UNIVERSITY OF CALIFORNIA, BERKELEY

BERKELEY • DAVIS • IRVINE • LOS ANGELES • MERCED • RIVERSIDE • SAN DIEGO • SAN FRANCISCO SANTA BARBARA • SANTA CRUZ

TELEPHONE: (510) 643-8678
TELEFAX: (510) 643-8919
E-MAIL: bea@ce.berkeley.edu
HTTP://www.ce.berkeley.edu/

CENTER FOR CATASTROPHIC RISK MANAGEMENT
DEPARTMENT OF CIVIL & ENVIRONMENTAL ENGINEERING
212 McLAUGHLIN HALL
BERKELEY, CALIFORNIA 94720-1710

APPROACHES TO ACHIEVE ADEQUATE QUALITY AND RELIABILITY

QUALITY & RELIABILITY

One of the important developments in life-cycle (concept development through decommissioning) assessments and management of engineered systems has been redefinition of the terms 'quality' and 'reliability' in a holistic way. These definitions help focus the efforts on a balanced and comprehensive understanding of the potential performance characteristics of a given engineered system.

Quality is defined as freedom from unanticipated defects in offshore systems. Quality is fitness for purpose. Quality is meeting the requirements of those who own, operate, design, construct, and regulate engineered systems. These requirements include those of serviceability, safety, compatibility, and durability. Quality is taken to be freedom from unanticipated defects in the serviceability, safety, durability, and compatibility of the offshore structure system.

Serviceability is suitability for the proposed purposes, i.e. functionality. Serviceability is intended to guarantee the use of the system for the agreed purpose and under the agreed conditions of use. Safety is the freedom from excessive danger to human life, the environment, and property damage. Safety is the state of being free of undesirable and hazardous situations. The capacity of a structure to perform acceptably during extreme demands and other hazards is directly related to and most often associated with safety. Compatibility assures that the structure does not have unnecessary or excessive negative impacts on the environment and society during its life-cycle. Compatibility also is the ability of the structure to meet economic, time, and environmental requirements.

Durability assures that serviceability, safety, and environmental compatibility are maintained during the intended life of the structure. Durability is freedom from unanticipated maintenance problems and costs. Experience with engineered systems has shown that durability is one of the most important characteristics that must be achieved; if insufficient durability is developed, then there are unanticipated and often undetected degradations in the other quality characteristics, and many times these degradations have disastrous results.

Note that concerns for safety have been integrated with the other quality attributes that largely dictate the operability / availability and financial viability of the system. In addition, engineered 'systems' have been defined in a holistic way to include: operating personnel, organizations (local, corporate), equipment (hardware), structures (supporting facilities), procedures (formal, informal, software), environments (internal, external, social), and the interfaces between the foregoing. The 'stop-rule' in the characterization of such systems is to stop the identifications and analyses when the effects of adding more elements / components produces relatively insignificant effects on quality and reliability.

Reliability is defined as the probability (likelihood) that a given level of quality will be achieved during the design, construction, operating, and maintenance life-cycle phases of an engineered system. Reliability is the likelihood that the structure system will perform in an acceptable manner. Acceptable performance means that the structure system has desirable serviceability, safety, compatibility, and durability. The compliment of reliability is the likelihood or probability of unacceptable quality; the probability of failure. This definition has linked the concepts of probability, uncertainty, and reliability with the holistic definition of quality to reflect upon the likelihoods of achieving acceptable quality in engineered systems.

Compromises in quality of an engineered system can occur in the structure itself and / or in the facilities it supports. These failures can be rooted in malfunctions developed by individuals (operators) in design, construction, operation, and / or maintenance. Individuals, the people who design, construct, operate, and maintain the systems have direct influence on malfunctions developed in these phases. However, the malfunctions developed by the individuals can be and often are caused (contributing factors) or compounded (propagating factors) by malfunction inducing influences from organizations, hardware, software (procedures), and environment (external, internal).

The quality and reliability of an engineered system can be directly influenced by two primary categories of factors: intrinsic, and extrinsic. Intrinsic factors are hazards that can result in compromises in the quality of the system that are 'natural' or due to inherent randomness (residual risk elements). Model, parametric, and state uncertainties are also included in intrinsic factors. Extrinsic factors are hazards that can result in compromises in the quality of the engineered system that are 'unnatural' or caused by human and organizational factors (HOF) and knowledge acquisition and utilization (unknown knowables and unknown unknowables). HOF can result in human and organizational errors (HOE); these are misadministrations or malfunctions that have unanticipated and undesirable outcomes. Human errors represent outcomes from interactions of a complex series of initiating, contributing, and compounding factors. Unknown knowables represent information access and utilization challenges; the information exists, but it is not accessed or accessed properly. Unknown unknowables represent limitations to predictability and knowability; the information does not exist and can not be known in advance of the events that challenge the system.

