CALIFORNIA MANAGEMENT REVIEW
Volume XXIX, Number 2, Winter 1987

Organizational Culture as a Source of High Reliability

Karl E. Weick

As organizations and their technologies have become more complex, they have also become susceptible to accidents that result from unforeseen consequences of misunderstood interventions.[1] Recent examples include Bhopal, the Challenger, and Three Mile Island.

What is interesting about these examples is that they involve issues of reliability, not the conventional organizational issues of efficiency. Organizations in which reliability is a more pressing issue than efficiency often have unique problems in learning and understanding, which, if unresolved, affect their performance adversely.

One unique problem is that a major learning strategy, trial and error, is not available to them because errors cannot be contained. The more likely an error is to propogate, the less willing a system is to use trial and error to understand that source of error firsthand. Because of this limitation, systems potentially know least about those very events that can be most damaging because they can propogate widely and rapidly. This article will explore an unconventional means by which organizations achieve error-free performance despite limited use of trial and error.

The point is that accidents occur because the humans who operate and manage complex systems are themselves not sufficiently complex to sense and anticipate the problems generated by those systems. This is a problem of "requisite variety,"[2] because the variety that exists in the system to be managed exceeds the variety in the people who must regulate it. When people have less variety than is requisite to cope with the system, they miss important information, their diagnoses are incomplete, and their remedies are short-sighted and can magnify rather than reduce a problem.

If the issue of accidents is posed this way, then there should be fewer accidents when there is a better match between system complexity and human complexity. A better match can occur basically in one of two ways:

either the system becomes less complex or the human more complex. This article has more to say about the latter alternative than about the former.

Since learning and reliable performance are difficult when trial and error are precluded, this means that reliable performance depends on the development of substitutes for trial and error. Substitutes for trial and error come in the form of imagination, vicarious experiences, stories, simulations, and other symbolic representations of technology and its effects. The accuracy and reasonableness of these representations, as well as the value people place on constructing them, should have a significant effect on the reliability of performance.

The basic idea is that a system which values stories, storytellers, and storytelling will be more reliable than a system that derogates these substitutes for trial and error. A system that values stories and storytelling is potentially more reliable because people know more about their system, know more of the potential errors that might occur, and they are more confident that they can handle those errors that do occur because they know that other people have already handled similar errors.

Training as a Source of Accidents

Training is often used to prevent errors, but in fact can create them.

Training for the operation of high reliability systems if often tough and demanding so that the faint of heart and the incompetent are weeded out. People training to be air traffic controllers, for example, are targets of frequent verbal abuse in the belief that this will better prepare them to deal with pilots who are hostile, stubborn, and unresponsive. As one trainer said, "When you are training to be an air traffic controller, you only walk away from your screen once in anger and you're out." The trainer assumes that people who walk away from training screens under instructor abuse would walk away from real screens under pilot abuse. The problem with that assumption is that its validity never gets tested.

Furthermore, trainees who are unable to handle trainer hostility may be good controllers because they are more likely to sense emotions conveyed by pilots and be better able to predict impending problems. Thus, a person who walks away from a training screen may be better able to deal with the more frequent problem of emotional communications than with the less frequent problem of pilot abuse. The other side of the argument, however, is that in situations where there is the possibility of catastrophic failure, the marginal gain made by keeping controllers who are more sensitive to emotions is less important than the possibility that the price of this sensitivity is occasional inadequate response.

Training settings themselves often have modest validity, as is shown by the widespread agreement that much that is learned during training for air traffic control has to be unlearned once the controller starts to work traffic. For example, simulators used in training do not accurately simulate change of speed in an airplane. When a plane on the simulator changes speed, the new speed takes effect immediately, whereas in real life the change in speed is gradual. That discrepancy could become consequential because people under pressure revert to their first-learned ways of behaving.[3] Under pressure, controllers who first learned to control planes that changed speed swiftly might systematically underestimate the time it actually takes for planes to execute the speed changes that are ordered. Thus, a controller will assume that developing problems can be resolved more quickly than in fact is the case.

When people are trained for high reliability, the first tendencies they learn are crucial because those may be the ones that reappear when pressure increases. If trainers are more concerned with weeding out than with the adequacy of initial responses, then training could again become the source of a breakdown in reliability which it was designed to prevent.

Even when training works, there are problems. If training is successful, there is usually no pattern to the errors trainees make once they actually operate the system. But if operational errors are randomly distributed because the training is good, then it is not clear how operators can reduce those errors that still occur, since there is no way to predict them.

