Science & Education (2006) 15:113–120
DOI 10.1007/s11191-005-8922-9

Discussion

Lawson's Shoehorn, Reprise

DOUGLAS ALLCHIN
Program in the History of Science and Technology and Minnesota Center for the Philosophy of Science, University of Minnesota, Minneapolis, MN 55455, USA (E-mail: allch001@umn.edu)

Abstract. Lawson's (Lawson, A.: 2004, Science & Education, **13**, 155–177) analysis of the meteorite hypothesis of dinosaur extinction exhibits flaws similar to his earlier (2002) analysis of Galileo's discovery of Jupiter's moons (Allchin, D.: 2003, Science & Education, **12**, 315–329).

Key words: Analogy, contingency, dinosaur extinction, hypothetico-predictive reasoning, scientific discovery

Lawson's (2003) "*T. rex*, the Crater of Doom, and the Nature of Scientific Discovery" parallels his earlier paper in this same journal (2002). While dealing with another case history (Alvarez's meteorite impact hypothesis on dinosaur extinction instead of Galileo's discovery of the moons of Jupiter), the primary theme is still to argue from case to the general nature of scientific discovery. To the degree that the framework and strategy of the paper are a virtual clone of the earlier work,[1] general criticisms of *the failure to respect history* in the earlier work (Allchin 2003) apply here again. In addition, numerous basic flaws and assumptions make the paper's broad conclusions unwarranted.

Most important, the paper does not use history to discern the nature of scientific discovery, but rather imports a philosophical model, aiming to 'characterize' a particular case (pp. 156–157) and, worse, *to* '*reconstruct*' an argument just along those lines (pp. 163, 173) (see Allchin 2004, pp. 180–186, on appropriating history). While many episodes *can* be reconstructed to fit a particular model of science, such exercises do not provide support for the model (Allchin 1995a; 2003, p. 318). Such theory-laden reconstructions hardly test the model, because the "evidence" is generated to fit the model: the reasoning from hypothesis to "evidence" back to hypothesis again is plainly circular (pp. 167, 172).

The fundamental problem with the reconstructive approach, as with the basic hypothetico-predictive (HP) model the paper endorses, is that positive evidence can be taken at face value. The possibility of error remains (e.g., Mayo 1996, on severe tests). Thus, philosophically, scientific discovery may include several relevant elements not captured in the HP model. Here, Lawson's own account of Alvarez's case can expose the very weaknesses in his characterization. Consider how the narrative of Alvarez's research presents, then disregards several significant elements (all italics added):

(1) 'they *unexpectedly* discovered ...' (p. 157)
'... their investigation *led to another surprising observation*' (p. 157)
'But as *luck* would have it ...' (p. 161)
'At each outcrop, they *noticed* ...' (p. 158)

How can investigations 'lead to' unintended observations, if such observations are prescriptively deemed irrelevant? How are such 'unexpected' observations conceptually distinct from observations based on nosing around or probing a particular phenomenon, 'not knowing where to look' (p. 171)? Why are some observations 'surprising', in contrast to merely failing to meet theoretical predictions? Why do some investigators 'notice' them and consider them surprising, while others do not? (Judson 1981, Ch. 4) Where does 'luck' fit (– especially if the process of scientific discovery is characterized in 7 discrete steps; pp. 169–170)?

(2) 'Alvarez was *well aware* of meteor impact craters on Earth' (p. 158)
'Alvarez was also *aware of* two papers in which the authors proposed ...' (p. 158)
'Alvarez *recalled reading* about the 1883 explosion of an Indonesian volcano' (p. 160)
'He *remembered reading* ...' (p. 168)
'thanks in part to *information provided in a talk* ...' (p. 158)
'... *a previously ignored*, but enormous, circular pattern of gravity anomalies' (p. 161)
'Importantly, core samples *previously extracted* ...' (p. 161)
'The Mexican geologists *had seen the* bubbles *before* ...' (p. 162)
'He *went to the library in search of* other KT sites.' (p. 159)

How does prior knowledge become relevant in research developed on an independent trajectory? How does someone become 'aware of' things if they are only to focus on evidence resulting from an HP-directed investigation? How does an investigator come to realize that such other information is relevant (– especially if 'looking for clues' is not allowed, as claimed on p. 171)? What constitutes relevant memory, or is it precisely that such relevance cannot be articulated in advance? Do scientists with a broad liberal arts education (rather than narrow technical expertise) have a wider knowledge base on which to draw in interpreting unique or unexpected events? How does one cultivate a habit of thinking across fields or searching for relevant information from unexpected sources? How could data that can help unknown future investigations be collected and documented, if relevant information is determined solely by its ability to either confirm or disconfirm an earlier hypothesis? What is the role of scientific communication and meetings between scientists, especially in sharing different background knowledge and experience? How is the scientific literature structured to allow effective searching?

