THE SOCIETY OF NAVAL ARCHITECTS AND MARINE ENGINEERS
and
THE SHIP STRUCTURE COMMITTEE

Paper presented at the Ship Structure Symposium '96
Sheraton National Hotel, Arlington, Virginia, November 18-20, 1996

A Decision Analysis Framework for Assessing Human and Organizational Error in the Marine Industries

Duane E. Boniface[1], Robert G. Bea[2]

[1]Lieutenant, U.S. Coast Guard Office of Marine Safety and Environmental Protection, Washington, D.C.
[2]Professor, Department of Civil & Environmental Engineering, University of California at Berkeley

Abstract

The marine industry has come under increased pressure in recent years to reduce the risks of accidents in both U.S. national and international waters as a results of disasters such as the HERALD OF FREE ENTERPRISE. This heightened emphasis has led to the identification of a common cause for the majority of accidents; human and organizational error. In this paper, the authors outline a decision making process for marine managers and engineers to assess the risks of human and organizational error in their operations, and to develop and prioritize appropriate countermeasures. The process is identified as HERROS: Human Error Risk Reduction Operating System.

1. Introduction

As a result of such disasters as the VALDEZ, the Ro-Ro ferries HERALD OF FREE ENTERPRISE and ESTONIA and other mishaps, the marine industry today is subject to tight internal and external scrutiny. At the core of these accidents and indeed, for approximately 85% of all major transportation mishaps is the human element. The import of human involvement in complex technological systems has been well established. However, until recently, few practical measures to address and attempt to control the risks associated with human and organizational error have been available. Research recently concluded by the authors has sought to provide the industry with tools to assess the risks of and countermeasures for human and organizational error [1, 2]. This project is one part of a long-term research program to learn how to improve the management of human and organizational factors in the marine industries. This program is conducted under the auspices of the *Marine Technology and Management Group* (College of Engineering and Haas School of Business) at the University of California at Berkeley.

In this paper, we describe the decision making process that could be used by marine managers (mates, masters, fleet and port managers, regulators) and engineers to assess the risks associated with the operations under their purview and to develop and prioritize potential countermeasures. We begin by outlining a generic risk based decision making system.

2. Generic Decision Support Process

Risk based decision making processes have been the subject of a great deal of interest in industry. Such processes provide an ability to encode and incorporate uncertainties inherent in today's highly complex and changing marine systems. Statistical decision making provides a process to help ensure that optimal decisions are reached, consistent with the goals and perceptions of those involved as well as all available information. Here, it is important to note that "optimal" decisions are not necessarily those that achieve the best outcome (which is a result of chance as much as decision making skill), but rather one that is most appropriate for the information, values, and goals for a particular situation. On average and over time, these decisions should provide the best outcomes more often than not. Furthermore, as stated previously, use of a risk based system allows for decisions to be made time and again, by different decision makers, which are consistent with the stated values of the particular organization.

For our research, we developed a generic decision support system, which was then adapted and applied to human and organizational error in the marine industry. To do so, we relied heavily on the procedures established by Howard and Matheson [3] (Fig. 1), and as modified and expanded by Ashley [4]. We will briefly describe these generic procedures before outlining their use in our framework.

2.1 The Decision Analysis Cycle

As shown in Fig. 1, the decision analysis cycle is broken down into seven major stages: 1) collection of prior information, 2) the deterministic phase, 3) the probabilistic phase, 4) the informational phase, 5) the decision phase, 6) information gathering, and 7) action. With this framework, the analyst has a powerful tool to frame the problem and ensure that all pertinent steps are taken.

2.2 Initial Information Collection

The first step in any decision analysis is the collection and analysis of all pertinent background information. One of the keys to proper decision analysis is to identify and incorporate any and all pertinent information, no matter how uncertain. To do so, review of the background literature (journals, newspapers, etc.) as well as unstructured and semi-structured interviews are used to develop the background in the subject matter and requisite details of the problem. As accurate a picture of the problem as possible should be developed at this earliest stage in order to ensure focus and direction for all future work. Time invested in this stage typically has great payoffs later in the analysis, as the resultant familiarity with the problem allows for better targeting of the analysis process.

2.3 Deterministic Phase

The next step in the analysis, the deterministic phase, is one whereby the decision problem is framed and the structural model developed for the analysis. Table 1 [4] identifies the steps involved with the deterministic phase of the analysis.

Table 1 Deterministic phase steps

Define and bound the decision problem
Identify alternatives
Establish the outcomes
Select decision variables and state variables
Build a structural model of the system
Specify time preferences
Eliminate dominated alternatives
Measure sensitivity to identify crucial state variables

In bounding the decision, the analyst specifies the focus of the problem (what decision must be made). With this, identification of the perceived and new alternatives can be made, and potential outcomes determined for sets of alternatives. Here, outcomes are defined as "whatever the decision-maker would like to know in retrospect to determine how the problem came out" [3].

The selection of system variables is an iterative process, whereby factors affecting and describing the outcome are defined, screened for importance, and redefined. System variables are sub-divided into state and decision variables, in order to differentiate between their control. While state variables, being determined by the environment of the problem, are not controllable, decision variables are the set of factors under the decision-maker's command. The focus of any decision analysis, therefore, is on the decision variables, as those are the only ones which can be affected by the decision maker.

Once the decision variables have been identified, construction of the model may begin. The system structural modeling approach that we utilized was the influence diagram. This approach facilitated interpretation by our subject matter experts (who did not know decision analysis techniques). The key feature of the influence diagram is its graphical illustration of the interrelationship between variables which, when constructed properly, can greatly aid in the understanding of the problem at hand. Fig. 2 [5] shows the six basic elements used in an influence diagram. Rules for creating and working with influence diagrams are included in McNamee & Celona [5], Bodily [6], Schachter [7] and elsewhere, and will not be repeated here.

Upon selection of the system variables, and completion of the framework step, the model is almost complete. All that remains is the selection of a model for the value measure and determination of the decision maker's time preference. In the former, one measure (e.g., profit, total cost, etc.) is selected based upon the variables involved and the particulars of the situation. This measure should be clear and readily discernible, in order to allow quantification. Finally, the time preference is encoded (through interviews typically) to depict the decision maker's willingness to wait for payoffs, and other similar considerations.

Once the pertinent information has been collected and classified, an initial sensitivity analysis is conducted to identify those decision variables which impact the specified outcome the most. These variables will then be carried through to the probabilistic phase for further screening and study. To run the sensitivity analysis, variables are fixed at their nominal values, then individually ranged from their high to low extremes, and the impact on the value measure noted. As mentioned previously, those variables which have the greatest impact (change the value measure the most) will be carried forward to future phases. These variables are known as aleatory variables, while the remaining factors are known as fixated (as they are held at their nominal values in future analyses). As noted by Howard and Matheson [3], if there is indecision concerning whether a decision variable(s) are worthy of consideration as aleatory variables, it is better to keep it (them) for future analysis, as they can always be screened out later. With this prioritized and screened set of aleatory variables, then, the probabilistic phase is begun.

2.4 Probabilistic Phase

In the probabilistic phase, the uncertainty in value and worth due to that of the aleatory variables is determined. To do so requires the series of steps as shown in Table 2 and described briefly below.

Table 2 Probabilistic Phase

Encode uncertainty on aleatory variables
Encode risk preference
Develop worth lotteries and certainty equivalent
Measure stochastic sensitivity
Measure risk sensitivity

As shown, the first step in the probabilistic phase is to determine the probability distributions (Probability Density Functions, PDF) for the aleatory variables. To do so, the analyst must make use of historical data (when available) and subjective judgment. While the analysis of historical data is relatively straightforward (especially given the proliferation of statistical analysis software available), encoding probability distributions from subjective judgment is more problematic. To do so, the analyst must take steps to resist any and all biases held by the subject matter expert (particularly the central bias commonly encountered) in order to get the most accurate portrayal of the PDF. Methods such as the probability wheel, interval technique, direct response mode and others can be used to facilitate the interview process, which in itself is a highly structured and detailed procedure. These and other procedures are described in detail in Spetzler and von Holstein [8] and elsewhere, and will not be duplicated here.

Once the PDFs for the aleatory variables are determined, the decision maker's risk preference must be identified (unless one variable dominates the others). To do so, the interview process is used to determine the decision maker's preference (utility) for any value of worth (value measure). By comparing the decision maker's preference between a lottery and a certain event, their level of risk acceptance or aversion can be identified, and the utility curve derived. Procedures for this process can be found in Howard and Matheson [3].

Analysis within the probabilistic phase is done to further screen and prioritize amongst the aleatory variables, in order to further focus the analysis. The stochastic sensitivity of all aleatory variables is determined in a manner similar to the sensitivity analysis conducted in the deterministic phase. Here again, one variable is swept along its potential values (as determined by its probability distribution), while the others are held at their appropriate conditional probabilities. If the resultant change in the value measure is large, the variable is kept for further analysis. Otherwise, the variable can be fixated and removed from further analysis. By doing so for all variables, the number of "important" variables can be further prioritized, and the next, informational phase begun with a reduced set of variables.

2.5 Informational and Decision Phases

The purpose of the informational phase is primarily to determine if it is worthwhile making further expenditures in research and information gathering efforts before making a decision. Here, the value of information (thus eliminating or at least reducing the inherent uncertainty in each of the variables) is determined. Decision trees are set up to define the decision (whether or not to conduct tests/research) and Bayes' rule (Equation 1) used to incorporate prior and posterior (after testing) information.

$$P[B|A] = \frac{[P[A|B] \times P[B]}{P[A]} \qquad (1)$$

The interpretation of this rule is as follows: A is the experimental evidence, B is the cause, P[B] is the prior or original estimate without evidence, and P[B|A] is the posterior estimate given the evidence.