In each of these examples, the benefits of training are either diluted or reversed due to unanticipated effects of emotional, social, and interpretive processes. Closer attention to these processes may uncover new ways to cope with conditions that set accidents in motion. To illustrate this possibility, following is an examination of three ways in which variations in the social construction of reality can affect the likelihood of error-free operations.

Reliability and Requisite Variety—As noted, to regulate variety, sensors must be as complex as the system which they intend to regulate. An interesting example of less complicated humans who try to manage more complicated systems is found in the repeated observation that the first action of many senior airline captains, when they enter the cockpit and sit down, is to turn up the volume control on the radio equipment to a level which the junior officers regard as unnecessarily high.[4] Data show that the number of aircraft system errors are inversely related to pilots acknowledging the information they receive from controllers. More errors are associated with fewer acknowledgements.[5] When pilots acknowledge a message, they are supposed to repeat the message to verify that they have received it correctly, but busy pilots often acknowledge a transmission by saying simply "Roger," and at other times, they make no acknowledgement at all.

While aircraft errors are often attributed to communication deficiencies, the observation that senior officers turn up the volume on the radio suggest that one source of error may be a hearing deficiency. Older commercial pilots often learned to fly in older, noisier aircraft; and chronic exposure to these conditions may cause current messages to be missed or heard incorrectly. (The hypothesis of a hearing deficiency was not ruled out in the Tenerife disaster on March 27, 1977, and is consistent with all the data assembled about that accident.)

Problems with hearing deficiency are not confined to the airways. Before people are licensed to operate the reactor at Diablo Canyon, they spend up to 5 years as Auxiliary Operators, which means they work on the floor servicing pipes and valves before they ever set foot inside a control room. As is true with most power generation plants, Diablo Canyon is noisy. This creates the same possible history for reactor operators as is created for senior pilots: they develop less sensory variety than is present in the systems of signals, alarms, voices, and strange noises that are symptoms of changes in the system they are trying to control.

Humans gain as well as lose the variety that is requisite for reliability in several ways. Daft and Lengel, for example, propose that the ways in which people receive information provide varying amounts of requisite variety.[6] Information richness is highest when people work face-to-face, and informational richness declines steadily as people move from face-to-face interaction to interaction by telephone, written personal communiques (letters and memos), written formal communiques (bulletins, documents), and numeric formal communiques (computer printouts). Effectiveness is postulated to vary as a function of the degree to which informational richness matches the complexity of organizational phenomena. Rich media provide multiple cues and quick feedback which are essential for complex issues but less essential for routine problems. Too much richness introduces the inefficiencies of overcomplication, too little media richness introduces the inaccuracy of oversimplification.

In the context of the Daft and Lengel argument, it becomes potentially important that communication between Morton Thiokol and NASA about the wisdom of launching Challenger in unusually cold temperatures was made by a conference telephone call,[7] a medium with less variety than a face-to-face conversation. With only voice cues, NASA did not have visual data of facial expressions and body cues which might have given them more vivid information about the intensity of Thiokol's concerns.

Face-to-face communication in high reliability systems is interesting in the context of the large number of engineers typically found in such systems. One way to describe (and admittedly stereotype) engineers is as smart people who don't talk. Since we know that people tend to devalue what they don't do well, if high reliability systems need rich, dense talk to maintain complexity, then they may find it hard to generate this richness if talk is devalued or if people are unable to find substitutes for talk (e.g., electronic mail may be a substitute).

Up to this point, we have been talking largely about the development of requisite variety in individuals, but in high reliability organizations, requisite variety is often gained or lost by larger groups. When technical systems have more variety than a single individual can comprehend, one of the few ways humans can match this variety is by networks and teams of divergent individuals. A team of divergent individuals has more requisite variety than a team of homogeneous individuals. In problems of high reliability, the fact of divergence may be more crucial than the substance of divergence. Whether team members differ in occupational specialities, past experience, gender, conceptual skills, or personality may be less crucial than the fact that they do differ and look for different things when they size up a problem. If people look for different things, when their observations are pooled they collectively see more than any one of them alone would see. However, as team members become more alike, their pooled observations cannot be distinguished from their individual observations, which means collectively they know little more about a problem than they know individually. And since individuals have severe limits on what they can comprehend, a homogeneous team does little to offset these limits. This line of argument, which suggests that collective diversity increases requisite variety which in turn improves reliability, may conflict with the commmon prescription that redundancy and parallel systems are an important source of reliability. That prescription is certainly true. But a redundant system is also a homogeneous system and homogenous systems often have less variety than the environments they are trying to manage and less variety than heterogeneous systems that try to manage those same environments.

The issues with collective requisite variety are fascinating as well as complex.