(3) 'thanks to *the analytical skills* of Berkeley chemist...' (p. 158)

What role does technical skill play in scientific discovery? Presumably, without this chemist's skilled analysis, the Alvarez team would have been stymied, at least along this research trajectory. Why do intellectual skills seem more valued than manual or craft skills? (Shapin 1989)

(4) 'Therefore, they *came up with a way to* indirectly measure the clay's deposition rate' (p. 158)
'It took several months of *generating and rejecting* possible mechanisms' (p. 160)

How, indeed, do new ideas originate? How do the 'if's in 'if-then' thinking develop? If analogical thinking is central, why peripheralize it by highlighting just prediction and testing (the HP method)? For example, why does the use and basis of analogy not get included in the summaries of the scientific reasoning (pp. 164–166)? Why are some ideas 'rejected' even before they are tested? How does one measure the plausibility of a hypothesis?

(5) 'a Princeton geologist who disputed nearly every *interpretation of the evidence* ...' (p. 162)

What does it mean to '*interpret*' evidence? Perhaps evidence is not plain enough on its own? What factors are relevant to interpretation, such that scientists may legitimately disagree? Is a discovery individual, or does it also involve acceptance by a community of qualified peers?

(6) 'there were always *serious questions* and nagging *doubts*' (p. 162)
'a *major challenge* came from a Princeton geologist who *disputed* ... ' (p. 162)
"an *attempt ... to replicate* the finding turned up empty" (p. 160)

What is the role of skepticism and/or criticism in the growth of knowledge? How do critics contribute to the development of evidence and the revision and refinement of theoretical explanations? (For this case, see Glen 1994). How do checks and balances serve a scientific community? Why should they be needed at all if interpreting evidence is unequivocable? Is a discovery complete if warranted alternative explanations persist?

Each element noted above marks a critical factor in the history of Alvarez's discovery. That is, without the elements of luck, noticing, prior knowledge, interpretation of evidence and new ideas, it is not clear how the history would have proceeded, nor how the discovery would have occurred (or even if events would have led to a discovery at all). This is how one may define a feature of the process of scientific discovery: something which, had it been different or missing, would have changed the outcome historically. Indeed, based on this historical case alone, the following skills would seem 'essential' to the nature of *discovery*, or the assembly of new knowledge:

- encountering novel phenomena
- perceiving their relevance to other information
- generating new patterns or explanations
- capitalizing on opportune observations
- applying experimental skills
- interpreting evidence effectively
- addressing criticism.

Science education would be greatly enriched by teaching these in lessons about the nature of science and scientific discovery. By contrast, characterizing science as alternating simply between bouts of speculation and testing is simplistic and uninformative – and certainly misleading (even if it is not demonstrably false). Alvarez's research obviously passed several empirical and conceptual benchmarks. *Yes*, hypotheses were *included*. *Yes*, tests were *included*. The question of process, however, is *how they were connected historically, or causally*. Other accounts of this same episode offer a far more robust and complex view of the *process* of scientific discovery (Raup 1992; Glen 1994; Powell 1999). Here, Lawson's analysis ultimately omits a wealth of information about how science *happens*. Historical facts have been selected and interpreted to fit a philosophical preconception: Lawson's shoehorn (Allchin 2003) strikes again.

Note further that the author's primary authority on nature of science (first sentence) is a science educator, not a philosopher of science. The claims throughout this paper are uninformed by work in the philosophy of science from the last several decades, which addresses (at least) the subtleties of experimental reasoning (e.g., Hacking 1984; Franklin 1986; Galison 1987; Rheinberger 1997) and the complex interplay between logic and observation (e.g., Hanson 1958; Kuhn 1970; Darden 1991, Bechtel and Richardson 1994; also see review by Allchin (1995b), in this journal). No such work is applied or cited. The paper does not integrate insights from professional philosophy of science into education, but rather fosters educators' own parochial "philosophy of school science" (e.g., note the further appeals on p. 170 to educators Lewis and Moore as presumed philosophical authorities). As I understand it, honoring such standards is fundamental to furthering the mission of this journal.