APPROACHES

There are three fundamental, complimentary and interactive approaches to achieving adequate and acceptable quality (serviceability, safety, durability, compatibility) and reliability in engineered systems:

- **Proactive** (activities implemented before malfunctions occur),
- **Reactive** (activities implemented after malfunctions occur), and
- **Interactive** or real-time. (activities implemented during occurrence of malfunctions

In the context of these three approaches there are three primary strategies to be employed:

- **Reduce incidence of malfunctions**,
- **Increase detection and correction of malfunctions**, and
- **Reduce effects of malfunctions**.

Proactive Approaches

The proactive approach attempts to analyze the system before it fails (unacceptable quality) in an attempt to identify how it could fail in the future. Measures can then be put in place to prevent the failure or failures that have been anticipated. Proactive approaches include well developed qualitative methods such as HazOp (Hazard Operability) and FMEA (Failure Mode and Effects Analyses) and quantitative methods such as PRA (Probabilistic Risk Analyses) and QRA (Quantified Risk Analyses). Each of these methods have benefits and limitations.

The author has been an active protagonist and practitioner of the proactive PRA/QRA approach for more than three decades. The author believed that this approach provided an ability to forecast how systems could go bad. Very sophisticated PRA/QRA models could be developed to help foster this belief. The results from these analyses seemed to have value and to enhance his abilities to address some types of variability and uncertainty. This approach was workable as long as the author dealt with systems in which the interactions of people with the systems were minimal or minimized. However, the problem changed radically when people began to exert major influences on the quality of the systems and in many cases on the physical aspects of the systems. In this case, lack of knowledge of the physics and mechanics of the complex behaviors of people that in the future would design, construct, operate, and maintain the system defined an ‘unpredictable’ system, or certainly one with very limited predictability. The author’s analytical models addressed systems that were essentially static and mechanical. Yet the real systems were dynamic, constantly changing, and more organic (emergent) than mechanical. The analytical models generally failed to capture the complex interactions between people and the systems that they designed, constructed, operated, and maintained.

The author found that there was no way to verify the numbers that came from PRA/QRA. If the results indicated that the system was ‘acceptable’, then nothing was done. If the results indicated that the system was ‘not acceptable’, then generally equipment and hardware fixes were studied in an attempt to define a fix or fixes that would make the system acceptable or ALARP (As Low As Reasonably Practicable). When the author went to the field to compare his analytical models with what was really there, he found little resemblance between his models and what was in the field.

The author does not advocate discarding the analytical - quantitative proactive approach. He advocates using different types of proactive approaches to gain insights into how systems might fail and what might be done to keep them from failing. The marked limitations of analytical models and quantitative methods must be recognized or major damage can be done to the cause of the quality and reliability of engineered systems. The potential for engineers to be ‘hyper rational’ and attempt to extend the applicability of PRA/QRA methods beyond their limitations must be recognized and countered. On the other hand, qualitative methods (e.g., HazOp, FMEA), in the hands of qualified and properly motivated assessors (both internal and external) can do much to help the causes of quality and reliability. Experience, judgment, and intuition of the assessors needs to be properly recognized, respected, and fully integrated into proactive qualitative and quantitative approaches.

Much headway has been made recently in combining the powers of qualitative methods with quantitative methods. The qualitative methods are able to more fully capture the dynamic, changing, organic, complex interactions that can not be analyzed using traditional PRA/QRA methods. Given input from the qualitative methods, the quantitative methods are able to provide numbers that can be used to assist development of judgments about when, where, and how to better achieve quality and reliability in engineered systems. But, even at this level of development, proactive risk assessment and management (RAM) methods are very limited in their abilities to truly provide quality and reliability in engineered systems. Other methods (e.g. interactive RAM) must be used to address the unknowable and unimaginable hazards.

It is the author’s experience in working with a wide variety of engineered systems for more than five decades, that many if not most of the important proactive developments in the quality and reliability of these systems were originated in a cooperative, trust-based venture of knowledgeable ‘facilitators’ working with seasoned veterans that have daily responsibilities for the quality of these systems. This cooperative venture includes design, construction / decommissioning, operations, and maintenance / inspection personnel. Yet, it also is the author’s experience, that many engineering and many well meaning reliability – risk analysis ‘experts’ are not developing a cooperative environment. This is very disturbing. The conduct of each operation during the life-cycle of an engineered system should be regarded as the operations of ‘families.’ Knowledgeable, trained, experienced, and sensitive outsiders can help, encourage, and assist ‘families’ to become ‘better.’ But, they can not make the families better. Families can only be changed

from within by the family members. PRA/QRA measures based on casual or superficial knowledge of a system or of an operation of that system should be regarded as tinkering. And, tinkering can have some very undesirable effects and results.

The crux of the problem with proactive PRA/QRA approaches is with the severe limitations of such approaches in their abilities to reasonably characterize human and organizational factors (HOF) and their effects on the performance of a system. PRA/QRA rely on an underlying fundamental understanding of the physics and mechanics of the processes, elements, and systems that are to be evaluated. Such understanding then allows the analyst to make projections into the future about the potential performance characteristics of the systems. And, it is here that the primary difficulties arise. There is no fundamental understanding of the physics and mechanics of the future performance – behavior characteristics of the people that will come into contact with a system and even less understanding of the future organizational influences on this behavior. One can provide very general projections of the performance of systems including the human and organizational aspects based on extensive assumptions about how things will be done, but little more. The problem is that engineers and managers start believing that the numbers represent reality.