As an example of collective requisite variety, the operating team in the control room at Diablo Canyon has five people who stay together as a team when they change shifts. The lead person on the team is the Shift Foreman whose responsibility is to maintain the "big picture" and not to get into details. There is a Shift Technical Advisor who has engineering expertise and a Senior Control Room Operator who is the most senior union person in the control room. Under the Senior Operator are the Control Room Operator and the Assistant Control Room Operator, the latter being the person who has the newest operating license and who most recently has worked outside the control room. What is striking about this team is the spread of attention among several issues induced by diverse roles.[8]

The issue of effective delegation of responsibility is crucial in high reliability systems. The most effective means for airline pilots to handle crisis, for example, is for the captain to delegate the task of flying the plane and then make decisions about how to handle the crisis without worrying about the details of flying. Obvious as this solution may seem, a failure to delegate positive responsibility for flying the plane has often meant that a crisis absorbed everyone's attention, no one flew the plane, and it crashed.[9]

The importance of collective requisite variety as a means to enhance reliability is one reason people are increasingly concerned about the reduction of flight crews from three to two people, and are especially concerned when that two-person crew is mixed gender. A female co-pilot adds considerable requisite variety, but if it is hard for a male to trust a woman communicating in an environment that has culturally always been a "man's world," then the two-person crew quickly loses variety and errors become more likely.

Collective requisite variety is higher when people both trust others, which enlarges the pool of inputs that are considered before action occurs, and themselves act as trustworthy reporters of their own observations to enlarge that same pool of inputs. Trust, however, is often difficult when diversity increases, because as people become more diverse they also become harder to trust and it is harder to be trusted by them.

Social psychologists have studied these issues in the context of the Asch conformity experiments. Collective requisite variety "is maximized when each person so behaves as to be in his turn a valid dependable model for the others. Each acts as both model and observer."[10] Translated to the Asch situation, this means that the best response for the sake of requisite variety is for the naive subject exposed to discrepant reports to say, "'You fellows are probably right, but I definitely see line B as longer,' i.e., both rationally respecting others as a source of information about the world, and so reporting that others can rationally depend on his report in turn. It is failure in this latter respect that instigates our moral indignation at the conformant chameleon character who parasitically depends upon the competence of others but adds no valid information, no clarifying triangulation, to the social pool."[11]

Requisite variety is enhanced by face-to-face communications for two reasons. First, face-to-face contact makes it easier to assess and build trust and trustworthiness. Second, face-to-face contact makes it easier to get more complete data once trust and trustworthiness have been established. Since people are the medium through which reliability is accomplished, signals relevant to reliability flow through them. When those people are both trusted and dealt with face to face, more information is conveyed, which should produce earlier detection of potential errors.

Building trust in high reliability systems is difficult because so much is at stake. People want to delegate responsibility, but not too soon and not without continued surveillance to see that continued delegation is warranted. A neat resolution of this dilemma is found in the comment of a Navy nuclear propulsion expert who parries a complaint from a second-class petty officer. The complaint goes, "Sir, I'm not stupid or incompetent. I've had over a year of specialized training and two years of experience, but no one trusts me. Everything I do is checked and double-checked." The engineer replies, "It's not a matter of trust. Your ability, training, and experience allow me to trust completely that in an emergency you will do the right thing—or at least a survivable thing. In a non-emergency situation, however, . . . we all make mistakes. . . . That is why your work is checked."[12]

This particular system builds reliability by institutionalizing an important bit of evolutionary wisdom: "Ambivalence is the optimal compromise."[13] The answer to the question "Don't you trust me?" is both yes and no: "Yes, I trust you if it's an emergency; no, I don't trust you if it's practice." Application of this rule presumably builds confidence, competence, and trustworthiness so that trust takes care of itself when the stakes rise dramatically.

Reliability Is a Dynamic Non-Event—Reliability is both dynamic and invisible, and this creates problems. Reliability is dynamic in the sense that it is an ongoing condition in which problems are momentarily under control due to compensating changes in components. Reliability is invisible in at least two ways. First, people often don't know how many mistakes they could have made but didn't, which means they have at best only a crude idea of what produces reliability and how reliable they are. For example, if a telephone switching unit normally loses dial tone for 12 minutes out of 72 hours, but could have potentially lost it for 15 minutes, that suggests very different information about its reliability than if it could have lost dial tone for 120 minutes during that same period. Reliability is also invisible in the sense that reliable outcomes are constant, which means there is nothing to pay attention to. Operators see nothing and seeing nothing, presume that nothing is happening. If nothing is happening and if they continue to act the way they have been, nothing will continue to happen. This diagnosis is deceptive and misleading because dynamic inputs create stable outcomes.