For the author, the whole process of discovery seems to begin with an *irreducible* "puzzling observation" (pp. 156, 163, 168–170, 172, 175). Often this reflects the very essence of the discovery itself: the awareness of something novel, only to be articulated later. Accordingly, the process seems to rely critically on "unexpected" and "surprising" events (pp. 157, 158, 170). But the formal account of the process (pp. 169–170) is purely algorithmic, following a step-by-step sequence. Indeed, the history of Alvarez's meteorite impact hypothesis is marked by contingency: unanticipated and unplanned convergences of people, places and events. What if Alvarez's father had not known about Krakatoa, for example, or mentioned it in informal conversation? Lawson's interpretation fails to acknowledge a substantial role for luck, coincidence, a keen eye or a perceptive mind (Kohn 1989; Roberts 1989; Merton & Barber 2003). Tracing a path backwards does not explain how the process moves forward, blind to the eventual outcome. The arbitrary notion of a linear process that seems by its nature to proceed solely by brute method, '*essentially* hypothetico-predictive in nature' (p. 170), with no happenstance or contingency, is unsupported by the historical data.

The author's chief complaint against non-HP models, Jung's in particular, is that the researcher will 'not know where to look' (pp. 167, 171). In terms of Jung's analogy with forensics, the author seems to claim that one cannot investigate a crime until a suspect, motive, etc., have been hypothesized and specific evidence predicted. One may wonder, then, how crime investigators function. As noted above, Lawson merely suppresses the role of latent clues (background information and analogies) in the Alvarez case. Speculating on the *prospective relevance* of a specific variable, such as the thickness of a clay layer (pp. 171–172), is not the same as a concrete prediction based on a clear hypothesis. Much depends on intent: is the research aimed at gathering relevant information, or seriously testing an explanatory hypothesis? Historically, fruitful questions have been posed without explicit anticipated answers. Indeed, productive research programs can be based on experimental systems and tinkering with instruments, rather than any explicit theory (Hacking 1984; Rheinberger 1997). Scientists follow *many* methods: namely, whatever works or seems appropriate to the task at hand. Hence, Rosemary and Peter Grant's work on the Darwin finches – massive data collection done without any explicit hypothesis (as one notable case) – has nonetheless led to significant and widely respected claims (Weiner 1994; also see Wivagg & Allchin 2002). Paleontologists had been *puzzled* by the extinction of dinosaurs for decades. They had posed *causal hypotheses* and *predictions* to test. Despite this, however, they apparently 'did not know where to look' for the ultimate answer. Paradoxically, Alvarez wasn't even looking at the problem of dinosaurs when he arrived at his hypothesis about their demise.

Further, the author denigrates enumerative induction, yet ironically hopes to persuade the reader chiefly by repeated examples of HP (pp. 156–163 and Table 1, pp. 164–166,). One might get the impression that sample size and any kind of enumerative argument did not matter in science, that iridium need only appear in a handful of KT boundary sites, rather than in most every instance. The same strategy of enumerative induction is used in citing many authors "in agreement with" the author (p. 170) – sidestepping the arguments and widespread critiques of those arguments.

The author seems to set up a pointless and simplistic competition between HP and induction (pp. 163, 166–168), as though there were no other alternatives. The author criticizes enumerative induction, apparently implying that all induction is worthless (p. 167), without ever quite acknowledging that analogy (pp. 168–169) is actually induction by other, non-enumerative means.

The whole approach seems buoyed by a metaphysical assumption that a single, simple, linear method can account for science. For example, the author is at pains to chastise Jung's model, hoping to prove that hypotheses must precede questions (pp. 170–171). The possibility that they arise together and are pursued simultaneously is never even entertained. The author suggests indirectly that alternative hypotheses must be ruled out (p. 170, #7) – a common posture of eliminative induction (*not* HP) – but again limits the process needlessly to a *sequential, step-by-step* method of dwindling plausibility (p. 170, 175). In some cases, experimental design and interpretations of evidence are explicitly based on alternative hypotheses: why limit the nature of science by fiat? The notion of a rigid, narrowly disciplined method hardly captures the diversity of scientific reasoning and practice.

I have tried to profile here only the most egregious flaws in this paper.[2] No one contests, I trust, the periodic role of hypothesis or prediction in science, or the benefit of teaching experimental thinking. But beyond that? While there is certainly a role for "if ... then/and-or-but" thinking in science, it is a far cry from an exclusive algorithm that characterizes *the* nature of scientific discovery. As I see it, a journal such as this should help educators – and, hence, students – develop beyond such simplistic conceptions.