To the author, the true value of the proactive PRA/QRA approach does not lie in its predictive abilities. The true value lies in the disciplined process PRA/QRA can provide to examine the strengths and weaknesses in systems; ***the objective is detection and not prediction***. The magnitudes of the quantitative results, if these results have been generated using reasonable models and input information, can provide insights into where and how one might implement effective processes to encourage development of acceptable quality and reliability. The primary problems that the author has with PRA/QRA is with how this method is used and what it is used to do. Frequently the results from PRA/QRA are used to justify meeting or not meeting regulatory / management targets and, in some cases not implementing clearly justified – needed improvements in the quality – reliability of an engineered system.

Perhaps the most severe limitation to proactive PRA/QRA regards 'knowability'. One can only analyze what one can know. Predictability and knowability are the foundation blocks of PRA/QRA analytical models. But, what about the unknowable and the unpredictable? Can we really convince ourselves that we can project into the future of engineered systems and perform analyses that can provide sufficient insights to enable us to implement the measures required to fully assure their quality and reliability? Or are some other processes and measures needed? This fundamental property of unknowability has some extremely important ramifications with regard to application of the ALARP principle.

The author has concern for PRA/QRA analyses that have been and are being used to define IMR (Inspection, Maintenance, Repair) programs for engineered systems. Such analyses can only address the knowable and predictable aspects that influence IMR programs. Such analyses frequently are used to justify reductions in IMR program frequencies, intensities, and costs. But what about the unknowable and unpredictable elements that influence IMR programs? We look for problems where we do not find them and we find them where we do not look for them. What about the host of major 'biases' (differences between reality and calculated results) that exert major influences on the results that come from such analyses? These elements are frequently referred to as being founded in 'gross errors'. Experience has adequately demonstrated that a very large amount, if not the majority of the defects and damages we encounter in engineered systems are not in any reasonable or practical sense 'predictable'. Other approaches (e.g. inductive information based) must be used to address the unknowable – unpredictable aspects that still must be managed in the operations of engineered systems.

Another important proactive approach that has been employed in engineering systems comes from the field of ergonomics; the art and science of interfacing people with the systems that they design, construct, operate, and maintain. This approach is fundamentally one that focuses on a proactive reduction in the likelihood of malfunctions that develop at people – hardware interfaces (American Society for Testing and Materials 1995). Recent experience has adequately demonstrated that configuration of *people friendly* interfaces with the other system components including procedures, environments, hardware, structure, and most recently, organizations (macro-ergonomics) can do much to help assure that desirable and acceptable quality and reliability in engineered systems are realized.

Experience has shown that one of the most important proactive strategies is that of creating robust – damage tolerant and fail-safe (intrinsically safe) systems. Engineers frequently have called this characteristic redundancy. But, we now understand that robustness requires much more than redundancy. Robustness in the structure, operating team, and organizational components of systems has been shown to be derived from four primary elements: 1) configuration, 2) ductility, 3) excess capacity, and 4) appropriate correlation (Bea 2000a, 2001). The elements are configured so that back-ups are provided for conditions in which the system may be defective or damaged; they are configured so that the full potential capacities of the elements can be developed. Configuration can involve redundancy, but it also involves many other aspects of the geometry and layout (topology) so that the structure, hardware, or organization can perform acceptably even though defective and damaged.

Ductility is the ability (and willingness) to carry overloads and shift the overloads to other parts of the system without loss of basic functionality. Excess capacity is provision of the ability of under-loaded elements in the system to carry abnormal demands or loads. Intrinsically safe systems are those that fail in ways that do not compromise the basic safety characteristics of the system; following a failure, the system can continue to be safely operated until repairs and/or modifications can be made. Appropriate correlation refers to how the various components in the system relate to and with each other. In a *series-type* element system (failure of one element leads to failure of the system), high degrees of correlation are desirable to help prevent the rogue element or elements that do not have desirable robustness characteristics. In a *parallel-type* element system (failure of system requires failure of all of the elements), low degrees of correlation are desirable to assure independence (requisite variety) in the elements so that low robustness elements do not lead to undesirable performance.

The most important proactive approach to help achieve acceptable quality and reliability in engineered systems is that of creation and maintenance of Higher Reliability Organizations (HRO). In HRO reduction in error occurrence is accomplished by the following:

- **Command by exception or negation,**
- **Redundancy,**
- **Procedures and rules,**
- **Training,**
- **Appropriate rewards and punishment**
- **Ability of management to "see the big picture".**

Command by exception (management by exception) refers to management activity in which authority is pushed to the lower levels of the organization by managers who constantly monitor the behavior of their subordinates. Decision making responsibility is allowed to migrate to the persons with the most expertise to make the decision when unfamiliar situations arise (employee empowerment). Redundancy involves people, procedures, and hardware. It involves numerous individuals who serve as redundant decision makers. There are multiple hardware components that will permit the system to function when one of the components fails.

