The Real World: Blooming Buzzing Confusion

Karlene H. Roberts, *University of California, Berkeley*

My purpose here is to discuss some of the issues we come face to face with when we try to deal with risk and tools relevant to risk in the real world. Imagine for a minute the manager or operator beleaguered from every side with the requirement to engage in "risk-based" management. What's wrong with this?

- First, if you're like me, you don't know what risk-based management means.
- Second, you may have some gnawing notion that some things need fixing.
- Third, there are a lot of would-be "Mr. Fix-Its" (and fewer Ms. Fix-Its) who probably want to help you fix things.

What do you do? Today I want to try to answer that question. Although I'm somewhat familiar with the commercial marine industry, I'm going to try to answer the question by drawing on experiences faced by both the marine industry and other industries. I don't buy the notion that the commercial marine industry is so different from anything else that lessons learned there can't at least be tried out in ports, on offshore platforms, and on the waterways. My assertion is at least partially based on research Bob Bea, his students, my students, and I have done showing that safety impediments in your industry are closely related to similar impediments in other industries.

Four questions gnaw at us in trying to operate nearly failure-free organizations:

- What things really need fixing (and how do I know)?
- How do I fix them?
- What are the impediments to fixing them?
- How much will it cost me to fix them?

Let me address the last question first because it is the easiest. What will it cost me to fix them? Plenty. Whatever the problem, the cost of fixing it will be more than you or your company wants to pay. The company may believe the worst-case scenario will never happen here and refuse to fix the problem, or it may rely on cost models of fixing the problem that greatly underpredict actual cost, which happened to Ford Motor Company in the case of the Pinto. Here's what Carolyn Libuser says about the problem (Roberts and Libuser, 1993):

> Traditional risk analysis utilizes mathematical modeling and expected values to aid in making decisions about risk or about pricing risk. While this is a valuable methodology, we argue that in some cases it is not the appropriate way to assess risk and other methods need to be utilized. This is illustrated with two examples.
>
> In the case of pricing life insurance conventional risk analysis is highly appropriate. Insurance companies use actuarial tables to assess the probability of an insured person's death at any particular age. Companies also know the probability that a policy will lapse (a large number lapse within five years). Insurance companies make money by establishing price so the expected value of the policy (the probability of death during the expected time period of the policy times its face value) is, on average, less than the accrued value of the premiums paid on the policy.
>
> On the other hand, the January 17, 1994, Northridge, California, earthquake points out how other types of in-

surance are risky bets for companies (the industry's losses from the quake are currently estimated at about 10 billion dollars according to various insurers such as Allstate, Farmers, and State Farm). Earthquake insurance is a tremendous loser for insurance companies because they are unable to predict accurately the probability of a quake or the amount of damage it inflicts. Thus, the expected value of policies cannot be determined. As a result insurance companies charged a premium based on best estimate of a quake and in 1994 those estimates were very wrong. Consequently, insurance companies absorbed huge losses at the time of the Northridge earthquake and most companies withdrew from writing homeowners policies in California.

This example can be applied to the case of the *Exxon Valdez*. Even if the expected value of that oil spill had been calculated correctly (which it wasn't) before the spill, once the spill occurred the only relevant cost was that of the spill and the cleanup that followed. The moral of this story is that it almost always costs more to fix something after the damage is done than it costs to fix it before the accident. How industries deal with this is a big problem. They usually fail to deal with it at all. Or they deal with it through legislation that is insufficient at best. It needs to be dealt with through shared cultural values about things like environmental protection or even more simply the impact of catastrophic organizational outcomes on the organization's reputation.

What things really need fixing? This really breaks down into three subquestions: (*a*) How do I find out what's wrong? (*b*) Where do I get my notions about how to think about what's wrong? and (*c*) How do I assess the match between what I think is wrong and what is really wrong?

First, how do I find out what's wrong? Generally, as an industry you know what's wrong because you have data on things such as accident rates. So, for example, you know that bulker accidents happen more frequently than tanker accidents and over time probably result in greater loss. And sometimes it's easy to find out what is wrong—CNN or Mike Wallace from "Sixty Minutes" pays you a visit and exposes what's wrong. Alternatively, a regulator may have exposed what's wrong. These things can more or less identify the big things—things that can overwhelm even the best-defended system. But it's better for the organization to figure out what's wrong before CNN or anyone else gets there.

