The Epistemic Value of Cautionary Tales

William M. Shields

Twice in NASA history, the agency embarked on a slippery slope that resulted in catastrophe. Each decision, taken by itself, seemed correct, routine, and indeed, insignificant and unremarkable. Yet in retrospect, the cumulative effect was stunning. In both pre-accident periods, events unfolded over a long time and in small increments rather than in sudden and dramatic occurrences. NASA's challenge is to design systems that maximize the clarity of signals, amplify weak signals so they can be tracked, and account for missing signals. For both accidents there were moments when management definitions of risk might have been reversed were it not for the many missing signals – an absence of trend analysis, imagery data not obtained, concerns not voiced, information overlooked or dropped from briefings. A safety team must have equal and independent representation so that managers are not again lulled into complacency by shifting definitions of risk . . . Because ill-structured problems are less visible and therefore invite the normalization of deviance, they may be the most risky of all.
– Vol. I, Section 8.5, *Report of the Columbia Accident Investigation Board* (August 2003).

For those involved in any way with the creation and management of modern technology, the report on the loss of the space shuttle *Columbia* (NASA 2003) should be required reading. Technically accurate and devoid of hype or newsroom exaggerations, the report's spare prose offers a warning that should shake both young engineers learning their trade and their older counterparts who may have drifted into management and finance. In the densely-packed pages of the report, we read of a faulty design never corrected, of precursor events that came to be accepted as ordinary, of many chances to assess the damage to the spacecraft left lying on the table, of warning voices ignored in the name of pushing ahead with the mission, and of the many counterfactual cases that might have led to a lost spacecraft but a living crew.

The cautionary tale has become something of a cottage industry in the past decade. To be sure, there has been plenty of material for these publications: Bhopal, Chernobyl, *Exxon Valdez*, Three Mile Island, and *Challenger* have entered the lexicon as virtual synonyms for "disaster." One of the more readable and fascinating of the cautionary-tale collections is *Inviting Disaster* by James Chiles (2001). Of the cautionary tales Chiles tells, I have personal knowledge only of Three Mile Island, and as to TMI his work is accurate. Most of the stories make chilling reading, yet not because of the deadly outcome. They are chilling because of their "disaster-waiting-to-happen" atmosphere. My own favorite is "The Really Bad Day," the story of American Airlines pilot Bryce McCormick attending training school as an introduction to the new DC-10 jumbo jet. In the course of inspecting the new airliner's cargo bay, McCormick observed that the triply-redundant hydraulic systems for the aircraft's control surfaces all ran in the same area of the cargo hold, and due to the aircraft's immense size, there were no manual backups. In terms of safety engineering, he had stumbled on a "common mode failure" that made him nervous. That nervousness led him to teach himself to fly the huge airliner by balancing engine power, simulating the loss of all control circuits. This new skill very soon became the only safety feature standing between total disaster and a safe if shaky landing. This story contains a least two valuable lessons: the ease with which well-intentioned design can be defeated by human error; and compensation for that critical error by another individual's powerful sense of personal responsibility.

After thirty years in the nuclear industry, both commercial and defense, I need very little convincing that cautionary tales of this kind are important. Much of my career has been focused on protecting nuclear materials from the consequences of fire. The near-disaster that was my first exposure to a cautionary tale was a dangerous fire at TVA's Browns Ferry Nuclear Station in 1985. (NRC 1975) This fire heavily damaged the plant's control systems, and a meltdown was averted only by a combination of human action and conservative design. The fire was started by a candle held inside a wall to look for ventilation leaks. A taper candle and the human being holding it nearly melted down the core of a nuclear generating station. Ten years and uncounted millions of dollars later, U.S. nuclear power plants had been redesigned and rebuilt to

prevent and mitigate fires.

Let me return now to the *Columbia* report. This was indeed an "accident waiting to happen." All of the elements were there: faulty design, failure to thoroughly consider the worst-case consequences of the design flaw, refusal to take actions to assess the possible damage after launch, decision to go ahead with reentry without any knowledge of the condition of the spacecraft. These elements can all be thought of as arrows or vectors, converging on a single event that cost the lives of the crew, destroyed the spacecraft, threatened the survival of the space station, and heavily damaged the reputation of NASA. Those vectors were created and aimed by *individuals* as well as management practices and organization charts, though we hesitate to assign personal blame for loss of life. Strangely, in the United States we seem more than willing to fix blame on individuals (e.g., the Enron executives) when mere money is lost. I don't believe anyone has been charged with negligence as a result of *Columbia*'s needless destruction.