Those variables which exhibit a high economic sensitivity to the information gathering (large return on investment ratio) can be considered for potential research. The final decision on whether or not to conduct the additional information gathering is made based upon these economic sensitivities and on the feasibility of conducting the research. If the gathering of additional information is desired, the analysis process is re-run, as shown in Fig. 1. This process may iterate for several (if not many) iterations, until the gaining of additional information is determined to cost more than it is worth. Once this stage is reached, the decision is made based upon the available information, and action taken accordingly.

3. Human Error Risk Reduction Operating System (HERROS)

Having assembled a generic decision making process, we modified it to fit the specifics of the marine industry and adapted it to our taxonomy and conceptual model for human and organizational error. We have identified this system as the Human Error risk Reduction Operating System (HERROS) [1,2].

The ultimate goal of any human error risk reduction system should not be the determination of a specific probability of failure (or survival, for that matter). Instead, the primary purpose of any such effort should be to reduce the risks associated with all potential error modes (human and otherwise) without sacrificing "economic" viability or social imperatives. Kuo [12] illustrated this goal as shown in Fig. 3.

Here, the objective is shown as reducing intolerable hazards (high probability and/or consequence) to a tolerable level. This acknowledges not only the inevitability of risk in any endeavor (particularly a technological one such as maritime commerce), but also the impossibility of ever attaining a degree of safety that is truly "acceptable" to all concerned.

While detailed analytical methods will be outlined herein, it is important to note the value of even the simplest of studies. Simple awareness of the basic tenets of human and organizational error theory as well as the general countermeasures described in Boniface [1], Bea [9] and Boniface and Bea [2] should go far in the improvement of safety.

In the following, a general risk analyses methodology will be developed and described. Within this framework, an increasingly detailed approach will be advanced for analysis of the failure modes, with particular attention paid to HOE. Data sources (including qualified subjective judgment) will be discussed, and initial rates supplied (for both general and specific application). Finally, integration of situational factors, error recovery, and countermeasures will be discussed.

The analysis of risk in any endeavor can rapidly become an incredibly complicated task. Those performing the analyses can rather easily become lost in all the details of the potential scenarios. To prevent such an occurrence, utilization of a screening methodology as shown in Fig. 4 should be used. Within each analysis stage, the investigation loop shown in Fig. 1 should be utilized. Each subsequent stage of the analysis would utilize more involved risk methodologies, which will require more time and expertise to implement. Each high risk item (for which management schemes were not readily available) from the previous assessment phase would be studied in greater detail to determine the sub-components which contribute the most to the identified risk. This added "cost" would be offset somewhat by the reduction in the number of risks provided by the progressive screening. The percentages given in Fig. 4 are representative of the suggested breakdown of time and effort between the stages of the analysis. Specific details concerning the methodologies and performance requirements at each stage will be discussed next.

3.1 Coarse Qualitative Stage

In this stage of the analysis, all potential mishap scenarios are identified, categorized, and ranked according to their risk. Here (and throughout this paper) risk is defined as the product of the likelihood of a mishap occurring and the expected costs associated with that mishap. This can be done by using simple screening methods. We recommend use of risk encoding techniques as discussed in the previous section. Details for the utilization of these techniques can be found in Howard and Matheson [3] and elsewhere, and will not be repeated here.

Here, those scenarios which present the greatest potential for system (vessel) failure (and for which risk management schemes are not readily available) can be identified for further analysis in the next stage; detailed qualitative analysis. The primary purposes of this stage are to set the scope (i.e., identify the problem and all pertinent variables), and to screen out non-critical variables. At this stage, the potential scenarios need be no more detailed than those shown in Table 3.

Table 3 Damage Scenarios

i =	Scenario
1	Grounding
2	Collision
3	Allision
4	Structural
5	Fire
6	Stability

However, if possible (and it likely will be with scenario analysis), determination of the most likely causes of failure/damage (categorized in Table 4) should be attempted. The final category in Table 4 (common mode failures) will largely be ignored in this report. Although these do play an important role in maritime mishaps, the refinement of HOE analysis is not such that these can be accounted for very readily.

With the level of detail used in coarse qualitative steps, successive analyses are greatly simplified. As stated by Lucas and Embrey [16], the primary functions of qualitative modeling are to: assist in accident investigation by allowing causes (and therefore countermeasures) to be determined, support design efforts by allowing for better incorporation of human performance factors, and to aid in risk assessment, thereby allowing identification and targeting of those processes most prone to mishap.

Table 4 Vessel Failure Causes

j=	Cause	Description-Failures Due To:
1	Human/ Organizational	Human and organizational factors
2	Material/Supply	Inadequate materials and supplies
3	Equipment/ Facilities	Machinery, piping, electrical & electronic, hydraulic systems, etc. (excluding errors due to design flaws)
4	Structural	Hull strength, defects and deterioration
5	Environmental	Natural forces and effects
6	Common Modes	Combinations of one or more of the above

3.2 Detailed Qualitative

In the detailed qualitative analysis, the most concerning damage scenario(s) are analyzed as to damage likelihood and consequence through means of such measures as reliability and maintainability screening and failure mode and effect analyses (FMEA). Here, failure modes are ranked by the number of component failures (or mishaps) required for system failure, and the potential consequence magnitude. At the end of this phase, a risk matrix should be developed to illustrate the risks involved in the process under consideration, as illustrated in Fig. 5.

Obviously, this will create a natural prioritization, with those scenarios falling higher and to the right being more pressing (e.g., Scenario 6 receives highest priority and Scenario 7 receives the lowest). Within the equivalent risk bands, prioritization is given by the sensitivities of each risk to countermeasures and the cost to implement countermeasures for that risk. Risks with a higher return (risk reduction) on investment (countermeasure implementation cost) should be selected first with such a band. The goal, therefore, is to shift these scenarios down (lower their potential consequence) and/or to the left (decreasing their likelihood) through use of countermeasures strategies as discussed previously.

There have been numerous methodologies developed and implemented for a detailed qualitative analysis. Of these, reliability and maintainability (R&M) screening and failure mode and effect analyses (FMEA) have been selected due to their widespread and relatively straightforward use.

Table 5 Reliability and Maintainability Screening Criteria

Number	Criteria	Description	Rating = 1	Rating = 4
1	Reliability	Based on Mean Time Between Failures (MTBF)	MTBF > 10^6 hrs	MTBF < 10^3 hrs
2	Maintainability	Based on Mean Time To Repair (MTTR)	MTTR < 1 hr.	MTTR > 1 day
3	Safety Effect	Based on impact of personnel or equipment safety (ignores redundancy)	Negligible Hazard	Potential Catastrophic
4	Hazard Class	Effect of item failure on personnel/system	Safe	Catastrophic
5	Shutdown Level	Degree of shutdown impact	Local Shutdown	Evacuate (Global)
6	Production Effect	Effect on operations	No Effect	Total Shutdown
7	Redundancy	Degree of reserve capacity	100%	No Redundancy
8	Complexity	Level of intricacy	Simple System	Very Complex
9	Environment	Sensitivity to operating environment	No Effect	Catastrophic
10	Contamination	Sensitivity to process contamination	No Effect	Catastrophic

Table 6 Example Reliability and Maintainability Worksheet

Equipment or Sub-System RO-RO FERRY BOW DOOR	1	2	3	4	5	6	7	8	9	10	HS*	SUM
Locking Devices	4	4	4	3	3	4	3	2	3	2	22	32
Scantlings	4	4	4	4	3	2	4	1	3	3	21	32
Hinge Mechanism	4	4	3	3	3	4	3	2	3	3	21	32
Gasket	4	4	3	3	3	2	4	1	3	3	19	30
Status Indicators	4	4	3	3	2	2	2	3	2	3	18	28
Positioning Horns	4	4	3	4	3	2	4	1	3	3	20	31

*HS = hazard source

Reliability and maintainability screening is a qualitative risk assessment tool applicable to initial screening efforts. Here, organizational and system flowcharts and process and instrumentation diagrams are used to identify critical subsystems for more detailed qualitative studies. The screening is carried out for all major subsystems, with a rating scale (1 = good, 4 = bad) being used to rank the sub-system in terms of basic criteria. These criteria are listed in Table 5 [10], and an example assessment for the bow door of a Ro-Ro ferry is shown in Table 6. As illustrated in Table 6, ratings between 1 and 4 are obtained by interpolating between the two bounds. The hazard score as noted is the sum of the scores for the first six criteria. The R&M screening process is extremely valuable for the following, detailed qualitative analysis phase, as the major functional relationships are identified and illustrated, thus making the failure mode and effect analysis much easier.

In the failure mode and effect analysis, the critical system (vessel) components and failure modes (from Table 2) are subdivided into the task level. The functional and reliability block diagrams developed previously are then used to direct the analyses. Fig. 6 summarizes the basic steps involved in the procedure, which will subsequently be discussed in greater detail.

The first step in the analysis (system definition) is perhaps the most critical step in the FMEA process. Here, the analyst(s) ensures that they have sufficient understanding of the system to perform the analysis. Walk-throughs, interviews, etc. are used to familiarize the analyst with the system components (including personnel) and their interactions, as well as to call upon the expertise of those directly involved in day-to-day processes. The final system description should include an account of normal and potential abnormal and emergency operations, as well as maintenance intervals, demand (operational and environmental) description, etc.

Once the analyst(s) are familiar with the details of the system, block diagrams are then constructed to graphically illustrate the functional and reliability relationships. As such, each system will be broken down into sub-systems, assemblies (or tasks) and components (or steps). The degree of detail utilized will depend upon the criticality of the system and the resources available. An example functional block diagram is shown in Fig. 7.

With the completion of the block diagrams, the process of documenting assumptions should be started. All premises made throughout the duration of the analysis should be meticulously documented in order to delineate the limitations of the derived results and show directions for future study. This list of assumptions should include those factors which are not varied throughout the analysis (instead being taken at some nominal value), simplifications made, etc.