Many of the things that creep up on us year after year until they finally result in the "big one" are cumulative and opaque. Jim Reason has produced an enormously popular model for thinking about these things, the Swiss cheese model. This model identifies more than simply human errors or the usual operator name-and-blame syndrome that is still prevalent in many industries. The important thing about Reason's contribution is that he identifies "fire walls" increasingly distant from the perpetrator of the error that contribute to some catastrophic outcome and asks us to look at what can go wrong at these successive fire walls—just as Karl Weick asked you to do. One would do well to lay over the static fire wall model some of the more dynamic ways of thinking about identifying processes that lead to error that Karl alerted us to earlier.

Consistent with Karl's approach but from more of an engineering perspective is the approach used in medicine and discussed in the February 1, 1999, *New Yorker* article "When Doctors Make Mistakes." On page 51, the article cites Jeff Cooper's efforts to uncover the nature of mishaps in anesthesiology. Jeff relied heavily on "critical incidents methodology," which had been used since the 1950s to analyze mishaps in aviation and is at the foundation of the Aviation Safety Reporting System (ASRS). Although critical incidents methodology has been used in the marine industry, it is not a tool most of you would think to pull off the shelf in your own diagnostic efforts. But those of us who have had surgeries are pleased it was used in medicine because it was part of the foundation for reducing errors in anesthesia.

Where do I get my notions about how to think about what's wrong? You need to find a way to think about what contributes to the things you don't want to happen, like the bulk carrier that hit the River Walk in New Orleans in 1996, the fire on the cruise ship *Ecstasy* in summer 1998, or the Tosco refinery fire in the San Francisco area that killed four people in February of this year. This is the same refinery where one person was killed in 1997 and where there have been four fatalities in 16 years.

Karl's reminder to us that we want methods that produce explanations that are at once general, accurate, and simple is a good place to begin, particularly as he tells us that we can at best get only two out of three. He also reminds us that there are different kinds of knowledge. And he shows us that the way we organize our thinking determines our actions. We need to develop or borrow some sort of model that offers us a starting place for organizing our thinking.

I want to offer a first. I think the focus on risk in this industry is misplaced. It is a loaded word that conjures up visions of wrongdoing, hapless behaviors, and other negative things. If we move to its opposite—in this context, improving reliable, safe operations—we go a long way toward reducing the defensiveness and self protections that get in the way of any real problem solving. Let's take a different tack to thinking about the outcomes we all want to realize—reliable, error-free operations—and search for mechanisms to improve reliability. A reliability model focuses on a different set of issues than does a risk model. The National Patient Safety Foundation (NPSF) of the American Medical Association takes this broader tack in its view of patient safety. Here's part of what NPSF thinks safety is:

> Safety emerges from the interaction of components of the system. It is more than the absence of adverse outcomes and it is more than the avoidance of identifiable "preventable" errors or occurrences. Safety does not reside in a person, device, or department. Improving safety depends on learning how safety emerges from the interactions of the components.

We need to focus on realizing the good as opposed to thwarting the bad or at least focus on the two simultaneously. Reason's Swiss cheese model does both and was my way of introducing you to another organizing principle, one that can show us where to look for causes of error. It tells us to look at the operator as embedded in his or her organization, and we can carry this further to talk about the organization as embedded in its systems of regulation and competition.

I'm sure if one looked closely at the air tragedy on April 3, 1996, that killed Commerce Secretary Ron Brown and 34 other people on a flight from Croatia to Dubrovnik one would find more than simply operator error. In a press release on the accident the Air Force stated that the field command approved the mission despite orders to the contrary from headquarters. Other questions I've not seen addressed raise other issues that direct our attention to the larger fabric within which the accident happened. Why did the aircraft fly to Dubrovnik in such bad weather? What was the urgency? Who determined the urgency? Why didn't the pilots turn back? Didn't the Dubrovnik airport perceive that a landing could be very risky? If not, why not?