Of what value are these cautionary tales? Do we read them with fascination just to experience the sense of relief that "It wasn't me" or "It wasn't my fault?" If so, then we gain no real value from them. They need to be read for more than that, and whatever that "more" is, it needs to be incorporated into engineering curricula, drummed into the heads of all technologists, and perhaps made the basis for Enron-like prosecutions of individuals. How do I as an engineer know when I may be participating in a cautionary tale *as it is unfolding*, and what can I do as a responsible human being to make a difference in the outcome? It is not enough, it seems to me, to argue that meeting a code of ethics is the limit of our responsibilities. That is not to dismiss the codes; they are important in their own sphere of relevance. But no code of engineering ethics imposes the generalized burden of watching for the precursors of failure and taking timely actions (even at the cost of career damage) to change the course of events.

Perhaps the fundamental difficulty in learning from cautionary tales of man-caused disasters is epistemological: what kind of *knowledge* do these tales constitute? As engineers and scientists, we are most comfortable thinking in a straightforward causal fashion, i.e., if I do X the likely result is Y. If I provide the code-required safety margin for structural strength of a steel beam, the likely result is that the beam will remain intact following an earthquake. Most engineers who have to consider safety are also comfortable with what is termed "failure modes and effects" analysis. If the beam does fail in an earthquake, what else will happen? How many beams must remain intact before structural collapse follows?

Cautionary tale scenarios involve a different kind of causality. In analyzing technology-related disasters after the fact, we find ourselves examining causal sequences that had no apparent connection to one another. For example, one causal factor may be a component part that was properly procured according to the correct specification. How could that be a cause? Because it may turn out that the specification itself is not adequate given an unanticipated series of events. Foam regularly broke off the shuttle booster tank and struck the vehicle on ascent, causing some minor tile damage. After a number of such occurrences, NASA managers assumed that the vehicle was built sturdily enough to withstand these impacts, without really considering what might happen if the impact occurred in a slightly different way. On the *Columbia* mission, the foam punched a hole in the wing rather than damaging a few tiles. The result was catastrophic failure once the decision was made to land without examining the hull.

Let me give another illustration from my own field of fire protection engineering. In July of 1998, a team of trained workers was preparing to conduct preventive maintenance on Idaho National Laboratory's Engineering Test Reactor. The high-voltage equipment to be serviced was located in a large room protected by a total deluge carbon dioxide suppression system, automatically actuated by a modern fire control panel. Naturally the power to the area was to be cut off before the work began. The workers entered the area with portable lights powered by long cables from an adjacent area. When they were in position to begin work, the power to the area was shut off, momentarily plunging the room into darkness. Before the workers could turn on their portable lamps, the CO-2 system discharged without warning, creating total whiteout conditions as the gas condensed into snow. Visibility fell to zero. Some workers near the exit door managed to escape before breathing the gas. One worker held his breath but ran the

wrong way, ending up at a locked door. He smashed a glass pane with his hand, severely lacerating his arm, then passed out. Several other workers collapsed before reaching the exit. Those who made it out found they had no breathing air equipment available; it had to be obtained from a cabinet some distance away that was found to be locked. By the time breathing air could be brought back to the area, one trapped worker had died. But for the selfless actions of fellow workers and others who were nearby, there could have been many more fatalities.

What went wrong? The CO-2 system had been disabled by a software command entered into the fire control panel. Why had it discharged instantly when power to the fire control panel was shut off? The system was designed to give an audible alarm 30 seconds before discharge; no alarm had sounded. This question turned out to be very difficult to answer, but eventually the cause was found. The alarm panel was equipped with batteries, and automatically switched to that power source when line power was lost. But in that short space of time while battery power was being activated, the panel's circuitry sometimes, but not always, generated an activation pulse that went directly to the CO-2 system's control valve, bypassing the alarm/delay circuit. This event was very hard to duplicate because the panel did not send a spurious pulse in every case of power transfer.