Utilizing the framework provided by block diagrams, data can then be collected describing system performance and failures. As shown in Fig. 8 (which is typical of an FMEA worksheet) [after 10 and 11], the minimum data required are failure modes and modal failure and detection rates. Exacting detail is not required at this phase, as it is the relative risk rankings that will be utilized, not the actual probabilities of failure. The goal of the data collection phase is to provide the information required by the FMEA worksheets. The failure rates used may be subjective or can be obtained from historical data. If historical data are used, the base error rates obtained must be modified to take into account the impact of the performance shaping factors as applicable. If individual component/step failure data are not available, qualified assessments may be obtained using the mapping in Table 7 and 8 [10].

Table 9 shows the ranking schema for use in the ranking of detection measures [10]. Once entries for all factors (failure modes for given steps/components) and all three value measures have been made, determination of the relative ranking is performed by multiplying all three values together. The resulting values are indicative of the relative importance of that step/component; the highest value indicates the most critical factor, the lowest the least important. If desired, these values may be normalized by the largest value for easier comparison.

FMEA worksheets, as shown, are merely the expanded and tabulated version of the block diagrams. Here again, the level of detail utilized will be dictated by the expected risks and the available resources. It is also important to note that Fig. 8 is only one example of an FMEA worksheet. The format for this worksheet can and should be modified to meet the particulars of the situation. However, no matter what format is chosen, the failure modes, rates and criticality's must be explicitly identified to allow ranking.

At this and every stage, a "big picture" review is extremely important. By looking for trends and commonalties between failure modes, high payoff countermeasures (i.e., those which reduce the consequence and/or likelihood of two or more scenarios/modes) can be identified for implementation. An example of such a countermeasure would be the implementation of a radar training program. Such a curriculum, if designed and implemented correctly, could reduce the likelihood's of collisions, allisions, and groundings, thus paying off in three ways for the same investment.

It is also important to note here that the analysis may be stopped at any point if it is determined that further, more detailed analysis may not be necessary or desirable (as

Table 7 Qualitative Severity Ranking Scales

Seriousness of Effect	Rating	Effect Realized at
Minor: Unreasonable to expect that this will impact any performance/evolution.	1	Local scene (e.g., workspace)
Low: Effect is limited only to one phase of life cycle (e.g., procurement of supplies, determination of navigational information, etc.). Results only impact own operations.	2 3 4 5	Primary operation; Secondary operation; Tertiary operation; or Other evolutions
Moderate: Effect is throughout life cycle chain (e.g., error made at design level which increases construction difficulty, operability, etc.)	6 7	Next life cycle phase Other than next phase
High: Effect impacts others in immediate operating area (e.g., vessels in harbor, etc.)	8 9	Local level State level
Extreme: Effect is injury/harm to anyone involved, widespread damage to the environment, etc.	10	National level

Table 8 Qualitative Guidelines For Occurrence Rating

Likelihood of Occurrence	Rating
Impossible or has never occurred previously	2
Remotely possible and similar events may have occurred previously	4
Has previously occurred rarely	6
Has previously occurred occasionally	8
Has previously occurred frequently	10

Table 9 Detection Ranking Index

Type of Detection	Ranking
No inspection or testing	10
Supplier certification- qualitative (conforms to spec)	9
Supplier certification- quantitative (date supplied)	8
First component use	8
100% manual inspection & testing- subjective	7
100% manual inspection & testing- objective	7
Incoming inspection- supplierÕs statistical data analyzed	7
Sample testing (normal sampling plans)	7
In-process attribute frequency sampling	7
Sample testing (tightened sampling plans)	6
100% automated inspection and testing	6
Statistical process control/statistical analysis	5
First/last piece layout	5
Continuous part/process monitoring	2

illustrated in Fig. 4). If risks (likelihood * consequences) are low, if one or two scenarios dominate, and/or if one set of countermeasures is clearly favorable (due to situational circumstances, projected effectiveness, etc.), the analyses need go no further. However, often this will not often be the case.

At the next phase of the analysis (coarse quantitative), the most important failure modes from the detailed qualitative stage are further subdivided into tasks and analyzed. Here, quantitative estimates are introduced to facilitate further ranking and screening. By using reliability methods, specific tasks (e.g., loading calculations, loading and discharge operations, etc.) will be highlighted as critical. These tasks can then be selected for further analysis, or can be targeted with specific countermeasures. To supplement these analyses, situation factors are quantified using Performance Shaping Factor (PSF) scales [1, 16,19] and used to modify base mishap rates. These analyses are carried down into the micro-task or step level to both refine the estimates on risk levels and to pinpoint the source of the high risk items. Again, the analysis can stop here, if further study is deemed to be cost inefficient.

3.3 Coarse Quantitative

In this phase of the analysis, the assessment of the error/failure costs and likelihood's initiated in the detailed qualitative stage is refined and expanded to provide a more accurate (and descriptive) account. In particular, the mean and standard deviation of the system and all levels of sub-components are obtained through detailed interviews (for subjective estimates) and/or historical data analysis (for objective estimates). With this data, fault tree analyses may be conducted to ascertain the reliability (or, conversely, the probability of failure).

If adequate historical data is available, determination of the mean (μ) and standard deviation (σ) is a relative straightforward process. Numerous computer programs exist to assist with data reduction and analysis. However, if data is inadequate (or even nonexistent), the analyst must resort to subjective estimation by experience qualified experts. Procedures for conducting the interviews required are detailed in Howard and Matheson [3], Spetzler et al. [8] and elsewhere. By using tools such as probability wheels, interval techniques, etc., the analyst can obtain cumulative distribution functions for each of the uncertain variables (likelihood's of failure and consequences, where consequences include non-monetary costs. With this description, fault trees may be developed and analyzed to determine the relative impact on system reliability.

When including HOE analysis by fault trees, care must be taken when determining the likelihood of failure. In human reliability analysis, human performance is typically taken as a binary fail-no fail system. However, in addition to our "faults," we also bring an extremely valuable ability to intervene when other modes (e.g., structural/equipment failures) go awry. In order to ascertain the true impact of human involvement, therefore, intervention should be modeled as well. However, scant direction in this field was uncovered during this research. As such, human intervention will not be explicitly modeled. Instead, the base error rate will be modified such that the situational error rate will be less than nominal for above average ratings on the PSF scales, as shown in Equations 19 through 21 [1].

Figures 9 through 15 (Figs. 11 - 15 at the end of this paper) show the general fault trees recommended for HOE analysis, in increasing levels of detail. Fig. 9 breaks the likelihood of failure for the vessel into the six failure modes listed in Table 3. Here, these failure modes are assumed to be statistically independent for notational convenience only. In actuality, common mode failures (i.e., failures involving two or more of the failure modes) would have to be addressed if the full lifetime probability of failure for all failure modes is desired.

Fig. 10 then utilizes the next level of detail (failure causes) to refine the analysis, a process continued in Fig. 11 (which looks at the failure contribution in each of the k tasks involved in that particular error mode i) and Fig. 12 (which itemizes the failure contribution from each of the m steps involved in a given task k). Figures 13 and 14 show trees for the analysis at the step level (without the possibility of error detection or correction) with and without redundancy. Finally, Fig. 15 shows an example decision problem whereby any of n countermeasures (or no countermeasure) can be implemented. Obviously, various forms and combinations of these fault trees can and should be developed to best describe the particular situation. Analysis of these trees is as discussed in Ang and Tang [30], Benjamin and Cornell [31] and elsewhere and will not be repeated here.

Equations 2 through 22 provide the corresponding general mathematical model for human and organizational analysis. While the exact relations are not (and may never be) known with certainty, research has identified some models as being generally appropriate for some of the probabilistic relationships involved. Others were adapted as part of this research to better address the marine situation. Each will be described and discussed individually in turn.

In Equation 2, the overall likelihood (probability) of failure for the vessel is shown to be the sum of the failure probabilities for each of the six failure modes, as shown in Table 12 and Fig. 11. Here, it is assumed that the failure modes are rare events and are statistically independent. Equation 3 continues this breakdown by using the total probability rule to determine the probability of failure of a given mode i for all five causes listed in Table 3.

In this research, only the first cause, human and organizational error, is addressed. As such, Equations 4 through 22 determine the likelihood of human and organizational error in the life of the vessel. Therefore, the "j" term was dropped in this and subsequent representations. Equation 4 shows this likelihood as being the sum (over all k tasks) of each of the likelihood's of HOE in a mode i given HOE in a task k (e.g., loading determination, scantlings design, etc.) and various detection and correction rates at the task levels. In the first part of that equation, the likelihood is based upon undetected events (which were therefore assumed to go uncorrected), while in the second, the probability was based upon the likelihood of detected errors that were not corrected. P[X|Y] is read as the probability of X given the occurrence of Y. P[$\cancel{X}$] is read as the probability that the event X does not occur. P[X$\cap$Y] is read as the probability that X and Y occur (intersection of the events X and Y).

As is the case with most of the probability distributions, the relationship between the likelihood of HOE given undetected and uncorrected HOE in any mode (i), task (k) and step (m) (shown in Equation 5 and Equation 12) is unclear. No relationships between these variables were uncovered during the course of this research effort, nor was any direction available from any other researchers. Thus, the sub-levels were modeled as being statistically independent, and the relationship between levels as shown in Equation 6. This relationship was used again in Equation 12, as the connection between the likelihood of HOE in a task (k) given HOE in a step (m) was felt to be similar.

Equation 7 [27] describes the probability of detection (POD) given the occurrence of an error (the complementary event is shown in Equation 8). The probability of error detection of an error has been the subject of many studies, such as Demsetz et al [23], Nowak [24], Stewart and Melchers [25, 26] and Engelund and Rackwitz [27] among others. Demsetz' work has been related to crack detection's in vessel inspections, in order to optimize inspection, maintenance and repair (IMR) activities. Included in her report is a summary of POD curves for cracks in aircraft, as developed by Dinkeloo [29] and as shown in Fig. 16. As seen, the data appears to fit a shifted negative exponential curve; a finding also made by Stewart and Melchers [27] in their study of design review checks.