In the real world do managers and other decision makers really use models to guide their behavior about what to look for? Yes. Here are two examples. In 1968, after United Airlines lost an aircraft in Portland, it began to look for some answers to reducing flight error and improving safety performance. United adopted findings from the then current social psychological literature on team performance in developing its version of crew resource management training. The concepts and models borrowed from social psychology did not suggest the necessity of looking beyond the skin of the cockpit but it did allow the airline industry to begin somewhere. Today those models and concepts are sufficiently developed that they are borrowed to inform bridge team training in the maritime industry and they're widely used in other industries as well.

In 1995, after losing three F-14D Tomcats from the same squadron the U.S. Navy wondered how to assess performance safety in its air community. They borrowed a model that was originally developed in the banking industry. Some of you have seen this simple model developed by Carolyn Libuser in her Ph.D. dissertation (Exhibit 1).

It goes a step beyond something like the original crew resource management approach of the commercial airlines by introducing notions that the behavior of the organization at large is as important as the behavior of the team. This is more in keeping with what Jim Reason and Karl Weick would have us look at. The model addresses the following processes, and in a minute I'll tell you something more about its Navy application.

Where do you get your models or schemes for thinking about error reduction and high performance behavior? You might get them from inside your industry, or you might unabashedly and unashamedly borrow them from other industries or directly from the fields in which they were developed. You use engineering models every day in the design and operation of what you do. It's possible that other kinds of models can be equally helpful. After all, crew resource management is based on social psychological concepts, and Carolyn's management model is based on research on organizations. Now that I have a scheme for organizing my thinking about how things should work, how do I find out if they work that way?

How do I assess the match between what I think is wrong and what is really wrong? Many times managers and other well-meaning people believe they have an adequate description of what is wrong (for example, no process auditing here, inattention to appropriate rewards there, and so forth) and they move to fix it, often at considerable cost. But that may not be what's wrong here or, more likely, a little more of one thing you think is wrong really is wrong, and a little less of another is wrong.

Part of the problem with the crew resource management approach is that it has been applied to settings where no one knows if it is useful. In many of these situations it would be tough to test its utility and so managers are left to go on blind faith that it improves things. In the real world new programs of any sort are rarely tested for utility. This is true for virtually all the quality programs instituted in the United States.

But you can run various kinds of tests of your notions about what will work, some more rigorous than others. All require someone to get their hands dirty and go into the organization and find out what's happening. Is something you're doing making a difference? You can't assess its value through probabilistic risk assessment. You want to assess "that something" against performance data, but often this is where the problem is. There are no performance data. I'm going to give you two examples of what can be done.

Greg Bigley at the University of Cincinnati has a problem of no performance data. He deals with community emergency service teams and is particularly interested in how incident command systems function. That's the sit-

EXHIBIT 1 Libuser's Risk Mitigation Model

Process Auditing

Establish a system for ongoing checks designed to identify expected and unexpected safety problems. Safety drills and equipment testing are important aspects of this audit. Follow up on problems identified in prior audits.

Appropriate Reward Systems

Reward systems must reward desired behaviors. Organizational reward systems have powerful influences on the behavior of individuals.

High-Quality Systems

The quality of the system must compare favorably with the quality of a referent system that is generally regarded as the standard for quality.

Risk Perception

Risk perception has two elements:

1. Knowledge that risk exists at all, and
2. If risk exists, the extent to which it is acknowledged appropriately and minimized.

Command and Control

Command and control have five subfactors:

1. Command by negation—this includes migration of decision making to the person with the most experience, not necessarily the highest-level person;
2. Redundancy—in people or hardware; backup systems exist;
3. Formal rules and procedures—a hierarchy exists but it is not a bureaucracy in the negative sense;
4. Training; and
5. Senior managers who have the "big picture"—the senior managers don't micromanage.

uation where a big community emergency happens and teams of experts from different geographic areas who probably have never worked together come together to solve the problem. Greg has a model that includes the notion that the way these people are successful depends on a number of things, two of which are the development of instantaneous or swift trust and the way people form a joint representation in their heads of the problem they are trying to solve.

Greg can measure these things. His problem is that emergency task forces don't keep performance data. The best he can do is describe what he thinks happens and ask operators if what he says makes sense to them. Remember, it's the operators, not the managers, who know what's going on where the rubber meets the road. From his analyses Greg can then help managers try to decide what changes they think might be useful. He can even try to develop some performance measures against which they can later assess the old versus the new ways of doing business. Sometimes those outcome measures should not include the usual about how many mishaps were prevented, how the accident rate was lowered, and so forth. Maybe a good outcome measure would be "worker contentedness" because a contented, trained work force is not as apt to walk off the job and leave the manager in a less safe situation.