This was a subtle failure mode. But is that what *really* went wrong in this cautionary tale? Would replacing the control panel be a suitable response? The fatality was also caused by an excessive reliance on the fire control panel, by a lack of planning for the eventuality of a system discharge, and by the failure of anyone to take a truly cautious attitude toward what was clearly a life-threatening hazard. The result was that workers had been placed in an unfamiliar room filled with CO-2 discharge heads isolated from pressurized tanks by a single valve controlled by computer software and circuitry. All of the "arrows" were pointing in the same direction, turning what should have been a routine electrical project into a disastrous situation depriving one man of his life and threatening many others. The post-accident report (DOE 1998) identified all of these factors, of course, and recommended specific and appropriate changes. But my point in relating this story is to ask the epistemic question: what kind of knowledge was needed a day before this fatal accident occurred to prevent it from happening, and what might cause that knowledge to be actionable. It is a strange sort of knowledge and not at all what engineers are used to dealing with, because it cannot be read in a textbook, calculated from an equation, or even acquired by reading cautionary tales. Perhaps *knowledge* is not the right word: what is needed is more a form of intuition arising from a suitable combination of attitude, assignment, experience, and technical understanding. This is not "quality assurance" or "discipline of operations," because these functions tend to be controlled by consensus standards and practices. Decisions to launch the space shuttle and to permit its reentry are tightly constrained by myriad procedures, checkpoints, concurrences and the like. Unfortunately, the end-oriented pressure of unrelated causes can overwhelm these well-intended precautions.

In cautionary tales, what we see is a causal sequence in which many unrelated factors seem to converge, almost conspire, to bring about an unexpected and usually undesirable result. The individual factors are *not* the common cause events engineers are trained to look for, such as the multiple redundant hydraulic systems running right alongside each other in a aircraft. As in the case of the *Columbia*, we see *after the fact* a combination of largely independent causes involving hardware, systems, operations, and human judgment. For this reason, the more typical responses to cautionary tales often do not prove effective in doing much more than preventing that particular event sequence from occurring again.

I am well aware that some critics of complex technological systems urge that we look for new, less complex forms of technology that will somehow be less vulnerable to failure. Unfortunately, these well-meaning suggestions are misguided. Nearly all of the cautionary tales I have studied do not reveal a pattern of complexity as the principal cause of calamity. We are not defeated by complexity itself, nor are complex systems necessarily more hazardous than simple ones. Claims that we have overreached ourselves, that our technologies have become inherently uncontrollable and hence dangerous, are in my view groundless. There is no turning back to a simpler time, because the history of technology shows that earlier eras

were in no measurable way more benign in terms of human safety. Carriages with spindly wooden wheels drawn by teams of massive horses threatened both riders and pedestrians. Sailing across huge spans of ocean in wooden ships bearing cloth sails was frightfully hazardous. Those beautiful, romantic steamboats ended many lives when their primitive boilers exploded. We will not find greater safety in a strategic retreat to less complex technologies.

We do need, however, to raise our level of thinking about the ways the modern technological systems can fail us and do harm, even after we have provided what we believe to be a conservative design and have tried to behave in ways we imagine to be safety conscious. Specifically, we need to learn new ways to approach the causality reflected in the cautionary tales of our times, to go beyond narrow, event-driven technical fixes and organizational changes. It will not be enough to approach the problem as one of design: the best design can always be defeated by human error and unusual circumstances. We will never succeed in preventing *Columbia* or TMI-type accidents by better design practices alone. Design fixes will prevent the accident that has already happened, but they will not anticipate or prevent a different cautionary tale from being entered into.

I believe that we need to develop a heuristic methodology that teaches us to think in somewhat teleological terms, as if we are looking backwards from a potential catastrophic event to see how current actions and events may be contributing to an as-yet-unrealized accident sequence. We have to find the vectors that point toward the unexpected failure before they are allowed to converge and reinforce each other. In short, we need to develop a sense of when "an accident is waiting to happen." This is easy enough to say, but very hard to do. Thinking in this backwards fashion is not part of our training and in a sense is counterintuitive. How can we evaluate what we are doing in light of an end-state that we cannot fully specify?

Time and effort will be needed to work out a methodology of this kind. One step has already occurred, at least to a degree. There is a growing literature on cautionary tales that provides valuable case studies. (Dörner 1989; Duffey 2003; Perrow 1999; Sagan 1993; Tenner 1996) More needs to be done, to be sure, and at a level of technical sophistication that is meaningful for engineers and other technologists. The *Columbia* report sets a high but not unattainable standard. We need to study all of the unexpected causal chains, all of the events that seem to have confounded our ability to design and operate technological systems safely.