Procedures that are correct, accurate, complete, well organized, well documented, and are not excessively complex are an important part of HRO. Adherence to the rules is emphasized as a way to prevent errors, unless the rules themselves contribute to error. HRO develop constant and high quality programs of training. Training in the conduct of normal and abnormal activities is mandatory to avoid errors. Establishment of appropriate rewards and punishment that are consistent with the organizational goals is critical. Lastly, Roberts and Roberts, et al. define HRO organizational structure as one that allows key decision makers to understand the big picture. These decision makers with the big picture perceive the important developing situations, properly integrate them, and then develop high reliability responses.

In recent organizational research has provided support for the following five hypotheses regarding HRO:

- Risk mitigating organizations will have **extensive process auditing procedures**. Process auditing is an established system for ongoing checks designed to spot expected as well as unexpected safety problems. Safety drills would be included in this category as would be equipment testing. Follow ups on problems revealed in prior audits are a critical part of this function.
- Risk mitigating organizations will have **reward systems that encourage risk mitigating behavior** on the part of the organization, its members, and constituents. The reward system is the payoff that an individual or organization gets for behaving one way or another. It is concerned with reducing risky behavior.
- Risk mitigating organizations will have **quality standards that meet or exceed the referent standard** of quality in the industry.
- Risk mitigating organizations **will correctly assess the risk associated with the given problem or situation**. Two elements of risk perception are involved. One is whether or not there was any knowledge that risk existed at all. The second is if there was knowledge that risk existed, the extent to which it was acknowledged appropriately or minimized.
- Risk mitigating organizations will have a **strong command and control system** consisting of five elements**: a) migrating decision making, b) redundancy, c) rules and procedures, d) training, and e) senior management has the big picture**.

Review of the literature and studies of HRO indicate that organizing in effective HRO's is characterized by:

- **Preoccupation with failure** – any and all failures are regarded as insights on the health of a system, thorough analyses of near-failures, generalize (not localize) failures, encourage self-reporting of errors, and understand the liabilities of successes.
- **Reluctance to simplify interpretations** – regard simplifications as potentially dangerous because they limit both the precautions people take and the number of undesired consequences they envision, respect what they do not know, match external complexities with internal complexities (requisite variety), diverse checks and balances, encourage a divergence in analytical perspectives among members of an organization (it is the divergence, not the commonalties, that hold the key to detecting anomalies).
- **Sensitivity to operations** – construct and maintain a cognitive map that allows them to integrate diverse inputs into a single picture of the overall situation and status (situational awareness, 'having the bubble'), people act thinkingly and with heed, redundancy involving cross checks, doubts that precautions are sufficient, and wariness about claimed levels of competence, exhibit extraordinary sensitivity to the incipient overloading of any one of it members, sensemaking.
- **Commitment to resilience** – capacity to cope with unanticipated dangers after they have become manifest, continuous management of fluctuations, prepare for inevitable surprises by expanding the general knowledge, technical facility, and command over resources, formal support for improvisation (capability to recombine actions in repertoire into novel successful combinations), and simultaneously believe and doubt their past experience.
- **Under-specification of structures** – avoid the adoption of orderly procedures to reduce error that often spreads them around, avoid higher level errors that tend to pick up and combine with lower level errors that make them harder to comprehend and more interactively complex, gain flexibility by enacting moments of organized anarchy, loosen specification of who is the important decision maker in order to allow decision making to migrate along with problems (migrating decision making), move in the direction of a garbage can structure in which problems, solutions, decision makers, and choice opportunities are independent streams flowing through a system that become linked by their arrival and departure times and by any structural constraints that affect which problems, solutions and decision makers have access to which opportunities.

The other side of this coin are LRO (Lower Reliability Organizations). It has been observed that these LRO are characterized by a focus on success rather than failure, and efficiency rather than reliability. In non-HRO the cognitive infrastructure is underdeveloped, failures are localized rather than generalized, and highly specified structures and processes are put in place that develop inertial blind spots that allow failures to cumulate and produce catastrophic outcomes. Efficient organizations practice stable activity patterns and

unpredictable cognitive processes that often result in errors; they do the same things in the face of changing events, these changes go undetected because people are rushed, distracted, careless, or ignorant. In LRO expensive and inefficient learning and diversity in problem solving are not welcomed. Information, particularly 'bad' or 'useless' information is not actively sought, failures are not taken as learning lessons, and new ideas are rejected. Communications are regarded as wasteful and hence the sharing of information and interpretations between individuals is stymied. Divergent views are discouraged, so that there is a narrow set of assumptions that sensitize it to a narrow variety of inputs.