As I mentioned previously the Navy has a simple management model and, based on its evaluation of the banking research at the foundation of the model, thought that the processes operating in high-performing banks might be equally important in high-performance air squadrons. The Navy decided there were two ways to assess whether encouraging operation of the processes in the model would be helpful to them. An initial step was to ask squadrons to assess their performance on these processes and have them discuss the value of each process to them in maintaining high safety performance.

From Carolyn's model they developed a questionnaire and gave it to 1,245 aviators in randomly selected squadrons. A team analyzed the data and told squadron commanding officers how well their squadrons were doing at process auditing, making sure the correct rewards were in place, ensuring high standards of quality, and so forth. The squadron commanders thought this way of conceptualizing what goes on in their squadrons and using this conceptualization as a discussion springboard with squadron members was helpful in alerting them to attend to the "right" safety characteristics.

The conundrum comes about when you think about the fact that the Navy (or your organization) doesn't often experience the awful things it tries to avoid. So it is difficult to match questionnaire responses to catastrophic outcomes. The appropriate step is to think about what might be precursors to the really awful things. Often you can measure them and match some of the organizational processes, in this case the five emanating from the model, to the outcomes. That is what the Navy is in the process of doing for all its aircraft squadrons. Clearly, the organization will discover that some processes are more important than others and that still others are simply irrelevant to the goal of aviation improving safety performance. But the organization will have in place, and in data banks, a way to assess the safety health of aircraft squadrons.

You may think such an approach won't work or that it is impossible to do in the marine industry. Not so. Bob Bea and his students developed a similar approach to assessing management processes in the marine industry. The approach has been tested in a marine terminal, on an oil drilling platform, in U.S. Navy diving operations, and in U.S. Coast Guard bridge operations.

We've discussed how you come to know what really needs fixing—through believing what CNN identifies as bad problems or, better yet, by engaging in some activity that can get at the source of the genotypic processes that underlie the phenotypic outcropping of bad practices. It is entirely possible to consistently treat the symptom (five people killed over a 2-year period in the same refinery) without ever getting to the underlying processes, which means the symptom will return. You get at these underlying processes by developing or applying some model of them and then adjust that model for your specific situation. Then you try to test whether the model is related to or possibly predicts the behavior you want in your situation.

How do I fix things? This is the difficult part because it usually involves changing the norms of an entire organization or set of organizations. For people with the kinds of backgrounds most people in the maritime industry have (engineering backgrounds) fixing some things is easy. You take a wrench and fix it. But if you believe what is often said, that "80 percent of the problems are people problems," tweaking an engineering solution one more time will not help. Surely you can work toward making some engineering improvements, such as replacing old equipment and making sure the equipment itself is safe.

But the rest is the difficult part. Many of you don't believe what every industry is learning—that the real problems reside in things such as the organization's structure, culture, training, and reward systems. Implementing a good process auditing system in an organization that doesn't have one will probably require changing its structure. Being sensitive to keeping the right quality standards in place rests on good training. These are things organizations don't want to put money into. But ignoring these things results in huge costs when the "big one" happens and even when some of the little ones happen. Tosco paid fines after the 1997 accident, and the company found itself living in a county that, because of its behavior, instituted stronger safety regulations; is disliked by the environmentalists; is threatened with shutdown; and today has one of the poorest reputations of any organization in its industry in the San Francisco Bay Area. Surely, these things are borne as cost to the organization.

What are the impediments to fixing things? I'm tempted to say I think the largest impediment is cost, but I don't think that's true. The largest impediment to organizations doing what they should be doing is the culture of the organization itself. It is in part created by the larger economic milieu in which the organizations live. In the Tosco case, the company tried to respond to falling oil prices with employee layoffs. Layoffs often contribute to lowered safety standards. Until companies take the long-term view by including a fix-it line in their accounting systems instead of the short-term view of "not on my watch," calamities will continue to occur. It is also in part contributed to by what I call "John Wayne management." That is, a sense of invulnerability on the part of top management, the feeling that if we just keep moving down this road things will turn around.