The next step is to work this line of inquiry into our academic institutions, where teachers and students alike may have the time, energy, intellectual prowess, and objectivity needed to make progress. I am proud to report that my alma mater, MIT, now offers a course to engineering students entitled " Colossal Failures in Engineering." The course is focused on:

> Case studies of known "colossal failures" from different engineering disciplines. Includes the collapse of the World Trade Center, the *Columbia* Space Shuttle accident, and the melt down at Chernobyl. Basic engineering principles are stressed with descriptions of how the project was supposed to work, what actually went wrong, and what has been done to prevent such failures from reoccurring.

This is headed in the right direction, but we must be cautious about being too quick to find "what went wrong" and "what can be done to prevent it." It is easy for engineers in particular to seek the single, dominant "cause" and then identify a "fix." Preventing the same failures from occurring is generally not that difficult once the analysis is done. I would like to see courses like this focus on how those involved with the colossal failure might have anticipated it and then prevented it. We want to learn how to close the barn door *before* the horses get out.

Any methodology we might come up with must be tested, of course, in the real world of engineering design and management of complex systems. This means taking the bold step of embedding in our technological infrastructure an intuitive, somewhat teleological function, assigned to highly-trained and experienced persons whose sole responsibility is to search out "accidents waiting to happen." Perhaps this notion would not meet with much resistance, as it sounds like—though it is not—a somewhat more fancy version of quality assurance. The hard part is ensuring that those who design, build and manage complex technological systems listen and respond to the voice of the Cautionary Tale Division. Consider: it would

have taken only one NASA senior manager to demand that *Columbia*'s true condition be ascertained before ordering the reentry.

What I am arguing for, in the end, is a new approach to the managing of technological risk, one that can identify a dangerous condition caused not by a single error in design or maintenance but by factors that may appear on the surface to be unrelated. If such an approach can be found, it must then be embedded in the management systems we use to control our most complex and hazardous technologies, from the space shuttle to the electric power grid. The cumulative impact of decisions made by individual managers are the divider between safety and disaster. These decisions can and should be informed, and when needed deflected, by the analytical capability of seeing *when* we are in a cautionary tale before the "accident waiting to happen" is upon us.

Dr. William M. Shields recently retired from the Department of Science and Technology in Society at Virginia Tech.

References

Chiles, James R. (2001). *Inviting Disaster: Lessons from the Edge of Technology*. New York: HarperCollins.

DOE (U.S. Department of Energy) (1998). *Type a Accident Investigation Board Report of the July 28, 1998, Fatality and Multiple Injuries, Resulting from Release of Carbon Dioxide at Building 648, Test Reactor Area*, Idaho National Engineering and Environmental Laboratory, September 1998.

Dörner, Dietrich (1989). *The Logic of Failure*. Reading: Addison-Wesley.

Duffey, Romney Beecher, & Saull, John Walton (2003). *Know the Risk: Learning from Errors and Accidents: Safety and Risk in Today's Technology*. New York: Butterworth-Heinemann.

MIT (Massachusetts Institute of Technology) (2005). Online course catalog: http://student.mit.edu/catalog/m22a.html.

NASA (National Aeronautics and Space Administration) (2003). *Report of the Columbia Accident Investigation Board*. http://anon.nasa-global.speedera.net/anon.nasa-global/CAIB/CAIB_lowres_full.pdf.

NRC (Nuclear Regulatory Commission) (1976). *Recommendations Related to Browns Ferry Fire*, NUREG-0050, February 1976.

Perrow, Charles (1999). *Normal Accidents: Living with High-Risk Technologies*. Princeton: Princeton University Press.

Petroski, Henry (1982). *To Engineer is Human*. New York: Vintage.

Petroski, Henry (1997). *Remaking the World*. New York: Vintage.

Sagan, Scott D. (1993). *The Limits of Safety: Organizations, Accidents, and Nuclear Weapons.* Princeton: Princeton University Press.

Tenner, Edward (1996). *Why Things Bite Back*. New York: Vintage.

Vaughan, Dianne (1997). *The Challenger Launch Decision: Risky Technology, Culture, and Deviance at NASA*. Chicago: University of Chicago Press.