In the Melchers and Stewart study, they recommended use of shifted exponential curves (Equation 7) for checking efficiency as a function of both checking time and error magnitude, where error magnitude is as given in Equation 9. In Equation 9, the true value is given as x_m, while the value used is given by x. t is a normalized variable representing either the time for the activity or the tolerance level for magnitude of the error (as shown in Equation 10).

If personnel involved are relatively inexperienced or untrained, Melchers and Stewart recommended use of an s-shaped curve to better characterize the probabilities of detection (Fig. 17). This is the classical "learning curve" first identified in educational studies.

In terms of cost effectiveness, Stewart and Melchers felt that checking efficiencies (POD's) between 0.6 and 0.9 to be optimal [27]. Efficiencies less than 0.6 were identified as producing limited reductions in the likelihood of error, while efficiencies greater than 0.9 were seen as producing a marginally smaller reduction in the error likelihood. Additionally, they felt that two design checks should be sufficient to reduce HOE to a minimal level.

When developing the HERROS model, it was unclear what checking relationship was applicable for the marine industry. It appears that some form of relationship between the bounds displayed in Fig. 18 should hold true. Furthermore, it appears that the most significant factors impacting this relationship are the training and experience of the inspector/checker and the time allotted for the check, although fatigue and others certainly play important roles as well. Research is currently being conducted to better identify the true nature of this relationship. However, until results from these studies are available, use of the model given in Equation 7 [25] is recommended and will be used in this paper for both the task and step levels.

Not correcting an undetected error was assumed to be a certain event, as shown in Equation 10. The probability of not correcting an error that has been detected falls into the violation category (at a level given by the seriousness of the violation). Although functional studies have not been identified, it is believed that the error rate will be most affected by the organizational, workload, impairment and personality performance shaping factors. Until more definitive relationships are known, a Poisson model will be used. This model (Equation 11) which is the same as the model for HOE (Equation 19) will be discussed in detail shortly. The likelihood of not correcting an undetected error was assumed to be 1.0.

λc is the base correction error rate. SLI is the Success Likelihood Index developed by Embrey [16] and Chien et al [17], and as modified by Zamanali et al [15]. w_n is the weighting factor for the directly acting PSF, v_n is the directly acting PSF value (from the PSF scales), w_o is the weighting factor for the indirectly acting PSF and v_o is the indirectly acting PSF value. Given the lack of information regarding the PSF effects, the w_n and w_o must be subjectively estimated for the particular situation. The distinction between directly and indirectly acting PSFs is made to address the differing degrees of impact on the system safety that they would have.

The model for the SLI was modified to allow a use of a 1 to 7 Likert Scale (instead of the 1-10 scale recommended by Embrey), which is less prone to a central bias [1]. Furthermore, adaptations have been made to account for the likelihood of performance improvement (and therefore human intervention in other error modes) by taking the difference between the observed PSF scale value (v_n) and the mean PSF scale value (which corresponds to the mean error rate), all of which is normalized by the standard deviation of the PSF scale values. The use of the normal cumulative density function (Φ) in Equation 13 allows for the modeling of performance enhancement (error rates below base rate) for situations where the PSF scale values being less than the average (lower being better). Until further research is available, it is recommended that values of 3.5 be used for both the mean (μn)and the standard deviation (σn)(the coefficient of variation shown in Table 10 is of the order of 100%) [1].

$$P_F = \sum_{i=1}^{6} P_f[\text{Mode } i]$$

Equation 2

$$P_f[\text{Mode } i] = \sum_{j=1}^{5} P_F\,[\text{Mode } i|\text{Cause } j] * P[\text{Cause } j]$$

Equation 3

$$\begin{aligned} P[HOE_i] = \sum_{k\ Tasks} &\{P[HOE_i|HOE_k \cap \overline{Detect}_k \cap \overline{Correct}_k] \\ &* P[\overline{Correct}_k|HOE_k \cap \overline{Detect}_k] * P[\overline{Detect}_k|HOE_k] * P[HOE_k]\} \\ + \sum_{k\ Tasks} &\{P[HOE_i|HOE_k \cap Detect_k \cap \overline{Correct}_k] \\ &* P[\overline{Correct}_k|HOE_k \cap Detect_k] * P[Detect_k|HOE_k] * P[HOE_k]\} \end{aligned}$$

Equation 4

$$P[HOE_i|HOE_k \cap \overline{Detect}_k \cap \overline{Correct}_k] = ???$$

Equation 5

$$P[HOE_i|HOE_k] = \sum_{k\ tasks\ in\ mode\ i} P[HOE_k]$$

Equation 6

$$P[Detect_k|HOE_k] = \frac{1}{1 + 3000 * \exp\{-m_e^{1/10} * 2.2765 * \sqrt{t}\}}$$

Equation 7

$$P[\overline{Detect}_k|HOE_k] = 1 - P[Detect_k|HOE_k]$$

Equation 8

$$m_e = \frac{|x - x_m|}{x_m}$$

Equation 9

$$t = \frac{|\text{Expected time/tolerance} - \text{Standard time/tolerance}|}{\text{Standard time/tolerance}}$$

Equation 10

$$P[\overline{Correct}_k|HOE_k \cap \overline{Detect}_k] = 1.0$$

Equation 11

$$P[\overline{Correct}_k|HOE_k \cap Detect_k] = \lambda_c * \exp\{\lambda_c * SLI*t\}$$

Equation 12

$$\begin{aligned} SLI = &\sum_{n=1}^{10} \frac{W_n}{\sum_{n=1}^{10} W_n} * \left[\Phi\left(\frac{v_n - \mu_n}{\sigma_n}\right) * \left(\frac{v_n - \mu_n}{\sigma_n}\right)\right] \\ &+ \sum_{o=1}^{10} \frac{W_o}{\sum_o W_o} * \left[\Phi\left(\frac{v_n - \mu_n}{\sigma_n}\right) * \left(\frac{v_n - \mu_n}{\sigma_n}\right)\right] \end{aligned}$$

Equation 13

$$\begin{aligned} P[HOE_k] = \sum_{m\ Steps} &P[HOE_k|HOE_m \cap \overline{Detect}_m \cap \overline{Correct}_m] \\ &* P[\overline{Correct}_m|HOE_m \cap \overline{Detect}_m] * P[\overline{Detect}_m|HOE_m] * P[HOE_m] \\ + \sum_{m\ Steps} &P[HOE_k|HOE_m \cap Detect_m \cap \overline{Correct}_m] \\ &* P[\overline{Correct}_m|HOE_m \cap Detect_m] * P[Detect_m|HOE_m] * P[HOE_m] \end{aligned}$$

Equation 14

$$P[HOE_k|HOE_m] = \sum_{m\ steps\ in\ task\ k} P[HOE_m]$$

Equation 15

$$P[\overline{Correct}_m|HOE_m \cap Detect_m] = \lambda_c * \exp\{\lambda_c * SLI*t\}$$

Equation 16

$$P[\overline{Correct}_m|HOE_m \cap \overline{Detect}_m] = 1.0$$

Equation 17

$$P[Detect_m|HOE_m] = \frac{1}{1 + 3000 * \exp\{-m_e^{1/10} * 2.2765 * \sqrt{t}\}}$$

Equation 18

$$P[\overline{Detect_m}|HOE_m] = 1 - P[Detect_m|HOE_m]$$

Equation 19

Equation 14 continues the breakdown further, down to the step or micro-task level (e.g., table look-ups in a loading calculation, monitoring equipment performance parameters, etc.). As seen, the relationships used were the same as at the task level.

The model for the likelihood of HOE in a given step (m) of task (k) and failure mode (i) was based upon work by Rackwitz [13], Ang and Yamamoto [14], and Zamanali et al [15]. In the HERROS modeling, this likelihood is as shown in Equation 20 (discrete) and 21 (continuous). Here, $_b$ represents the base error rate for the task (skill based, rule based, knowledge based, violation, etc.).

$$P[HOE_k] = \lambda_b * \exp\{\frac{\lambda_b * SLI}{t}\}$$

Equation 20

$$F[HOE_k] = \int_0^{t_1} \lambda_b * \exp\{\frac{\lambda_b * SLI}{t}\} * f(\lambda_b)\partial t$$

Equation 21

Use of Poisson processes (Equations 20 and 21) was chosen based upon work by Ang and Yamamoto [14]. Although Poisson processes require events in non-overlapping time intervals to be statistically independent (not entirely the case here), this was believed to be a relatively minor difference. Other models were reviewed, including normal [18, 19], lognormal [20], SLIM (Equation 22) [21] and others, but all were found to be inadequate. Further research in this area should shed more light on the true nature of the relationships between base error rates, PSFs and modal failure rates.

$$P_f = 1 - \exp\{a * SLI + b\}$$

Equation 22

As seen in Equations 4 through 19, the likelihood of failure in a level is conditioned upon the occurrence and nonoccurrence of errors in each sub-level. Thus, the likelihood of HOE as contributing to failure mode i is conditioned upon the likelihood of HOE in each task involved in mode i as well as the detection and correction events. (The model as developed makes the assumption that all corrections undertaken are effective for the sake of simplicity, although this is not required.) The error rate at the task level is in turn conditioned on the error likelihood in the step (or micro-task) level, as well as the detection and correction events. By handling all probabilities as conditional events, the effects of correlation between steps, tasks and modes are implicitly included, which greatly simplifies the analysis. The conditional probability of redundant steps, for varying degrees of dependency (shown graphically in Fig. 14) is shown in Table 10 [18].

Table 10 Dependency model

Level of Dependency	Actual P[m2|m1]	Approximate P[m2|m1]
Zero	P[m2]= P	P
Low	(1+19P)/20	0.05
Moderate	(1+6P)/7	0.14
High	(1+P)/2	0.5
Complete	1	1

In order to supply the data for these analyses (base error rates), extensive research was conducted into existing maritime, aviation, civil engineering and nuclear power plant databases [1].