A nuclear power plant operator: "I'll tell you what dull is. Dull is operating the power plant." Another operator describing his plight: "I have total concentration, for five hours, on nothing happening." A senior officer on a nuclear carrier: "When planes are missing the arresting wire, and can't find the tanker where they are to refuel, and the wind is at 40 knots and the ship is turning, there are no errors." This latter experience is confirmed in studies of air-traffic controllers. There tend to be more errors in air traffic control under light traffic load than under heavy load because, under high load, controllers visually sweep the entire radar screen whereas in low load they don't. When they fail to sweep the screen, problems can build to more extreme levels at the edges.

When there are non-events, attention not only flags, it is often discouraged. Consider the homespun advice: if it ain't broke, don't fix it. The danger in that advice is that something that isn't broken today, may be tomorrow. Just because a two-engine Boeing 767 airplane hasn't crashed yet while crossing the Atlantic Ocean doesn't mean that a system with two rather than three pilots having two rather than three engines is a reliable system. What it means is that there hasn't been any trial and error on two-engine transatlantic flights, which means people don't yet have any idea of what they know about such flying. That uncertainty can give way to an illusion that since there have been no errors, there must be nothing to learn, which must mean we know what there is to know.

More attentiveness and more reliability might be induced if we were able to shift the homespun advice from its static form to a more dynamic form: if it isn't breaking, don't fix it. Such a modification might alert observers to the dynamic properties of reliable situations, to the fact that small errors can enlarge, to the possibility that complacency is dangerous, to the more active search for incipient errors rather than the more passive wait for developed errors. Both the early explosions of the de Haviland Comet jet airliners as well as the capsizing of the newly designed Alexander L. Keilland oil rig were traced to small cracks that enlarged gradually and then catastrophically under high stress.[14] The Comet aircraft exploded when a crack, which started at the edge of one cabin window after repeated pressurization and depressurization, suddenly enlarged and ripped open the skin of the aircraft.[15] The oil rig collapsed when a 3-inch crack in the frame, which was painted over rather than re-welded, enlarged during a North Sea gale.[1]

Part of the mindset for reliability requires chronic suspicion that small deviations may enlarge, a sensitivity that may be encouraged by a more dynamic view of reliability. People need to see that inertia is a complex state, that forcefield diagrams have multiple forces operating in opposed directions, and that reliability is an ongoing accomplishment. Once situations are made reliable, they will unravel if they are left unattended.

While it is a subtle point, most situations that have constant outcomes—situations such as a marriage, or social drinking, or an alcohol rehabilitation program—collapse when people stop doing whatever produced the stable outcome. And often what produced the stable outcome was continuous change, not continuous repetition. We all smile when we hear the phrase, "the more things change, the more they stay the same." The lesson in that truism for problems of reliability is that sameness is a function of change. For a relationship to stay constant, a change of one element must be compensated for by a change in other elements.

When people think they have a problem solved, they often let up, which means they stop making continuous adjustments. When the shuttle flights continued to depart and return successfully, the criterion for a launch—"Convince me that I should send the Challenger"—was dropped. Underestimating the dynamic nature of reliability, managers inserted a new criterion—"Convince me that I shouldn't send the Challenger."

Reward structures need to be changed so that when a controller says, "By God, I did it again... not a single plane collided in my sector today," that is not treated as a silly remark. If a controller can produce a dull, normal day, that should earn recognition and praise because the controller had to change to achieve that outcome.

People aren't used to giving praise for reliability. Since they see nothing when reliability is accomplished, they assume that it is easier to achieve reliability than in fact is true. As a result, the public ignores those who are most successful at achieving reliability and gives them few incentives to continue in their uneventful ways.

Reliability as Enactment—A peculiar problem of systems is that people in them do not do what the system says they are doing. John Gall illustrates this problem with the example of shipbuilding. "If you go down to Hampton Roads or any other shipyard and look around for a shipbuilder, you will be disappointed. You will find—in abundance—welders, carpenters, foremen, engineers, and many other specialists, but no shipbuilders. True, the company executives may call themselves shipbuilders, but if you observe them at their work, you will see that it really consists of writing contracts, planning budgets, and other administrative activities. Clearly, they are not in any concrete sense building ships. In cold fact, a *system* is building ships, and the *system* is the shipbuilder."[17]

If people are not doing what the system says they are doing, then they know less about what is dangerous and how their own activities might create or undermine reliability. Just as nurses commit medical errors when they forget that the chart is not the patient, operators commit reactor errors when they forget that the dial is not the technology.

