities in the fluid? Here, Perrin explicitly drew on previous experimental experience to discount these factors, though he checked them for himself once again (pp. 240-41). Perrin could confidently rule out error, or artifact, due to the experimental conditions.

One inevitable experimental factor is the observer himself. Could Perrin have biased the results by unconsciously selecting only certain types of grains? He established a procedural rule to preclude any potential error (whether it existed or not): the observer would not choose, but instead follow the first grain that drifted into the center field of the viewing scope. Error based on observer bias could be excluded, as well.

Next, to evaluate the motion for randomness, Perrin had to transform the recorded particle positions into a set of measurements that could be analyzed mathematically. That is, he had to model the data in a form to compare with a probability distribution. The manipulations were transparent, and the mistakes would be primarily arithmetic. Still, Perrin knew that such errors could occur (and sometimes had occurred), especially in selecting misleading measurements. He elected to model the same data in several different ways, allowing them to crosscheck each other (pp. 228-29). He thereby addressed known pitfalls in modeling data.

When Perrin turned to calculate Avogadro's number from these same observations, he relied on experience to manage yet another set of errors. Here, he ensured that he had both measured the size of the grains properly (pp. 237-38) and derived a theoretical value accurately. Virtually every particular of Perrin's experiment was determined by potential errors. And for good reason: if *any* of these errors had occurred, Perrin's conclusions could have unraveled. For guarding against potential mistakes (in contrast to Greg in his formative stage), Perrin was a master of error.

Coping with error is a hallmark of some of the most highly regarded work in science. Consider Ignaz Semmelweis's (1861) revered study of childbed fever in a Viennese hospital. Much of his now classic treatise recounts his tests that "failed" to show what caused the disease: position of delivery, rough handling by inexperienced interns, priests walking through the maternity ward, etc. He thereby succeeded in building a stronger argument, by showing that proposed alternative explanations did not hold. Robert Millikan's renowned "oil-drop experiment," Franklin (1996) has shown, was governed by careful attention to experimental conditions that would otherwise have led to error: room temperature (hence, convection currents), large drops, nonuniform electric field, dust particles, asymmetrical drops, pressure, voltage irregularities, cumulative measurement uncertainties, etc. Likewise, Bernard Kettlewell, managed error in his classic study of peppered moths. He explored specific ways his conclusions about the selective predation of dark and light moths could err: migration of moths from the study area, biased recapture methods, differential longevity (Rudge 1999, 19-20). In supplemental studies, Kettlewell also addressed possible error in his assumptions about how conspicuous moths were and how this affected predation by birds. Other scientists praise all these studies and parade them as exemplary investigations, I contend, largely due to their masterful handling of error.

Suppe (1998) recently analyzed in detail how scientists structure their arguments in published papers. One striking feature, he found (not addressed in most other philosophical accounts), was how authors regularly consider specific doubts or alternative interpretations of data (questions, Q) and then impeach them (through rejoinders, R). Addressing and dismissing potential error is thus a critical element in scientific discourse and thinking. I claim that this process of ruling out error is critical for deepening reliability—that is, for filling the familiar gap between verification and ultimate reliability (Mayo 1996, 174-213, Allchin 2000b). An essential skill of an effective scientist, then, is being able to anticipate and manage error, as vividly epitomized in Perrin's work (and many other "classic" studies). One aphorism suggests that good auto mechanics are marked by what they know about cars—not about how they run, but about all the ways they *don't* run. Similarly, one might characterize skilled scientists in part by the depth of their knowledge of how experiments, interpretations or reasoning can *err*.

4. From Error Repertoires to Methodological Heuristics: How Standards Escalate. As profiled in the cases above, error repertoires serve to warn of specific mistakes. On other occasions, however, error repertoires may be distilled into more general methodological rules or heuristics (Mayo 1996, 18, 452-53). Consider, for example, the now standard practice of double-blind drug studies. Here, one error is mistaking a psychologically mediated effect (based on belief in a purported remedy) for a physiological effect of the medication in question.