Cost alone is decisively a factor. Until organizations recognize that the cost of not fixing is higher than the cost of fixing, fixing won't get done. The cost of fixing Tosco now is much higher than it was in January 1999. And one wonders what the decision makers at Chernobyl in 1986 think today about preaccident versus postaccident fix-it costs.

Another impediment to change is that often we don't know what to change. The regulators don't know what to change, the organizations don't know what to change, and the industry doesn't know what to change. The commercial nuclear power industry attempted to address this problem when it formed INPO and its international counterpart. INPO is the industry's clearinghouse for training and other activities relevant to nuclear power plant safety. The FAA attempts to do this through the ASRS, offering airlines and everyone else an opportunity to learn by studying near misses. Today there is a growing amount of research coming from a number of different fields about what to change. And today I've offered you a primer about how to figure out what to change in your setting.

To find out more about some of the things that go wrong and some things various industries have learned about change, see the short reading list in the Bibliography.

REFERENCES

Bogner, M. S. (ed.) *Human Error in Medicine.* Lawrence Erlbaum, Hillsdale, N.J., 1994.

Flin, R. *Sitting in the Hot Seat: Leaders and Teams for Critical Incident Management.* Wiley, New York, 1996.

Gawande, A. When Doctors Make Mistakes. *New Yorker Magazine,* Feb. 1, 1999, p. 40.

Grabowski, M., and K. H. Roberts. Risk Mitigation in Large Scale Systems. *California Management Review,* Vol. 39, 1997, pp. 152–162.

Moran, K. Danger Aboard. *Houston Chronicle,* Aug. 20, 1996, pp. 1A, 6A.

Morris, J. Pressure to Set Sail. *Houston Chronicle,* Aug. 18, 1996.

Morris, J., and K. Moran. Lost at Sea. *Houston Chronicle,* Aug. 18, 1996, pp. 1A, 24A, 25A.

Moss, T. H., and D. L. Sills. *The Three Mile Island Accident: Lessons and Implications.* New York Academy of Sciences, 1981.

Pauchant, T. C., and I. I. Mitroff. *Transforming the Crisis Prone Organization.* Jossey Bass, San Francisco, 1992.

Perrow, C. *Normal Accidents: Living with High-Risk Technologies.* Basic Books, New York, 1984.

Pfeiffer, J. The Secret of Life at the Limits: COGs Become Big Wheels. *Smithsonian Magazine,* July 1989, pp. 38–48.

Phillips, D. F. "New Look" Reflects Changing Style in Patient Safety Enhancement. *Journal of the American Medical Association,* Vol. 281, No. 3, Jan. 20, 1999, pp. 217–219.

Read, P. P. *Ablaze: The Story of the Heroes and Victims at Chernobyl.* Random House, New York, 1993.

Reason, J. *Managing the Risks of Organizational Accidents.* Ashgate, London, 1997.

Rees, J. V. *Hostages of Each Other: The Transformation of Nuclear Safety Since Three Mile Island.* University of Chicago Press, Chicago, 1994.

Roberts, K. H. Managing Hazardous Organizations. *California Management Review,* Vol. 32, Summer, 1990, pp. 101–113.

Roberts, K. H. (ed.). *New Challenges to Understanding Organizations.* Macmillan, New York, 1993.

Roberts, K. H., and C. Libuser. From Bhopal to Banking: Organizational Design Can Mitigate Risk. *Organizational Dynamics,* Vol. 21, 1993, pp. 15–26.

Sagan, S. *The Limits of Safety: Organizations, Accidents, and Nuclear Weapons.* Princeton University Press, Princeton, N.J., 1993.

Short, J. F., and L. Clarke (eds.). *Organizations, Uncertainties, and Risk.* Westview, Boulder, Colo., 1992.

Shrivastava, P. *Bhopal: Anatomy of a Crisis.* Ballinger, Cambridge, Mass., 1987.

Vaughan, D. *The Challenger Launch Decision: Risky Technology, Culture, and Deviance at NASA.* University of Chicago Press, Chicago, 1996.

Weick, K. E. Organizational Culture as a Source of High Reliability. *California Management Review,* Vol. 29, 1987, pp. 116–136.