These misunderstandings, however, are not inevitable. There are systems which achieve high reliability because, for them, the chart is the patient. They provide one model of ways to restructure other systems which have reliability problems.

The system that will illustrate the argument is the air traffic control system. One striking property of air traffic control is that controllers are the technology, they don't watch the technology. Controllers build their own system in the sense that they build the pattern of aircraft they manage by interacting with pilots, using standard phraseology, and allocating space. The instructions that controllers issue are the system and hold the system together. Controllers do not suffer the same isolation from the world they work with as do people in other systems. For example, air traffic controllers on the carrier Carl Vinson make an effort to learn more about the quirks of their carrier pilots. As a result, they are able to separate quick responders from slow responders. This knowledge enables controllers to build a more stable environment when they line up pilots to land on a carrier under conditions where high reliability performance is threatened. Because controllers also use voice cues, they often are able to build a more complete picture of the environment they "face" because they are able to detect fear in voices and give a fearful pilot more airspace than is given to a confident pilot.

The ability of controllers to enact their environments can be interpreted as the use of slack as a means of increase reliability. Controllers can add slack to the system by standardizing their customers so that they expect more delays. Overload comes not so much from the number of planes that a controller is working as from the complexity of the interactions with the pilot that occur. A complicated reroute can monopolize so much time that a nearly empty sky can become dangerous when the few remaining planes are totally neglected. If controllers can reduce the number of times they talk to a pilot and the length of time they talk to a pilot, they can add slack to their system.

Controllers can also hold planes on the ground, slow them, accelerate them, turn them sooner, line them up sooner, stack them, or refuse to accept them, to build an environment in which reliability is higher. Airplanes stacked into holding patterns provide a perfect example of an enacted environment. Space which had previously been formless and empty now is structured to have layers 1000 feet apart, a shape, and a pattern in which planes enter at the top and exit from the bottom. An environment has been created by the controller which then constrains what he or she does.

While a stack is a good example of an enacted environment, it also illustrates that when people construct their own environments, they create problems as well as solutions. When a controller creates slack by stacking airplanes, this solution creates at least two problems. First, stacks "get loose," which means that planes drift outside the circular pattern as well as up or down from their assigned altitude. Second, a stack, when viewed on a radar screen, creates lots of clutter in a small space so it is harder for the controller to keep track of all the players.

As if discretion and looseness were not enough heresy to introduce into a discussion of high reliability, it is also important to make the point that the air traffic control system works partly because it is an exercise in faith. Unless pilots and controllers each anticipate what the other is going to say, the clipped phraseology they use would never work. Their communiques usually ratify expectations rather than inform, which means that if an unexpected remark is made, then people begin to listen to one another.

For example, foreign national pilots (e.g., China Air) who fly into San Francisco International Airport, and for whom English is a distant second language, are hard to understand. Controllers are unsure what those pilots have heard or what their intentions are. In cluttered skies, this uncertainty increases the probability of error. The system around San Francisco is held together by faith in the sense that the pilots and controllers each anticipate what they will be told to do and each tries to meet these anticipations as much as possible. The elegant solution adopted when the language problem is especially severe is that the foreign pilot is directed to fly straight to the airport and land, and all other aircraft, piloted by people who have better command of English, are routed around the straight-in flight.

The importance of faith in holding a system together in ways that reduce errors has been discussed for some time as "The Right Stuff." The right stuff often creates reliability, and the way it does so is important to identify, partly because that process is currently in jeopardy at NASA.

No system can completely avoid errors. Any discussion of reliability must start with that as axiomatic. Actors frequently underestimate the number of errors that can occur. But if these same actors are dedicated people who work hard, live by their wits, take risks, and improvise, then their intense efforts to make things work can prevent some errors. Because they are able to make do and improvise, they essentially create the error-free situation they expected to find. What they fail to see is that their own committed efforts, driven by faith in the system, knit that system together and create the reliability which up to that point existed only in their imaginations.

While this mechanism is sometimes interpreted as macho bravado,[18] it is important to remember that confidence is just as important in the production of reliability as is doubt. The mutually exclusive character of these two determinants can be seen in the growing doubt among astronauts that they have been flying the safe system they thought they were. Notice that the system itself has not suddenly changed character. What has changed is the faith that may have brought forth a level of commitment that created some of the safety that was anticipated. Obviously, there are limits to faith. The Challenger did explode. But whatever increments to safety the process of faith may have added are no longer there as astronauts see more clearly the shortcuts, problems, and uncertainties which their committed efforts had previously transformed into a temporarily functioning system.