Investigators therefore test two groups: one group receives the drug, while a parallel group receives an inert placebo. The patients are "blind" to whether they actually receive a drug. Comparing the two groups allows the experimenter to cancel out, or nullify, the placebo effect as a potential source of error. Other errors can occur when the investigator (who knows which treatment is real and which is placebo) either inadvertently cues patients or unconsciously biases the evaluation of patients' responses. The epistemic remedy is to construct a system where the investigator, too, remains blind to the treatment versus placebo groups. These elements of drug-study design are so commonplace now that their origin in past error may escape notice. Such methodological principles seem to have a priori status. Yet they are steeped in the contingencies of how human brains function. Psychosomatic healing and observer bias are particular facts of human biology. The discovery of how they can subvert drug studies has a history. The justification for compensating strategies, while philosophically sound, was an afterthought. Blind and double-blind studies, then, essentially embody concrete experience with errors. An error repertoire has been transformed into a methodological principle.

Many common methodological norms derive historically from particular types of error. Consider controlled experiments. Logically, they fit a structure appropriate to inferences about the relevance of individual causal factors. However, the awareness of the need for controls was based originally on mistakes about causal factors. Scientists first noticed cases where inference based on only demonstrative instances misled us (that is, where experiments were *not* controlled). Again, we had to discover historically the limits or deficits of such reasoning. Such knowledge does not come "naturally." Indeed, we need to instruct each new generation in this pattern of reasoning. An experimental control, while supporting reasoning towards a positive conclusion, is fundamentally a mechanism for regulating error. Null statistical models, likewise, allow one to guard against mistaking a chance for a real effect (avoiding sampling error; Mayo 1996).

Most methodological norms distill an error repertoire into a simple practical guide, coupling "*do*"s and "*don't*"s. The prevalence of such experimental and inferential norms, or heuristics—so widespread and ordinary that they have become virtually invisible—should alert phiosophers to the importance of regulating error. Avoidance of error, as much as the search for truth, guides the course of scientific investigation.

Error repertoires thus contribute to the development of reliable knowledge in at least two fundamental ways. First, error repertoires constitute knowledge, albeit (as noted earlier) "negative" knowledge. They are products of research. Knowledge of different errors and their nature involves experimental justification; the knowledge is reliable (Allchin 2000b). We know that electricity is not a fluid, though under some conditions it seems so. We know what generates bacterial mesosomes, polywater, apparent gravity-wave signals, etc. Knowing about specific errors and error types can be valuable, as illustrated in the two case studies (Sections 2-3). One should not discount the significance of error repertoires as knowledge. Accordingly, such knowledge (like any knowledge) should be documented, registered in institutionalized memory, and effectively distributed (see Section 5 below).

Second, as error repertoires are transformed into methodological rules, standards of evidence escalate. Some of what would once have passed as evidence becomes viewed as error. Results must meet more stringent conditions to be considered reliable. Hence, where error repertoires grow, the reliability of the corresponding "positive" conclusions can also increase, or deepen. At the same time, content will almost certainly be lost, as some claims fail to measure up to the new standards. Ever since Kuhn's *Structure of Scientific Revolutions*, it has been fashionable to deny progress in science. But this view disregards how error repertoires contribute to more rigorous epistemic standards. Thus, even if one contends that "positive" knowledge content does not uniformly accumulate, the overall quality or reliability of knowledge, filtered through escalating standards, may very well rise. This all depends, of course, on effective memory of error (Section 5). Error repertoires, as forms of local memory, thereby contribute indirectly to the development of even "positive" knowledge.

Taton's criticism of error (noted in Section 1) epitomizes a view that science can, ideally, be error-free. In a more pragmatic view, error is not entirely eliminable. Scientists will continue to encounter new forms of error, at least as they develop new instruments and explore new topics. However, they might limit *repeating* errors. This is the function of error repertoires. Error repertoires, and their corresponding methodological norms, may thus be an important mechanism of progress in science.

As encapsulated in Piet Hein's poem, knowledge of error accumulates: scientists "err/and err/and err again." By documenting their errors in error repertoires they can also use history to guide further research. They potentially avoid succumbing to the same error again, while narrowing the scope or shifting the focus of their trials. Hence, they should err "less/and less/and less." Trial and error in science is not wholly blind or haphazard. By using error repertoires (and the methodological heuristics derived from them) the process becomes, rather, trial and *learn*—and, hence, progressive.