The nuclear industry's NUCLARR human error probability (HEP) database was classified using the HERROS taxonomy [2] and analyzed to determine central tendency and dispersion measures. Although not all base error rates were able to be obtained (as they were not contained in the database), and the data presentation (mean, median, and confidence band values for approximately 1,300 data points which were already reduced) precluded all but a fairly crude analysis, the results shown in Table 11 and Fig. 19 were felt to be fairly representative of reality.

As shown in Table 11, there were significantly more data points for the skill based and rule based error rates, thus providing a stronger sense of confidence in these values. It should be noted here that the data points were derived from thousands of events and the actual number of points involved is significantly higher. Given that the results realistically need only be in terms of an order of magnitude accuracy (particularly for the coarse quantitative analysis), the results obtained should be more than suitable. Furthermore, their relative ranking (as shown in Fig. 19) appears to be correct intuitively, with the knowledge based rate being greater than the rule based rate, which is in turn greater than the knowledge based rate. The wide variation (as shown in Fig. 20) is also consistent with intuition.

Table 11 Results of NUCLARR Data Analysis

Category	Mean Rate	Coef. of Variation %	No. Data Pt.s
Skill Based	0.0269	322.24	558
Rule Based	0.116	175.81	655
Knowledge Based	0.153	168.89	73
Ingrained Violations	0.0468	93.85	4
Routine Violations	0.0768	132.06	16
Exceptional Violations	0.0219	30.79	16
Overall Violations	0.104	243.81	36
Physical-Coordination	0.0073	311.87	55

As can be seen, the level of effort for even the coarse quantitative analysis is rather substantial. It is for that reason that careful screening in the qualitative phases is so important. However, from the coarse quantitative analysis completed, a fairly accurate ranking of the risk is available. If desired, the analyses can be stopped at this stage and countermeasures developed and implemented to address those items with the most risk, thus reducing the costs involved. However, by continuing the analysis into the detailed quantitative phase a more refined and more accurate prioritization of risk countermeasures can be developed. This final phase will now be described.

3.4 Detailed Quantitative Analysis

In this phase, the most important errors (it is recommended that only 2-3 variables be analyzed here to minimize computational costs) be reviewed. Here, the errors (and countermeasures) analyzed can be at the modal, task or step level, although the latter is more likely to be the case. These errors will be further screened in this stage for their costs (in terms of net present value or similar measure) as well as their risk and sensitivities to and costs of countermeasures. In this stage, the risk assessment will take the form of a Level III Reliability Analysis, with economic analysis included as an integral part. Level III reliability analyses are covered in great detail in Mansour [22] and Bea [9], and will not be covered here. However, economic (including both monetary and non-monetary costs) and sensitivity analyses must be conducted in conjunction with these risk analyses. To begin, the process of sensitivity analysis will be outlined.

Sensitivity analyses are generally conducted in two phases; deterministic and probabilistic. In the deterministic phase, the risk analysis is conducted with all values at their median value. The expected cost obtained (by resolution of the decision/fault tree) is then the nominal or base case. One by one, the variables involved (usually payoffs/costs) are changed to their high and low values, and the percentage change in expected cost versus the base case is plotted versus the percentage change in the input variable. By doing so for all variables, it is a straightforward process to develop the deterministic sensitivity diagram, as illustrated in Table 12 and Fig. 21. Here the retrofit and newbuild options are seen as most sensitive (as indicated by the maximum slope), while the do nothing option is least sensitive (which has minimum slope). Positive slopes indicate that the variable and expected profit are positively correlated, while negative slopes show negative correlation (inverse relationship). Probabilistic or stochastic sensitivity analysis is virtually the same as the deterministic sensitivity. The only real difference between the two is that in stochastic sensitivity analysis, the probabilities are varied to their upper and lower bounds, with the percentage change in input versus percentage change in measure plotted. The resulting graph is interpreted in the exact same manner as described above. With these sensitivity tests, the degree of response of the system (vessel) to various countermeasures can be predicted and ranked. Obviously, those countermeasures which have the greatest sensitivity will be preferable, as they provide the most return on investment.

The final portion of the detailed quantitative analysis is the incorporation of the time value of money into the investigation. Knowing the firm / organizations minimum attractive rate of return, and estimating the inflation rate and costs associated with various countermeasure options, the fault tree (with countermeasures) can be re-analyzed using common dollar values. For most problems, this should not change the prioritization of countermeasures. However, if cash flows are significantly different between countermeasure options, or if long-term schemes are to be implemented, the priority ranking may change.

Table 12 Sensitivity table

Vessel Option	Low Cost	Nominal Cost	High Cost
Do Nothing	$150M	$250M	$350M
Retrofit Existing	$21M	$169M	$534M
Build New Vessel	$80M	$228M	$532 M

4. Conclusion

With this modeling system as outlined, the marine decision maker (owner, operator, regulator, etc.) may rationally allocate resources to minimize risk in their purview. Although no methodology can guarantee success, utilization of these reliability and decision analysis techniques will ensure that good decisions are made. On the average, these should prove to provide more favorable results over time.

References

1 Boniface, D. *An Analytical Methodology to Assess the Risks of and Countermeasures for Human and Organizational Error in the Marine Industry*, Master's Thesis, The University of California at Berkeley, 1996

2 Boniface, D. and Bea, R. G. "Assessing the Risks of and Countermeasures for Human and Organizational Error," 1996 SNAME Annual Meeting Proceedings, Society of Naval Architecture and Marine Engineers, Jersey City, 1996

3 Howard & Matheson. *An Introduction to Decision Analysis*, Decision Analysis Group, Stanford Research Institute, Menlo Park, 1977.

4 Ashley, D. *Project Risk Management*, CE268F Class Reader, Department of Civil & Environmental Engineering, University of California, Euclid Street Copy Central, 1996.

5 McNamee & Celona. *Decision Analysis with Supertree*, Scientific Press, San Francisco, 1990.

6 Bodily, S. *Modern Decision Making*, McGraw Hill, Inc., New York, 1985.

7 Schachter, R.D. "Evaluating Influence Diagrams," Operations Research, November 1986.

8 Spetzler & von Holstein. "Probability Encoding in Decision Analysis," ORSA-TIMS-AIEE 1972 Joint National Meeting, Atlantic City, November 1972

9 Bea, R. *Reliability Based Design Criteria for Design and Maintenance of Marine Structures*, Class Reader (CE290C), Department of Civil & Environmental Engineering, University of California, Euclid Street Copy Central, Berkeley, 1995.

10 Rao, G. "Failure Modes and Effects Analysis for Chemical Plant Processes," Proceedings Society for Engineering Risk Analysis Annual Meeting, American Society of Mechanical Engineers, New York, 1993

11 Andrews. *Reliability and Risk Assessment*, John Wiley & Sons, New York, 1993.

12 Kuo, C. "Managing the Safety of Ships and Marine Vehicles," The Royal Institute of Naval Architects, London, 1995.

13 Rackwitz, R. "Planning for Quality- Concepts and Numerical Tools,"

14 Ang, A. and Yamamoto. "Significance of Gross Errors on Reliability of Structures," Proceedings 4th International Conference on Structural Safety & Reliability, ICOSSAR 1985.

15 Zamanali, et al. "Evolutionary Enhancement of the SLIM-MAUD Method of Estimating Human Error," Transactions of the 1992 Meeting of the American Nuclear Society, American Nuclear Society, Boston, 1992.

16 Embrey. *SLIM-MAUD: An Approach to Assessing Human Error*, NUREG/CR-3518, Nuclear Regulatory Commission, Washington, DC, 1984.

17 Chien, et al. "Quantification of Human Error Rates Using a SLIM Based Approach," Proceedings of IEEE 4th Conference on Human Factors, IEEE, New York, 1989.

18 Swain and Guttman. *Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications*, NUREG/CR-1278F, Nuclear Regulatory Commission, Washington, DC, 1983.

19 Dougherty and Fragola. *Human Reliability Analysis*, John Wiley & Sons, NY, 1988.

20 Drager et al. "PROF: A Computer Code for Prediction of Operator Failures," *Human Factors & Decision Making*, Sayers et al, ed., Elsevier Applied Science, London, 1988.

21 Chapanis, A. *Human Factors in Systems Engineering*, Wiley Interscience, New York, 1996.

22 Mansour, A. *An Introduction to Structural Reliability Theory Directed at the Marine Industry*,

23 Demsetz, et al. "Inspection of Marine Structures," Presentation given at U.C. Berkeley, Berkeley, CA, 1994....could cite recent SSC report

24 Nowak, A. "Effect of Human Error on Structural Safety," American Concrete Institute Journal, Sep. 1979.

25 Melchers, R. and Stewart. "Checking Models in Structural Design," Journal of Structural Engineering, American Society of Civil Engineers, New York, June 1989.

26 Takyi. "Total Quality Management for Public Transit Systems," Transportation Quarterly, April 1983.

27 Stewart and Melchers. "Error Control in Member Design," Structural Safety, Vol VI, Elsevier Science Publishers, Amsterdam, 1989.

28 Engelund & Rackwitz. "Quality Assurance in Structural Design," *Structural Safety & Reliability*, Schueller, Shinozuka & Yao, eds., Balkema, Rotterdam, 1994.

29 Dinkeloo, 1978.

30 Ang, A. and Tang, W. *Probability Concepts in Engineering Planning and Design*, John Wiley & Sons, New York, 1975.

31 Benjamin, J. and Cornell, C. A. *Probability, Statistics and Decision for Civil Engineers*, McGraw-Hill, New York, 1970.