While the activity of air traffic control can be viewed in many ways, I have chosen to emphasize that qualities such as discretion, latitude, looseness, enactment, slack, improvisation, and faith work through human beings to increase reliability. The air traffic control system seem to keep the human more actively in the loop of technology than is true for other systems in which reliability is a bigger problem. It is not immediately clear what the lesson in design is for a nuclear power generation facility, but neither is it self-evident that such a design question is nonsensical. The air traffic control system, because it has not been taken over by technology, accommodates to human limitations rather than automates them away.

But there are threats to the enacted quality of air traffic control and they come from plans to automate more control functions so that the system can be speeded up. Any automated system that controls traffic can go down without warning, which forces controllers to intervene and pick up the pieces.[19] If automated control allows planes to fly closer together (e.g., separated by 10 seconds), the problem is that then when the control fails, humans will not be able to pick up the pieces because they are not smart enough or fast enough. A system in which looseness, discretion, enactment, and slack once made reliability possible will have become a system in which reliability is uncontrollable.

Automation, however, is not at fault. Automation makes a 10-second separation possible. But just because such separation is possible, doesn't mean it has to be implemented. That remains a human decision. The heightened capacity isn't dumb, but the decision to heighten capacity may be.

Cultures of Reliability

In many of the problems we've looked at, the recurring question is, "What's going on here?" That's not so much a question of what decision to make as it is a question of what meaning is appropriate so we can then figure out what decision we need to make. It is important to underscore that difference because more attention is paid to organizations as decision makers than to organizations as interpretation systems that generate meaning. That's one reason why the recent interest in organizational culture is important, because it has redistributed the amount of attention that is given to issues of meaning and deciding. That shift in emphasis is summarized in Cohen, March, and Olsen's observation that "an organization is a set of procedures for argumentation and interpretations as well as for solving problems and making decisions."[20]

One reason organizational theorists have had trouble trying to think clearly about issues of reliability is that they have made a fundamental error when they think about meaning and decisions. A discussion by Tushman and Romanelli illustrates the problem.[21] In their presentation of an evolutionary model of organization, they argue that managers are concerned both with making decisions and with managing meaning. They accommodate these two rather different activities by arguing that managers make decisions when environments are turbulent and make meanings when environments are stable. I think they've got it backwards, and that's symptomatic of the problems people have when they think about reliability and how to achieve it.

To make decisions, you need a stable environment where prediction is possible, so that the value of different options can be estimated. The rational model works best in a stable environment. When environments become unstable, then people need first to make meaning in order to see what, if anything, there is to decide. When there is swift change, you either label the change to see what you should be paying attention to, or you take action in an effort to slow the change so that you can then make a rational decision. Stabilization and enactment make meaning possible, which means they necessarily precede decision making.

Making meaning is an issue of culture, which is one reason culture is important in high reliability systems. But culture is important for another reason. Throughout the preceding analysis, I have highlighted the importance of discretion and have played down the necessity for centralization.

But the real trick in high reliability systems is somehow to achieve simultaneous centralization and decentralization. People need to benefit from the lessons of previous operators and to profit from whatever trials and errors they have been able to accumulate. And when errors happen, people need a clear chain of command to deal with the situation. These are requirements of centralization. There has to be enough centralization that no one objects when the airline captain delegates authority for flying the plane while he tries to focus his full attention on the crisis. A control room full of people all shouting contradictory diagnoses and directions, as was the case at Three Mile Island, does little to clarify thinking.

But a system in which both centralization and decentralization occur simultaneously is difficult to design. And this is where culture comes in. Either culture or standard operating procedures can impose order and serve as substitutes for centralization. But only culture also adds in latitude for interpretation, improvisation, and unique action.

Before you can decentralize, you first have to centralize so that people are socialized to use similar decision premises and assumptions so that when they operate their own units, those decentralized operations are equivalent and coordinated.[22] This is precisely what culture does. It creates a homogeneous set of assumptions and decision premises which, when they are invoked on a local and decentralized basis, preserve coordination and centralization. Most important, when centralization occurs via decision premises and assumptions, compliance occurs without surveillance. This is in sharp contrast to centralization by rules and regulations or centralization by standardization and hierarchy, both of which require high surveillance. Furthermore, neither rules nor standardization are well equipped to deal with emergencies for which there is no precedent.

The best example of simultaneous centralization and decentralization remains Herbert Kaufmann's marvelous study of the Forest Ranger[23] which shows, as do many of Philip Selznick's analyses,[24] that whenever you have what appears to be successful decentralization, if you look more closely, you will discover that it was always preceded by a period of intense centralization where a set of core values were hammered out and socialized into people before the people were turned loose to go their own "independent," "autonomous" ways.