5. Instantiating Error Repertoires in Practice. The abstract concept of an error repertoire as a way to deepen reliability is simple enough. But such a conceptual mechanism must also be instantiated in practice. Epistemic norms have social or institutional dimensions, as well (e.g., Hull 1988, Goldman 1999). As noted earlier, to be effective, error repertoires depend critically on institutional memory. The challenge, then, is really no different than for preserving, communicating and applying any knowledge. But knowledge of error, viewed as "negative" knowledge or even non-knowledge, may not always circulate beyond the lab where it originates. The analysis above suggests, by contrast, that scientists might profit from deeper appreciation of error repertoires and the dynamics of error at the institutional level.

One factor is whether error is published and registered along with fact in the conventional archive of journals and research reports. I do not refer, of course, to results that are unreliable because of some specific, identified error. Rather, one concern is whether negative results and new sources of error, when they occur and have been adequately validated, are published (Bowman 1999, Gould 1995). I contend that such results deserve public note under certain circumstances. Underwood (1999) recommends two operational criteria: (1) one must clearly articulate the study design so that someone else can clearly interpret the meaning of negative results, and (2) one must discuss the statistical power of the study or other bases for validating negative conclusions. I concur that focusing on the power of any statistical analysis is essential (Mayo 1996). This characterizes the ability of the given data to discern, or discriminate, a difference statistically. For example, if measurements are imprecise (measurement uncertainty is high), or sample size is low (sampling uncertainty is high), the results may be unable to *resolve* important differences (Allchin 2000b). In general, a result should be clear and relevant, whether the conclusion is positive or negative. Ideally, an experimenter designs the study in advance such that any result will be informative. This is less likely when the investigation is more exploratory, and the scope of possible results is uncertain.

A basic heuristic for deciding whether to publish some results, then, might be how much they rule out, or how productively they may limit continued search: are other researchers liable to encounter this potential error or need to manage it? Does the error fundamentally reorient assumptions in the field? As in all scientific publication, novel results are valued: confirming known error, without characterizing it more deeply (or from a different perspective), is only weakly valuable. Journals now seem to publish deeply significant negative results. "Notable" error is recognized. Concepts and practice do thereby change. One may wonder, however, whether more modest negative results and errors are adequately documented in the literature. Are the informal mechanisms of communication among scientists adequate in developing communal error repertoires?

Instantiating knowledge of error also extends to the level of individual practitioners. Institutional memory must be properly distributed to each relevant researcher. The constant turnover of practicing scientists makes this a challenge. As exemplified in the case of Greg, the scientist's apprentice, each successive generation must learn again what it already generally known. It can be important to distinguish between what "is known" and who knows it (Goldman 1999). One major feature of the current apprenticing system in science (illustrated in Section 2) is learning field-specific error repertoires. Some labs regularly check for error at all levels, and newcomers simply adopt the habit by immersion. Still, the quality of error education seems to vary widely from lab to lab and is usually implicit only. Such learning should perhaps be more systematic and explicit, as is often found in engineering education (e.g., Petroski 1994; Rabins 2000). One strategy is to teach error repertoires through case studies. Another is to survey several *error types* (Allchin 2000c). Knowledge of how scientists err is also important for public understanding of the nature of science, especially in shaping effective business and policy decisions. Science education even for non-scientists should thus include models or case studies of error in science (Allchin 2000d, 2001).

Error is an integral feature of science. I hope to have profiled its positive role by articulating the concept of an error repertoire and describing how scientists use, or apply, knowledge of error, especially in the form of methodological norms. The challenge remains to educate scientists and non-scientists about error.

Acknowledgements. My deep appreciation to Greg Meisener and Deborah Mayo. Work was supported in part by NEH #FS-23146 and NSF #SES-0096073.

References

Allchin, Douglas, (1999a), "To Err and Win a Nobel Prize: Paul Boyer and ATPase", International Society for the History, Philosophy and Social Studies of Biology Meetings (Oaxaca, July, 1999).

Allchin, Douglas, (1999b), "Negative results as positive knowledge. *Marine Ecology Progress Series* 191:303-305.

Allchin, Douglas, (2000a), "To Err is Science", American Association for the Advancement of Science, Washington, DC, Feb. 21, 2000.

Allchin, Douglas, (2000b), "The Epistemology of Error", Philosophy of Science Association Meetings (Vancouver, B.C., November, 2000).

Allchin, Douglas, (2000c), "Error Types", ms.

Allchin, Douglas, (2000d), "How *Not* to Teach Historical Cases in Science", *Journal of College Science Teaching* 30: 33-37.