In LRO success breeds confidence and fantasy, managers attribute success to themselves, rather than to luck, and they trust procedures to keep them appraised of developing problems. Under the assumption that success demonstrates competence, non-HRO drift into complacency, inattention, and habituated routines which they often justify with the argument that they are eliminating unnecessary effort and redundancy. Often down-sizing and out-sourcing are used to further the drives of efficiency and insensitivity is developed to overloading and its effects on judgment and performance. Redundancy is eliminated or reduced in the same drive resulting in elimination of cross checks, assumption that precautions and existing levels of training and experience are sufficient, and dependence on claimed levels of competence. With outsourcing, it is now the supplier, not the buyer, that must become preoccupied with failure. But, the supplier is preoccupied with success, not failure, and because of low-bid contracting, often is concerned with the lowest possible cost success. The buyer now becomes more mindless and if novel forms of failure are possible, then the loss of a preoccupation with failure makes the buyer more vulnerable to failure. LRO tend to lean toward anticipation of 'expected surprises,' risk aversion, and planned defenses against foreseeable

Reactive Approaches

The reactive approach is based on analysis of the failure or near failures (incidents, near-misses) of a system. An attempt is to made to understand the reasons for the failure or near-failures, and then to put measures in place to prevent future failures of the system. The field of worker safety has largely developed from application of this approach.

This attention to accidents, near-misses, and incidents is clearly warranted. Studies have indicated that generally there are about 100+ incidents, 10 to 100 near-misses, to every accident. In some cases, the incidents and near-misses can give early warnings of potential degradation in the safety of the system. The incidents and near-misses, if well understood and communicated provide important clues as to how the system operators are able to rescue their systems, returning them to a safe state, and to potential degradation in the inherent safety characteristics of the system. We have come to understand that responses to accidents and incidents can reveal much more about maintaining adequate quality and reliability than responses associated with successes.

Well developed guidelines have been developed for investigating incidents and performing audits or assessments associated with near-misses and accidents. These guidelines indicate that the attitudes and beliefs of the involved organizations are critical in developing successful reactive processes and systems, particularly doing away with 'blame and shame' cultures and practices. It is further observed that many if not most systems focus on 'technical causes' including equipment and hardware. Human – system failures are treated in a cursory manner and often from a safety engineering perspective that has a focus on outcomes of errors (e.g. inattention, lack of motivation) and statistical data (e.g. lost-time accidents).

Most important, most reactive processes completely ignore the organizational malfunctions that are critically important in contributing to and compounding the initiating events that lead to accidents. Finding 'well documented' failures is more the exception than the rule. Most accident investigation procedures and processes have been seriously flawed. The qualifications, experience, and motivations of the accident assessors are critical; as are the processes that are used to investigate, assess, and document the factors and events that developed during the accident. A wide variety of biases 'infect' the investigation processes and investigators (e.g. confirmational bias, organizational bias, reductive bias).

In the author's direct involvement with several recent major failures of engineered systems (casualties whose total cost exceeds U.S. $1 billions each), the most complete information develops during legal, regulatory induced, and insurance investigation proceedings. Many of these failures are 'quiet.' Fires and explosions are 'noisy' and frequently attract media, regulatory, and public attention. Quiet failures on the other hand are not noisy; in fact, many times overt attempts are made to 'keep them quiet.' These quiet failures frequently are developed during the design and/or construction phases.

The author recently has worked on two major quiet failures that involved international EPC (Engineering, Procurement, Construction) project failures that developed during construction. A third major failure involved an EPCO (add Operation) project that failed when the system was not able to develop the quality and reliability that had been contracted for. In both of these cases, the initial 'knee jerk' reaction was to direct the blame at 'engineering errors' and a contended 'lack of meeting the engineering standard of practice.' Upon further extensive background development (taking 2 and 3 years of legal proceedings), the issues shifted from the engineering 'operating teams' to the 'organizational and management' issues. Even though 'partnering' was a primary theme of the formation of the contractors and contracting, in fact partnering was a myth. Even though ISO (International Standards Organization) certifications were required and provided, the ISO Quality Assurance / Quality Control (QA/QC) guidelines were not followed. The international organizations involved in the work developed severe 'cultural conflicts' and communication breakdowns. Promises were made and not honored. Experienced personnel were promised and not provided ('bait and switch'). There was a continually recurring theme of trying to get something / everything for nothing or next to nothing. As ultimately judged in the courts, these three failures were firmly rooted in organizational malfunctions, not engineering malfunctions. The problem with most legal proceedings is that it is very rare that the results are made public. Thus, the insights important to the engineering profession is largely lost, and in some cases, seriously distorted.