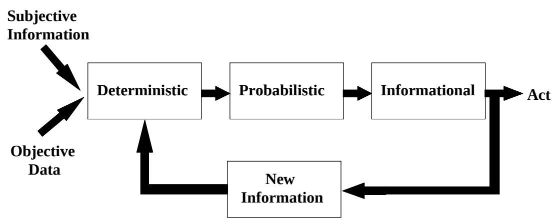


Figure 1
The Decision Analysis Cycle

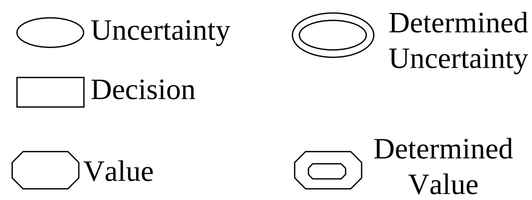


Figure 2
Basic Influence Diagram Elements

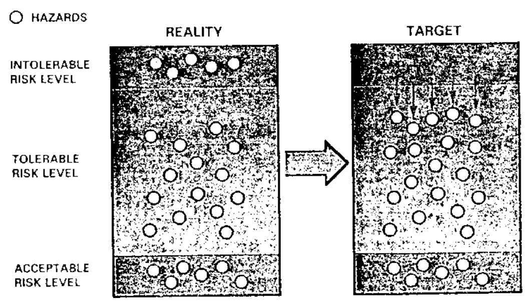


Figure 3
Objective for Risk Reduction Measures

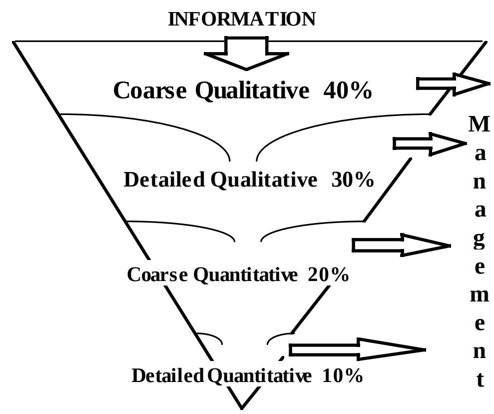


Figure 4
Screening Stages and Suggested

CONSEQUENCE			
High	Scenario 1	Scenario 4	Scenario 6
Medium	Scenario 3	Scenario 2	Scenario 5
Low	Scenario 7	Scenario 9	Scenario 8
RISK	High	Medium	Low

Figure 5
Risk Matrix

Step 1	Define thesystem to be analyzed andits required reliability performance
Step 2	Constuct functional and reliability block diagrams to illlustrate relationships between sub-components
Step 3	Note assumptions made during the analysis and define system and sub-system failure modes
Step 4	List components, identifying failure modes and approximate modal failure rates and typical effects for each
Step5 or	Complete FMEA worksheets analyzing effects of each task
Step 6	Enter severity rankings and failure rates (or ranges) on worksheets and evaluate criticality of each failuremode to sytem performance

Figure 6
FMEA Procedure Outline

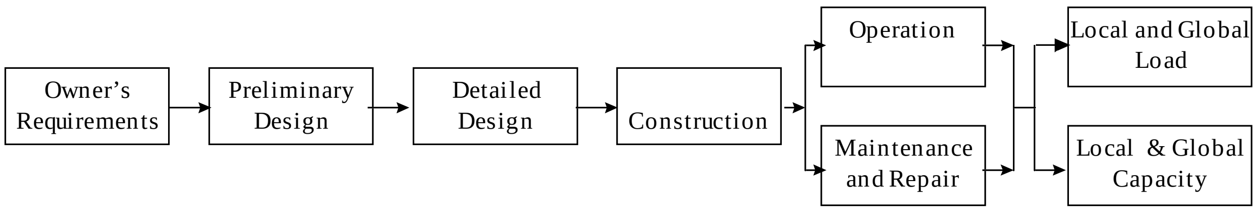
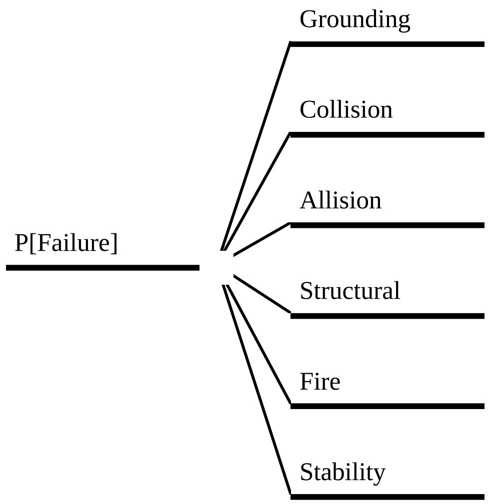


Figure 7
Ro-Ro Ferry Bow Lock Functional Block diagram

Cause	Mode	Failure Effects	Controls	OC	SV	DT	RPN
Overload	Stability	Marginal Stability	Weigh Station	8	4	4	128
Uneven Load	Stability	Marginal Stability	Weigh Station	8	3	7	168
Insecure Load	Stability	Marginal Stability	Training, Regulation	8	4	7	224
Door Not Secured	Stability	Loss of Vessel	Automated Checking	8	10	10	800
Ballasting Incomplete	Stability	Marginal Stability	None	8	3	10	240
No Course/Speed Optimization With Respect to Load	Strength	Structural Overload	None	10	6	10	600
Insufficient Structural Capacity	Strength	Structural Overload	Inspections, Plan Reviews	10	6	8	480

Figure 8
FMEA Worksheet Example



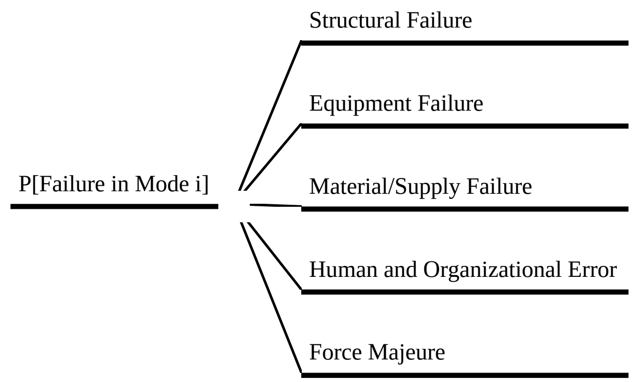


Figure 9
Failure Modes Fault Tree

Figure 10
Failure Causes Fault Tree

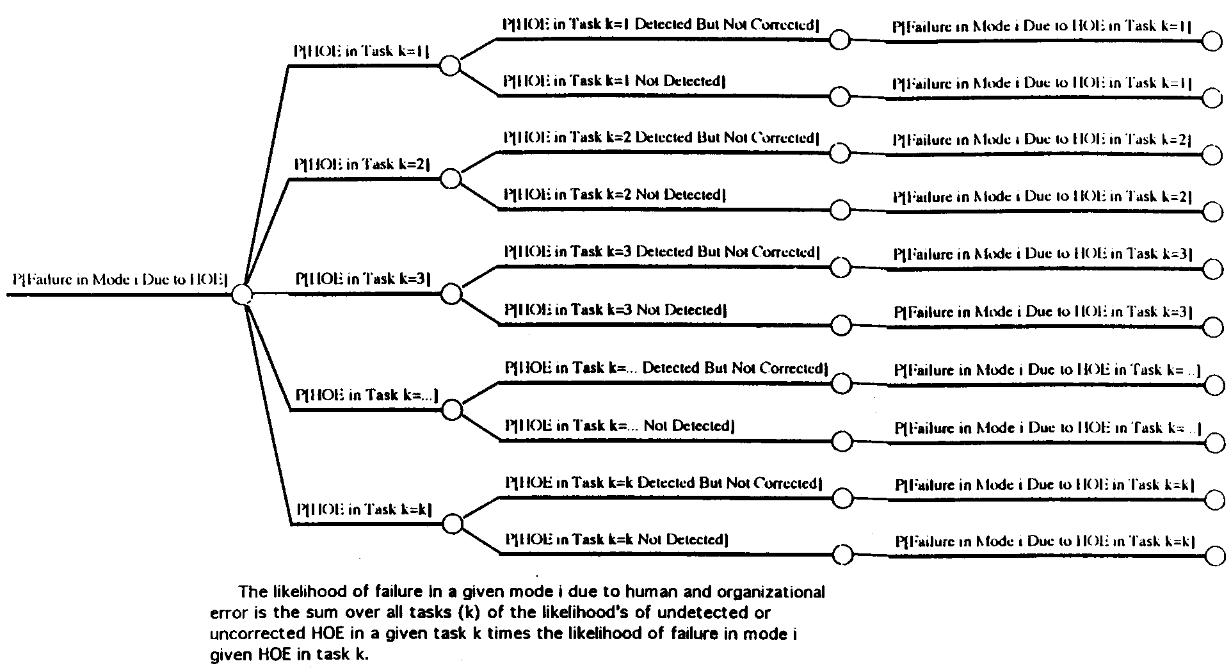
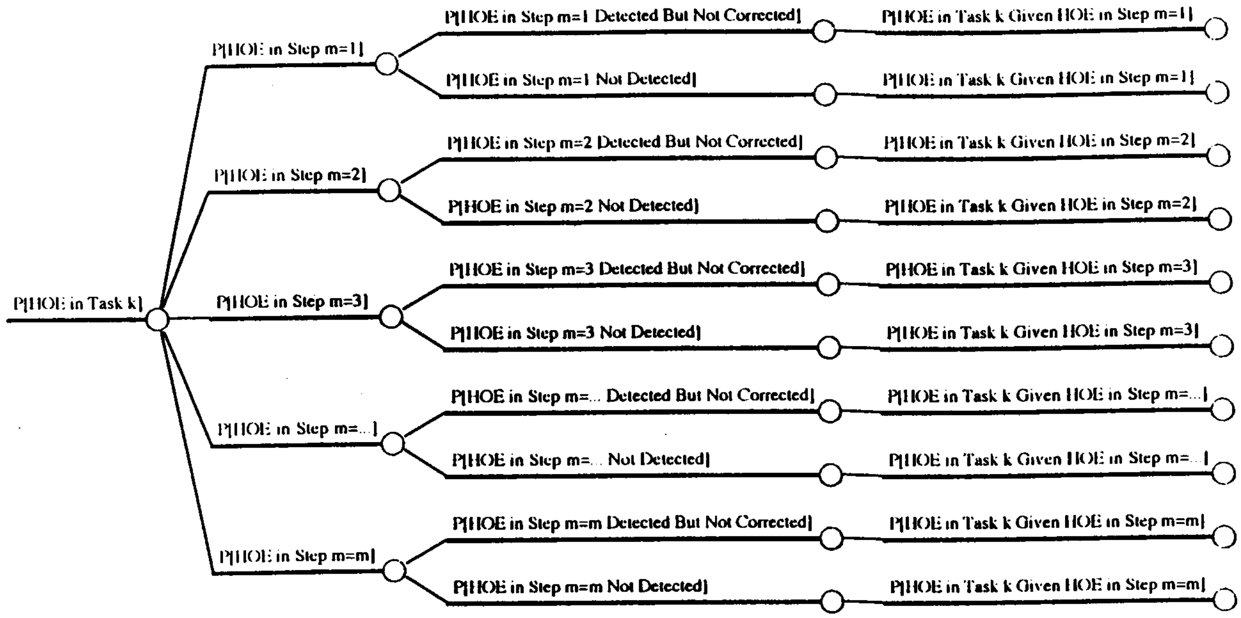
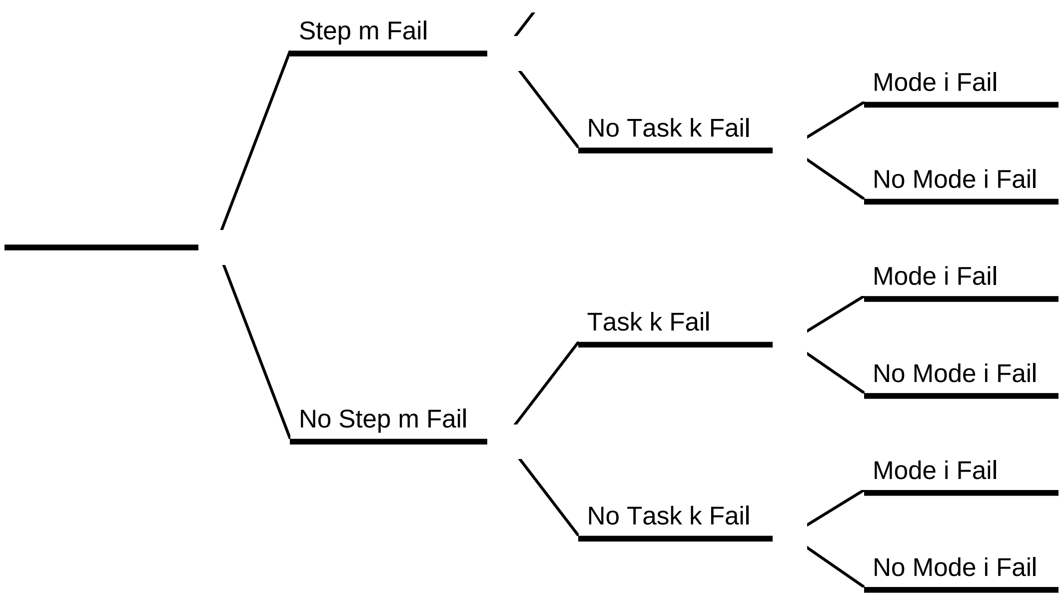