It is potentially relevant that operators and managers in many nuclear power reactors (those with fewest errors?) have had prior Navy nuclear experiences and that many FAA controllers are former military controllers. In both cases, there are previously shared values concerning reliability which then allow for coordinated, decentralized action. The magnitude of shared values varies among power stations as does the content of the values which are shared. A research question that may predict the likelihood of errors is the extent of sharing and the content that is shared on operating teams.

Culture coordinates action at a distance by several symbolic means, and one that seems of particular importance is the use of stories. Stories remind people of key values on which they are centralized. When people share the same stories, those stories provide general guidelines within which they can customize diagnoses and solutions to local problems.

Stories are important, not just because they coordinate, but also because they register, summarize, and allow reconstruction of scenarios that are too complex for logical linear summaries to preserve. Stories hold the potential to enhance requisite variety among human actors, and that's why high reliability systems may handicap themselves when they become preoccupied with traditional rationality and fail to recognize the power of narrative rationality.[25]

Daft and Wiginton have argued that natural language, metaphors, and patterns that connect have more requisite variety than do notation, argumentative rationality, or models.[26] Models are unable to connect as many facts as stories, they preserve fewer interactions, and they are unable to put these interactions in motion so that outcomes can be anticipated.

Richard Feynman tells a story about the Challenger disaster when he dips O-ring material from the booster into a glass of ice water and discovers that it becomes brittle. Rudolph Pick, a chemical engineer writing to the *New York Times* on January 14, 1986, observed that the only way he could impress people with the danger of overfilling vessels with chemicals was to use what he called the psychological approach. "After I immersed a piece of chicken meat for several minutes in the toxic and corrosive liquid, only the bone remained. Nobody took any short cuts to established procedures after this demonstration and there were no injuries." Pick tells this story about hydrofluoric acid and the message remains with people once they scatter to their various assignments. Thus, the story coordinates them by instilling a similar set of decision premises. But the story also works because, from this small incident, people are able to remember and reconstruct a complicated set of chemical interactions that would be forgotten were some other medium, such as a set of regulations, used.

When people do troubleshooting, they try to tell stories that might have, as their punch lines, the particular problem that now confronts them.[27] When stories cannot be invented, troubleshooting and reliability become more difficult. For example, Diablo Canyon has a poor memory for some past decisions and relatively few stories. This creates trouble when people find that a problem can be traced to an odd configuration of pipes. When this happens, they face the disturbing possibility that the odd configuration may solve some larger, more serious problem that no one can remember. Rerouting may solve the immediate problem, but it might also set in motion an unexpected set of interactions that were once anticipated and blocked, though no one now can recall them. Stories about infrastructure are not trivial.

What all of this leads to is an unusual reconstruction of the events of the night of January 27, 1986, when NASA was arguing with Morton Thiokol about whether freezing weather would disable the booster rocket. That conversation apparently took the traditional course of people arguing in linear, sequential fashion about the pros and cons of a launch. If, somewhere in those discussions, someone had said, "That reminds me of a story,"[28] a different rationality might have been applied and a different set of implications might have been drawn. Those, in turn, might well have led to a different outcome. There are precedents in history. The solution of the Cuban Missile crisis by a surgical airstrike was dropped when Robert Kennedy recalled the story of Pearl Harbor, and portrayed a U.S. attack on Cuba as Pearl Harbor in reverse.[29]

We have thought about reliability in conventional ways using ideas of structure, training, and redundancy, and seem to be up against some limits in where those ideas can take us. Re-examination of the issue of reliability using a less traditional set of categories associated with an interpretive point of view seems to suggest some new places to attack the problem.

References

Author's Acknowledgement:
I am grateful to Lisa Berlinger, Larry Browning, George Huber, Todd LaPorte, Reuben McDaniel, Karlene Roberts, and Sim Sitkin for comments on an initial draft of this manuscript.

The analyses in this article represent work in progress and are derived from interaction with a group at Berkeley that is concerned with hypercomplex organizations and a group at Texas that is concerned with narrative rationality. The key people in the Berkeley group include Karlene Roberts, Todd LaPorte, and Gene Rochlin. Key people at Texas include Larry Browning, George Huber, Reuben McDaniel, Sim Sitkin, and Rich Cherwitz. The data with which I am working come from observations and interviews with people who operate the Diablo Canyon nuclear reactor, the Nuclear Carrier U.S.S. Carl Vinson, and the air traffic control center at Fremont, California, as well as workshops, literature reviews, and discussions.