As the result of studying more than 600 ‘well documented’ major failures of engineered systems, some interesting insights have been developed:

- Approximately 80% of the major failures are directly due to HOF and the ‘errors’ that develop as a result of these factors (‘exherent’); only about 20% can be regarded as being ‘natural’ or ‘inherent’ (represent residual risk).
- Of the 80 % of the major failures that are due to HOF, about 80% of these occur during operations and maintenance activities; frequently, the maintenance activities interact with the operations activities in an undesirable way.
- Of the failures due to HOF that occur during operations and maintenance, more than half (50%) of these can be traced back to seriously flawed engineering design; engineered systems may be designed according to ‘accepted industry standards’ and yet are seriously flawed due to limitations and imperfections that are embedded in the industry standards and/or how they are used; engineered systems are designed that can not be built, operated, and maintained as originally intended; modifications are made ‘in the field’ in an attempt to make the structure workable, and in the process additional flaws or ‘bugs’ can be introduced. Thus, during operations and maintenance phases, operations personnel are faced with a seriously deficient or a defective structure that can not be operated and maintained as intended.
- The accident development process can be organized into three categories of events: 1) initiating, 2) contributing, and 3) propagating. The dominant initiating events are developed by ‘operators’ performing erroneous acts of commission or interfacing with the hardware – structure components that have ‘embedded pathogens’ that are activated by such acts of commission (about 80%); the other initiating events are acts or developments involving acts of omissions. The dominant contributing events are organizational; these contributors act directly to encourage or ‘cause’ the initiating events. In the same way, the dominant propagating events are also organizational; these propagators are generally responsible for allowing the initiating events to unfold into an accident. A taxonomy (classification system) will be developed for this malfunctions later in this paper. It is also important to note that these same organizational aspects very frequently are responsible for development of ‘near-misses’ that do not unfold into accidents.
- Most accidents involve never to be exactly repeated sequences of events and multiple breakdowns or malfunctions in the components that comprise an engineered system. These events are frequently dubbed ‘incredible’ or ‘impossible.’ After accidents, it is observed that if only one of the protective ‘barriers’ had not been breached, then the accident would not have occurred. Experience has adequately shown that it is extremely difficult, if not impossible to accurately recreate the time sequence of the events that actually took place during the period leading to the accident. Unknowable complexities generally pervade this process because detailed information on the accident development is not available. Hindsight and confirmational bias are common as are distorted recollections. Stories told from a variety of viewpoints involved in the development of an accident seem to be the best way currently available to capture the richness of the factors, elements, and processes that unfold in the development of an accident.
- The discriminating difference between ‘major’ and ‘not-so-major’ accidents involves the ‘energy’ released by and / or expended on the accident. Not-so-major accidents generally involve only a few people, only a few malfunctions or breakdowns, and only small amounts of energy that frequently is reflected in the not-so-major direct and indirect, short-term and long-term ‘costs’ associated with the accident. Major accidents are characterized with the involvement of many people and their organizations, a multitude of malfunctions or breakdowns, and the release and / or expenditure of major amounts of energy; this seems to be because it is only through the organization that so many individuals become involved and the access provided to the major sources of this energy. Frequently, the organization will construct ‘barriers’ to prevent the accident causation to be traced in this direction. In addition, until recently, the legal process has focused on the ‘proximate causes’ in accidents; there have been some major exceptions to this focus recently, and the major roles of organizational malfunctions in accident causation have been recognized in court. It is important to realize that the not-so-major accidents, if repeated very frequently, can lead to major losses.

A primary objective of incident reporting systems is to identify recurring trends from the large numbers of incidents with relatively minor outcomes. The primary objective of near-miss systems is to learn lessons (good and bad) from operational experiences. Near-misses have the potential for providing more information about the causes of serious accidents than accident information systems. Near-misses potentially include information on how the human operators have successfully returned their systems to safe-states. These lessons and insights should be reinforced to better equip operators to maintain the quality of their systems in the face of unpredictable and unimaginable unraveling of their systems.

Root cause analysis is generally interpreted to apply to systems that are concerned with detailed investigations of accidents with major consequences. The author has a fundamental objection to root cause analysis because of the implication that there is a single cause at the root of the accident (reductive bias). This is rarely the case. This is an attempt to simplify what is generally a very complex set of interactions and factors, and in this attempt, the lessons that could be learned from the accident are frequently lost. Important elements in a root cause analysis includes an investigation procedure based on a model of accident causation. A systematic framework is needed so that the right issues are addressed during the investigation. There are high priority requirements for comprehensiveness and consistency. The comprehensiveness needs to be based on a systems approach that includes error tendencies, error inducing environments, multiple causations, latent factors and causes, and organizational influences. The focus should be on a model of the system factors so that error reduction measures and strategies can be identified. The requirement for consistency is particularly important if the results from multiple accident analyses are to be useful for evaluating trends in underlying causes over time.

There is no shortage of methods to provide a basis for detailed analysis and reporting of incidents, near-misses, and accidents. The primary challenge is to determine how such methods can be introduced into the life-cycle RAM of engineered systems and how their long-term support can be developed (business incentives).