Figure 11
Task Level Fault Tree



The likelihood of HOE in a given task (k) of failure mode (m) is obtained
by taking the product of the likelihood of an undetected or uncorrected error
in a given step (m) times the likelihood of HOE in task (k) given HOE in step i.

Figure 12
Step / Component Level Fault Tree

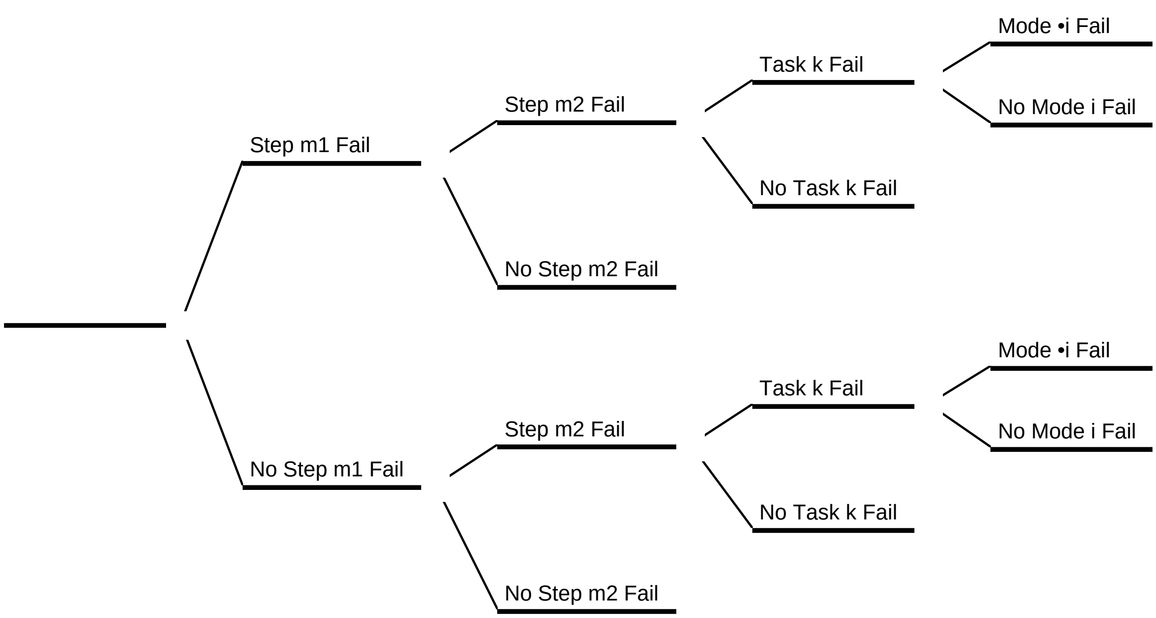


Failure analysis tree for a given step m which has •no
redundant steps, error detection or correction possibilities.

Figure 13
Simplified Step Fault Tree: No Redundancy, Detection or Correction

figure received for publication not legible

Figure 14
Countermeasure Decision Model



Failure analysis tree for a given step m1 which has a redundant step (m2), but which has no error detection or correction possibilities.

Figure 15
Simplified Step Fault Tree: with Redundancy, But No Detection or Correction

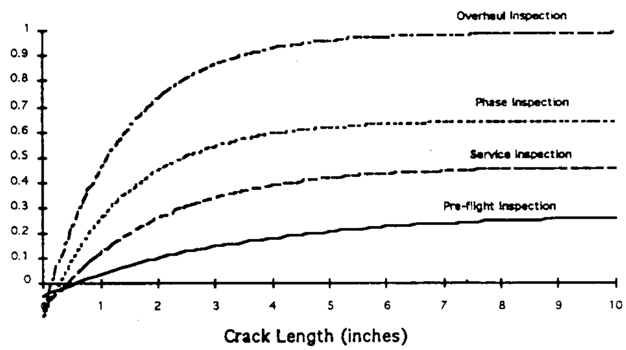


Figure 16
Aircraft Crack POD Curves

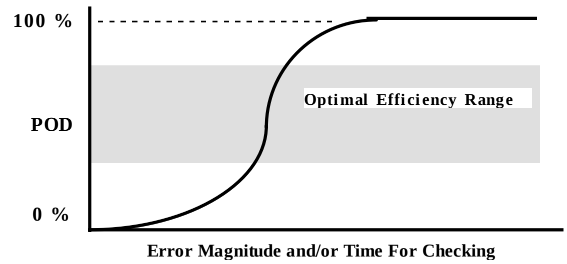


Figure 17
POD Curve for Inexperienced Personnel

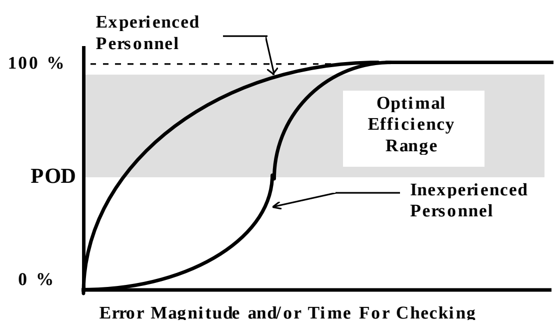


Figure 18
POD Bounds for Maritime Personnel

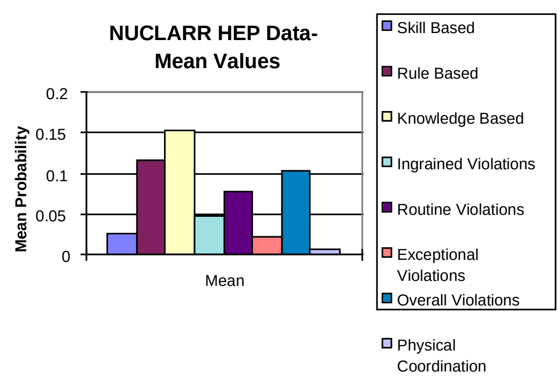


Figure 19
NUCLARR Mean Human Error Frequencies

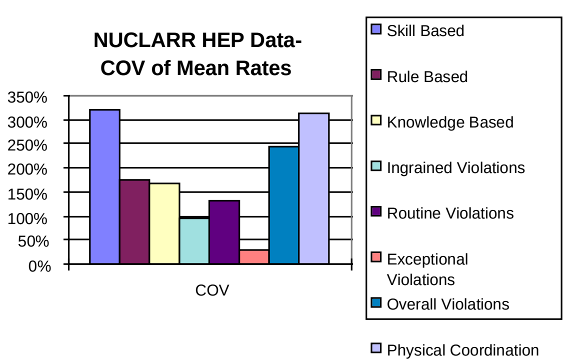


Figure 20
NUCLARR Human Error Frequency

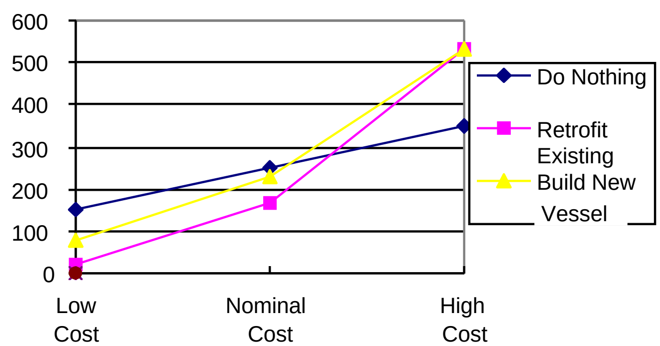


Figure 21
Sensitivity Diagram

Discussion

by Robert Sielski, Marine Board, National Academy of Engineering, National Research Council, Washington, DC