1. C. Perrow, *Normal Accidents* (New York, NY: Basic Books, 1984).
2. See, for example, W. Buckley, "Society as a Complex Adaptive System," in W. Buckley, ed., *Modern Systems Research for the Behavioral Scientist* (Chicago, IL: Aldine, 1968), pp. 490–513.
3. K. E. Weick, "A Stress Analysis of Future Battlefields," in J. G. Hunt and J. D. Blair, eds., *Leadership on the Future Battlefield* (Washington, D.C.: Pergamon, 1985), pp. 32–46.
4. R. Hurst, "Portents and Challenges," in R. Hurst and L. R. Hurst, eds., *Pilot Error: The Human Factors*, 2nd Ed. (New York, NY: Aronson, 1982), p. 175.
5. Ibid., p. 176.
6. R. L. Daft and R. H. Lengel, "Information Richness: A New Approach to Manager Information Processing and Organization Design," in B. Staw and L. L Cummings, eds., *Research in Organizational Behavior: Vol. 6* (Greenwich, CT: JAI Press, 1984), pp. 191–233.
7. R. J. Smith, "Shuttle Inquiry Focuses on Weather, Rubber Seals, and Unheeded Advice," *Science*, 231 (1986): 909.

8. N. W. Biggart and G. G. Hamilton, "The Power of Obedience," *Administrative Science Quarterly*, 28 (1984): 540–549.
9. See, for example, E. L. Wiener, "Mid-Air Collisions: The Accidents, the Systems, and the Realpolitik," in R. Hurst and L. R. Hurst, eds., *Pilot Error: The Human Factors*, 2nd Ed. (New York, NY: Aronson, 1982), pp. 101–117.
10. D. T. Campbell, "Conformity in Psychology's Theories of Acquired Behavioral Dispositions," in I. A. Berg and B. M. Bass, eds., *Conformity and Deviation* (New York, NY: Harper, 1961), p. 123.
11. Ibid.
12. J. D. Jones, "Nobody Asked Me, But...," *United States Naval Institute Proceedings*, November 1977, p. 87.
13. D. T. Campbell, "Ethnocentric and Other Altruistic Motives," in D. Levine, ed., *Nebraska Symposium on Motivation, 1965* (Lincoln, NE: University of Nebraska), pp. 2283–2311.
14. H. Petroski, *To Engineer Is Human* (New York, NY: St. Martin's Press, 1985).
15. Ibid., p. 178.
16. Ibid., p. 174.
17. J. Gall, *Systematics: How Systems Work and Especially How They Fail* (New York, NY: New York Times Book Co., 1977), pp. 32–33.
18. M. Allnutt, "Human Factors: Basic Principles," in R. Hurst and L. R. Hurst, eds., *Pilot Errors: The Human Factors*, 2nd Ed. (New York, NY: Aronson, 1982), pp. 14–15.
19. J. M. Finkelman and C. Kirschner, "An Information-Processing Interpretation of Air Traffic Control Stress," *Human Factors*, 22 (1980): 561.
20. M. D. Cohen, J. G. March, and J. P. Olsen, "People, Problems, Solutions, and the Ambiguity of Relevance," in J. G. March and J. P. Olsen, eds., *Ambiguity and Choice in Organizations* (Bergen, Norway: Universitetsforlaget, 1976), p. 25.
21. M. L. Tushman and E. Romanelli, "Organizational Evolution: A Metamorphosis Model of Convergence and Reorientation," in L. L. Cummings and B. M. Staw, eds., *Research in Organizational Behavior: Vol. 7* (Greenwich, CT: JAI Press, 1985), pp. 196, 209–212.
22. C. Perrow, "The Bureaucratic Paradox: The Efficient Organization Centralizes in Order to Decentralize," *Organizational Dynamics*, 5/4 (1977): 3–14.
23. H. Kaufman, *The Forest Ranger* (Baltimore, MD: Johns Hopkins, 1967).
24 P. Selznick, *Leadership in Administration* (New York, NY: Harper & Row, 1957).
25. K. E. Weick and L. B. Browning, "Argument and Narration in Organizational Communication," *Journal of Management*, 12 (1986): 243–259.
26. R. L. Daft and J. Wiginton, "Language and Organizations," *Academy of Management Review*, 4 (1979): 179–191.
27. N. M. Morris and W. B. Rouse, "Review and Evaluation of Empirical Research in Troubleshooting," *Human Factors*, 27 (1985): 503–530.
28. M. H. Brown, "That Reminds Me of a Story: Speech Action in Organizational Socialization," *Western Journal of Speech Communication*, 49 (1985): 27–42.
29. P. A. Anderson, "Decision Making by Objection and the Cuban Missile Crisis," *Administrative Science Quarterly*, 28 (1983): 211–212.