Allchin, D. 2001 (forthcoming). Kettlewell's missing evidence, a study in black and white.

Journal of College Science Teaching.

Barz, H., H.C. Ku, G.P. Meisner, Z. Fisk and B.T. Matthias, (1980), "Ternary Transition Metal Phosphides: High-temperature Superconductors", *Proc. Nat. Acad. Sci. USA* 77: 3132-3134.

Bowman, Howard (ed.), (1999), "Negative Results", *Marine Ecology Progress Series* 191: 301-309.

Collins, Harry and Trevor Pinch, (1993), *The Golem*, Cambridge: Cambridge University Press.

Dolby, R.G.A., (1996), "A Theory of the Pathologies of Science", in *Uncertain Knowledge*, Cambridge: Cambridge University Press, pp. 227-44.

Franklin, Allan, (1986), *The Neglect of Experiment*, Cambridge: Cambridge University Press.

Goldman, Alvin, (1999), *Knowledge in a Social World*, Oxford: Clarendon Press.

Gould, Stephen J., (1995), "Cordelia's Dilemma", in *Dinosaur in a Haystack*, New York: Harmony Books, pp. 123-32.

Hacking, Ian, (1984), *Representing and Intervening*, Cambridge: Cambridge University Press.

Harding, Sandra, (1991), *Whose Science? Whose Knowledge?*, Ithaca, NY: Cornell University Press.

Hull, David, (1988), *Science as a Process*, Chicago: University of Chicago Press.

Kottler, Malcolm Jay, (1974), "From 48 to 46: Cytological Technique, Preconception, and the Counting of Human Chromosomes", *Bulletin of the History of Medicine* 48:465-502.

Langmuir, Irving, (1989), "Pathological Science", *Physics Today* (October): 36-48.

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

Meisner, G. P., (1981), "Superconductivity and Magnetic Order in Ternary Rare Earth Transition Metal Phosphides", *Physica* 108B: 763-64.

Meisner, G.P., H.C. Ku and H. Barz, (1983), "Superconducting Equiatomic Ternary Transition Metal Arsenides", *Materials Research Bulletin* 18: 983-91.

Petroski, Henry, (1994), *Design Paradigms: Case Histories of Error and Judgment in Engineering*, Cambridge: Cambridge University Press.

Rabins, Michael J. (ed.), (2000), "Engineering Ethics", URL: lowery.tamu.edu/ethics/ (August 20, 2000; accessed December 11, 2000).

Rousseau, Denis, (1992), "Case Studies in Pathological Science", *American Scientist* 80: 54-63.

Rudge, David Wyss, (1999), "Taking the Peppered Moth with a Grain of Salt", *Biology and Philosophy* 19: 9-37.

Semmelweis, Ignaz, ([1861] 1983), *The Etiology, Concept, and Prophylaxis of Childbed Fever*, Trans. and with an introduction by K.C. Carter. Madison: University of Wisconsin Press.

Star, Susan Leigh and Elihu Gerson, (1986), "The Management and Dynamics of Anomalies in Scientific Work", *Sociological Quarterly* 28: 147-169.

Stewart, G.R., G.P. Meisner and H.C. Ku, (1982), "Specific Heats of the New High T_c Phosphide Superconductors HfRuP and ZrRuP", in *Superconductivity in d- and f-Band Metals*. Kernforshungszentrum Karlsruhe, pp. 331-335.

Suppe, Frederick, (1998), "The Structure of a Scientific Paper", *Philosophy of Science* 65: 381-405.

Taton, René, (1962), *Reason and Chance in Scientific Discovery*, New York: Science Editions.

Thagard, Paul, (1999), *How Scientists Explain Disease*, Princeton: Princeton University Press.

Underwood, A.J., (1999), "Publication of So-called 'Negative' Results in Marine Ecology", *Marine Ecology Progress Series* 191:307-309.

Wimsatt, William C., (1987), "False Models as a Means to Truer Theories", in M. Nitecki (ed.), *Neutral Models in Biology*, Oxford: Oxford University Press.

Wimsatt, William C., (2000), "The Use of False Models to Elaborate Constraints on Processes: the Case of Blending Inheritance", Philosophy of Science Association Meetings (Vancouver, B.C., November, 2000).

Younston, Robert, (1998), *Scientific Blunders*, New York: Carroll and Graf.