Inspections during construction, operation, and maintenance are a key element in reactive RAM approaches. Thus, development of IMR (Inspection, Maintenance, Repair) programs is a key element in development of reactive management of the quality and reliability of engineered systems. Deductive methods involving mechanics based PRA/QRA techniques have been highly developed. These techniques focus on 'predictable' damage that is focused primarily on durability; fatigue and corrosion degradations. Inductive methods involving discovery of defects and damage are focused primarily on 'unpredictable' elements that are due primarily to unanticipated HOE such as weld flaws, fit-up or alignment defects, dropped objects, ineffective corrosion protection, and collisions. Reliability Center Maintenance (RCM) approaches have been developed and are continuing be developed to help address both predictable and unpredictable damage and defects. Some very significant forward strides have been made in development and implementation of life-cycle IMR database analysis and communications systems. But, due to expense and cost concerns, and unwillingness or inability of the organization to integrate such systems into their business systems, much of this progress has been short lived.

The reactive approach has some important limitations. It is not often that one can truly understand the causes of accidents. If one does not understand the true causes, how can one expect to put the right measures in place to prevent future accidents? Further, if the causes of accidents represent an almost never to be repeated collusion of complex actions and events, then how can one expect to use this approach to prevent future accidents? Further, the usual reaction to accidents has been to attempt to put in place hardware and equipment that will help prevent the next accident. Attempts to use equipment and hardware to fix what are basic HOF problems generally have not proven to be effective. It has been observed that progressive application of the reactive approach can lead to decreasing the accepted 'safe' operating space for operating personnel through increased formal procedures to the point where the operators have to violate the formal procedures to operate the system.

Interactive Approaches

Experience with the quality and reliability of engineered systems indicates that there is a third important approach to achieving quality and reliability that needs to be recognized and further developed. Until recently, it was contended that there were only proactive and reactive approaches. The third approach is interactive (real-time) RAM in which danger or hazards builds up in a system and it is necessary to actively intervene with the system to return it to an acceptable quality and reliability state. ***This approach is based on the contention that many aspects that influence or determine the failure of engineered systems in the future are fundamentally unpredictable*** **and** ***unknowable.*** These are the incredible, unbelievable, complex sequences of events and developments that unravel a system until it fails. We want to be able to assess and manage these evolving disintegrations. This approach is based on providing systems (including the human operators) that have enhanced abilities to rescue themselves. This approach is based on the observation that people more frequently return systems to safe states than they do to unsafe states that result in accidents.

Engineers can have important influences on the abilities of people to rescue systems and on the abilities of the systems to be rescued by providing adequate measures to support and protect the operating personnel and the system components that are essential to their operations. Quality assurance and quality control is an example of the real-time approach. QA is done before the activity, but QC is conducted during the activity. The objective of the QC is to be sure that what was intended is actually being carried out.

Two fundamental approaches to improving interactive RAM performance are: 1) providing people support, and 2) providing system support. People support strategies include such things as selecting personnel well suited to address challenges to acceptable performance, and then training them so they possess the required skills and knowledge. Re-training is important to maintain skills and achieve vigilance. The cognitive skills developed for interactive RAM degrade rapidly if they are not maintained and used.

Interactive RAM teams should be developed that have the requisite variety to recognize and manage the challenges to quality and reliability and have developed teamwork processes so the necessary awareness, skills and knowledge are mobilized when they are needed. Auditing, training, and re-training are needed to help maintain and hone skills, improve knowledge, and maintain readiness. Interactive RAM teams need to be trained in problem 'divide and conquer' strategies that preserve situational awareness through organization of strategic and tactical commands and utilization of 'expert task performance' (specialists) teams. Interactive RAM teams need to be provided with practical and adaptable strategies and plans that can serve as useful 'templates' in helping manage each unique crisis. These templates help reduce the amount and intensity of cognitive processing that is required to manage the challenges to quality and reliability.

Improved system support includes factors such as improved maintenance of the necessary critical equipment and procedures so they are workable and available as the system developments unfold. Data systems and communications systems are needed to provide and maintain accurate, relevant, and timely information in 'chunks' that can be recognized, evaluated, and managed. Adequate 'safe haven' measures need to be provided to allow interactive RAM teams to recognize and manage the challenges without major concerns for their well being. Hardware and structure systems need to be provided to slow the escalation of the hazards, and re-stabilize the system.

One would think that improved interactive RAM system support would be highly developed by engineers. This does not seem to be the case. A few practitioners recognize its importance, but generally it has not been incorporated into general engineering practice or guidelines. Systems that are intentionally designed to be stabilizing (when pushed to their limits, they tend to become more stable) and robust (sufficient damage and defect tolerance) are not usual. Some provisions have been made to develop systems that slow the progression of some system degradations.

Effective early warning systems and 'status' information and communication systems have not received the attention they deserve in providing system support for interactive RAM. Systems need to be designed to clearly and calmly indicate when they are nearing the edges of safe performance. Once these edges are passed, multiple barriers need to be in place to slow further degradation and there should be warnings of the breaching of these barriers. More work in this area is definitely needed.

Combined Approaches

The results of the experience and work on which this paper is based clearly indicate that a combination of proactive, reactive, and interactive approaches should be used to improve the quality and reliability of engineered systems. Each of these approaches has its strengths and weaknesses and their strengths need to exploited. The results of this work also clearly that in most cases, these approaches are not being used as well as they could be used.