Finally, the reader may note that these comments are adapted (verbatim in many places) from a referee's report provided to the editor and the author before acceptance of Lawson (2004) for publication. Allchin's comments above were submitted for peer review. The category of "Discussion" is unprecedented in this journal. The reader is left to assess editorial decisions and practices in this case.

Acknowledgements

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

Notes

[1] For example, the opening example appeals to Lawson (2002) as pardigmatic exemplar. The repeated reconstruction of "if ... then ... therefore ..." formats in the table on pp. 164–166 echoes Lawson (2002, p. 16). The process listed on pp. 169–170 follows Lawson (2002, p. 17).

[2] Other problems include: (a) restricting the aims of science to causal explanations; (b) lack of discussion articulating the creative process of experimental design, by which a hypothesis, prediction or model is (re)expressed in material terms for investigation; (c) failure to effectively distinguish between a hypothesis about a natural phenomenon (scientific) and a hypothesis about the meaning of numerical results (statistical) (pp. 173–174) (see Suppes 1969; Mayo 1996, pp. 128–173). There is an additional irony that none of the section-title questions addressed in the case scenario (pp. 156–163) refer to dinosaur extinction. Lawson criticized Jung's analysis of the development of Alvarez's hypothesis (p. 171), but gives no alternative to the same question.

References

Allchin, D.: 1995a, 'How NOT to Teach History in Science Education', 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, pp. 13–22.

Allchin, D.: 1995b, Review of Lindley Darden's *Theory Change in Science* and William Bechtel and Robert Richardson's *Discovering Complexity*, *Science & Education* **4**, 399–402.

Allchin, D.: 2003, 'Lawson's Shoehorn, or Should the Philosophy of Science Be Rated "X"?', *Science & Education* **12**, 315–329.

Allchin, D.: 2004, 'Pseudohistory and Pseudoscience', *Science & Education* **13**, 179–195.

Bechtel, W. & Richardson, R.: 1994, *Discovering Complexity*, Princeton University Press, Princeton.

Darden, L.: 1991, *Theory Change in Science: Strategies from Mendelian Genetics*, Oxford University Press, Oxford.

Franklin, A.: 1986, *The Neglect of Experiment*, Cambridge University Press, Cambridge, UK.

Galison, P.: 1987, *How Experiments End*, University of Chicago Press, Chicago.

Glen, W.: 1994, *The Mass Extinction Debates: How Science Works in a Controversy*, Stanford University Press, Stanford, CA.

Hacking, I.: 1984, *Representing and Intervening*, Cambridge Unviersity Press, Cambridge, UK.

Hanson, N.R.: 1958, *Patterns of Discovery*, Cambridge University Press, Cambridge, UK.

Judson, H.F.: 1981, *The Search for Solutions*, Holt, Rinehart, Winston, New York.

Kohn, A.: 1989, *Fortune or Failure: Missed Opportunities and Chance in Scientific Discoveries*, Basil Blackwell, Oxford.

Kuhn, T.S.: 1970, *The Structure of Scientific Revolutions*, University of Chicago Press, Chicago.

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.: 2004, 'T. rex, the Crater of Doom, and the Nature of Scientific Discovery', *Science & Education* **13**, 155–177.

Mayo, D.G.: 1996, *Error and the Growth of Knowledge*, University of Chicago Press, Chicago.

Merton, R. & Barber, E.: 2003, *The Travels and Adventures of Serendipity*, Princeton University Press, Princeton, NJ.

Powell, L.: 1999, *Night Comes to the Cretaceous*, Harvest Books.

Raup, D.M.: 1992, *Extinction: Bad Genes or Bad Luck?*, W.W. Norton, New York.

Rheinberger, H.J.: 1997, *Towards a History of Epistemic Things*, Stanford University Press, Stanford, CA.

Roberts, R.M.: 1989, *Serendipity: Accidental Discoveries in Science*, John Wiley, New York.

Shapin, S.: 1989, 'The Invisible Technician', *American Scientist* **77**, 554–563.

Suppes, P.: 1969, 'Models of Data', in *Studies in the Foundation and Methodology of Science*, Reidel (Dordrecht), pp. 24–35.

Weiner, J.L.: 1994, *The Beak of the Finch*, Random House, New York.

Wivagg, D. & Allchin, D.: 2002, 'The Dogma of "The" Scientific Method', *American Biology Teacher* **64**, 484–485.