In many instances, the reactive approach has resulted in development extensive rules and regulations that have become so cumbersome that they either are not used or are not used properly. Systems are more normally operated by informal local operating procedures than by following the book. Accident investigations frequently have turned into 'witch hunts' many times with the sole purpose of 'killing the victims.' Management can mobilize the power to stop accident investigations with identification of the proximate causes and actors. Due to critical flaws in the accident investigation and recording processes, accident databases frequently fail to properly or reasonably capture the essence of how accidents develop or are caused. Near-miss incidents have not received nearly the attention that they should.

In many instances, the proactive approach has developed into a quantitative paper chase that has not yielded the benefits that it could yield. Numbers have been taken to represent the realities of future quality and reliability. Insights about how one might defend the system against unpredictable and unanticipated developments are lost in the complexities of the analyses. Experts are brought in to inspect and analyze the system and many times these experts do not possess the requisite experience or insights about how the system can unravel and fail. The experts are empowered and the system operators are 'depowered.' Fixes are general hardware oriented. Rarely do the HOF aspects receive any direct or extensive attention. Frequently, the attitude is 'this is not an engineering problem, it is a management problem' (or at least, someone else's problem).

In general, the interactive approach has not received the attention that it deserves. In some 'non-engineering' communities it has received extensive attention. These communities are those that daily must confront crises or the potential for crises. These crises all involve unpredictable and unknowable situations. Many of the communities have learned how to in most cases turn crises into successes. This research has not disclosed one instance in which the interactive approach has been used to address HOF in design engineering activities. Rarely has it been used in operations.

Conclusions

It should be apparent to all engineers and managers that HOF are of fundamental importance in development of engineered systems that will have acceptable and desirable quality and reliability during their life cycles. Engineers and managers alike have fundamental responsibilities to address HOF as an integral part of the life-cycle processes intended to develop and maintain adequate quality and reliability in engineered systems. It should also be apparent to all concerned with the quality and reliability of engineered systems that organizations (industrial and regulatory) have pervasive influences on the assessment and management of threats to the quality and reliability of such systems. Management's drives for greater productivity and efficiency need to be tempered with the need to provide sufficient protections to assure adequate quality and reliability.

The threats to adequate quality and reliability in engineered systems generally emerge slowly. It is this slow emergence that generally masks the development of the threats to quality and reliability. Often, the participants do not recognize the emerging problems and hazards. They become risk habituated and loose their wariness. Often, emerging threats not clearly recognized because the goals of quality and reliability are subjugated to the goals of production and profitability. This is a problem, because there must be profitability to have the necessary resources to achieve quality and reliability. Perhaps, with present high costs of lack of quality and reliability, these two goals are not in conflict. Quality and reliability can help lead to production and profitability. One must adopt a long term view to achieve the goals of quality and reliability, and one must wait on production and profitability to follow. However, often we are tempted for today, not tomorrow.

The second important thing that we have learned about RAM to help achieve management desirable quality and reliability is organizing the 'right stuff' for the 'right job.' This is much more than job design. It is selecting those able to perform the daily tasks of the job within the daily organization required to perform that job. Yet, these people must be able to re-organize and re-deploy themselves and their resources as the pace of the job changes from daily to unusual (it's improv time!). Given most systems, they must be team players. This is no place for 'super stars' or 'aces.' The demands for highly developed cognitive talents and skills is great for successful crisis management teams. In its elegant simplicity, Crew Resource Management has much to offer in helping identify, train, and maintain the right stuff. If properly selected, trained and motivated, even 'pick-up ball teams' can be successful design engineering teams.

The final part of the 20+ year stream of research and development on which this paper is based addressed the issues associated with implementation. A case-based reasoning study of seven organizations that had tried implementation for a significant period of time identified five key attributes associated with successful implementation:

1) **Cognizance** – of the threats to quality and reliability,

2) **Capabilities** – to address the HOF and HRO aspects to improve quality and reliability,
3) **Commitment** – to a continuing process of improvement of the HOF and HRO aspects,
4) **Culture** – to bring into balance the pressures of productivity and protection and to realize trust and integrity, and
5) **Counting** – financial and social, positive and negative, ongoing incentives to achieve adequate and desirable quality and reliability.

It is interesting to note that of the seven organizations that tried implementation, only two succeeded. It is obvious that this is not an easy challenge, and that at the present time, failure is more the rule than success. It is also interesting to note that the two organizations that succeeded recently have shown signs of 'backsliding.' Organizational – management evolution has resulted in a degradation in the awareness of what had been accomplished and why it had been accomplished. The pressures of doing something 'new,' downsizing, outsourcing, merging, and other measures to achieve higher short-term profitability have resulted in cutbacks in the means and measures that had been successfully implemented to reduce the costs associated with lack of adequate and acceptable quality and reliability. Perhaps, all organizations are destined to continually struggle for the balance in production and protection, and accidents represent a map of that struggle to succeed and survive.

Robert Bea
Center for Catastrophic Risk Management
University of California Berkeley
bea@ce.berkeley.edu