The authors have defined a process for decision making, that although at first glance may seem complicated, if implemented in the four stages from coarse qualitative to detailed quantitative as shown in Figure 4, then the process can be manageable and useful. The key as the authors have brought out, is to perform only enough analysis in order to make a decision. Such staged analysis is not always the case in application of risk analysis in decision making. For example, in the required Safety Case analysis for offshore platforms in the North Sea, typical analyses cost millions of dollars to perform because detailed qualitative analyses are performed when a decision can be made on the basis of coarse quantitative analysis because the analyst believed such detail was necessary to satisfy regulators. *How do the authors believe that such a staged approach can be implemented so that a good process for obtaining insight to key variables doesn't become a complicated process where all insight is lost?*

As I understand the authors' approach, they recommend a seven-phase cycle starting with information gathering and ending with final action as is illustrated in Figure 1. This basic cycle can be repeated in up to four stages from coarse qualitative to detailed quantitative. However, such a path does not seem reasonable considering the processes involved in each phase. For example, it is not until information gathering is completed that the first step of the deterministic phase, defining and bounding the decision problem is taken. *Doesn't it seem more reasonable to define and bound the problem before setting out to obtain all available information? How else can the investigator know if the information is relevant to the problem?*

In their discussion of the coarse quantitative phase the authors imply that the statistical analysis of data is a trivial task. They ignore such difficult questions as to the sufficiency of data, applicability of data to the problem, and bias of data. If the analysis involves rare events, then typically the interest is with the tails of probability distributions, where mean and standard deviation are of little benefit. The authors use probability of detection curves from aircraft inspection as an illustration of how such data could be used if properly developed for the marine industry. *But, can the data from aircraft be applied to marine with a few supplemental data points, or will the characteristics be entirely different?* In an ongoing risk assessment study, available data on propulsion reliability of the world-wide fleet was deemed inappropriate for the U. S. steam-powered fleet. Considerable effort was expended to obtain data because the existing data was considered as inappropriate. Bias in data can come in many ways, one of which is the human tendency not to report information that could be harmful to the reporter. *Can the authors reflect on the triviality of data analysis?*

The authors have defined a generic process for decision making that appears to apply to any situation. In the discussion, particularly of coarse quantitative analysis, several specific instances of tailoring the process to the marine industry are made. Data on the probability of crack detection is an example, although it is unclear throughout the paper how much the process has been modified to reflect the marine situation. *Will the authors please indicate what characteristics of the marine field make it sufficiently different so as to modify the process and techniques of decision making and of risk assessment for that field?*

The authors have described a process that has the potential to become complex, and that is heavily influenced by many human and organizational factors that can affect the decision making within the process. For example, the authors selected influence diagrams as the most appropriate means to model the many different problems to be encountered in the marine field. However, there are other modeling techniques, such as system simulation, that may be more appropriate or a particular situation. I have already alluded to the difficulties in analysis of data, which often arise from human and organizational considerations, such as data bias. Certainly there is considerable room for human and organizational error in the conduct of interviews to obtain data, a process to which he authors devoted only one half of a paragraph. And there are certainly human and organizational factors involved in determining if the analysis has included sufficient rigor with which to make a decision *How do the authors feel that the human and organizational error inherent in their process to analyze human and organizational error can be minimized?*

My final question reflects my personal unfamiliarity with the subject matter. The authors place considerable emphasis on Performance Shaping Factor (PSF) scales and ratings. However I found no definition or discussion of this term in the paper. *Could the authors please explain this concept?* Perhaps then I could provide more insightful comments on this excellent paper.

Author's Reply

Guidelines and protocols need to be developed that will encourage general use of the four stage process we have proposed. The goals of all concerned need to be clearly stated and agreed to; e.g., to identify potentially important challenges to the safety of systems at the simplest stage possible, and to remedy these challenges in order of the risks they present to the extent that resources can and should be made available to achieve the desired level of safety. Information systems need to be developed and implemented that will allow us to document and monitor the results of the safety measures and experiences with the process. Qualifications and training need to be developed for those that perform these assessments. These assessors need to have formal training in HOF.

Developing information for an analysis is a recursive problem. It is desirable to be as complete and comprehensive as possible at the start of the analysis. This is because the analysis can not be any more perceptive or complete than the data and information that is assembled to guide the analysis. The extent of the data and information assembled at the start of the analysis should be extensive enough to prevent misdiagnosis and misdirection of the analysis. As the analysis progresses, in all likelihood it will be desirable to gather additional information to answer the questions raised by the analysis.

The authors did not mean to imply that analysis of data is trivial. The issues raised by Dr. Sielski are correct. Our discussion of data analysis was brief due to length limitations on this paper. Our experience with data analysis involving a wide variety of marine systems indicates that very frequently the data is incomplete and in many cases, potentially misleading. The processes used to gather, validate, and encode the data frequently are seriously flawed. This is particularly true for accident and near miss data, data on and from inspections, and data gathered in the field. To the authors knowledge, there is not one complete database available for reliability and risk analyses of marine systems, including ship structures.

The means and standard deviations of probability distributions are quantities that are used to determine the shapes of these distributions, including the tails of the distributions. Some probability models have additional parameters to help shape the distributions. It is important to recognize that a probability distribution is a model. In many cases, there is no inherent reason why one probability distribution model is “correct.” The justification for a probability model should be rooted in an understanding of the physics and mechanics of the processes of concern, the ability of the model to fit the data in the regions of concern, and ease of use.

It is not likely that data from airframe inspections can be used directly in quantifying the probabilities of detection of flaws in ship structures. The conditions under which the data is gathered in these two systems is dramatically different. The inspection methods, and qualifications and training of inspection personnel are generally very different. However, the general trends from airframe inspections are useful: under the best of conditions, small flaws are difficult to detect Large flaws are easier to detect. Flaws frequently occur where they are not expected. The airframe industry has developed a variety of methods to help guide inspections. But, the most important method still used is visual. The most important aspect of airframe inspections is the inspector and the methods that are used by the inspector to disclose both anticipated and unanticipated flaws and problems. The most frequent cause of unanticipated and perhaps unanticipatable flaws are the unanticipated and unpredictable actions and inactions of people. We are working on development of inspection guidelines for ship structures that will address both anticipated and unanticipated flaws and that will take advantage of the lessons in inspections gathered in other fields such as airframe inspections and medical inspections.

Bias in data has many potential sources. Such bias can have sources in: 1) the personnel that perform the data gathering, 2) the organizations that become involved in the data gathering, encoding, and analysis, 3) the procedures used in the data gathering, encoding, and analysis, 4) the equipment used, 5) the procedures used, and 6) the environments in which the data are gathered, encoded, and analyzed. Data gathering, encoding, analysis, and recording is not a trivial undertaking. It is because many of these factors have been ignored that there are so few databases and data on marine and ship structures that can be used in reliability and risk assessments. Much more use of experimental design methods is needed to help guide gathering data on ship structures. Long-term commitments are needed by industry and government to allow sufficient data to be developed. The experience of the commercial and military airframe industries have much to offer the designers of database systems and data gathering for advanced engineering of ship structures.

Influence diagram methods are only one method that can be used to help formulate and express the logic of risk and reliability analyses. Event tree, fault tree, logic diagram, and other similar methods have been developed. Each has its applications, powers, and limitations. The method should be chosen that best fits the problem being addressed. A number of computer based analysis programs have been developed that facilitate applications. It is important that the user understand the theory behind these aids and use them appropriately.

The authors believe that decision making processes should be tailored to those that make the decisions. Important corporate, local, and national cultural and experience factors influence how decisions are made. The metrics from such processes must address the unique concerns and considerations associated with each decision problem. We are learning that many informed and experienced decision makers frequently do not believe nor rely on engineering probabilities nor expected risk values derived from such analyses. Decision makers know that risk analyses do not have the ability to forecast the future in any reliable way. The variance of the risk frequently has more meaning, but unfortunately this variance is frequently so large that it too can loose its meaning to decision makers. Scenarios associated with plausible upper and lower bounds on the risk estimates are frequently more meaningful. Perhaps it is the decision making *process* that is most important. It is the communications, information accessed, deliberation, thought, and discipline that is used that may be the most important aspect. The measures of results from reliability and risk analyses should not become the objective of performing such analyses. It is the process that holds the most promise for helping achieve serviceable, durable, safe, and compatible ship structures.

Human and organizational error can be minimized in performing a process by a variety of means that include qualifying and training those that perform the process, monitoring and independent checking and verification, development of meaningful and detailed guidelines and examples for performing the analyses, and providing incentives and motivations for performing and conducting high quality analysis processes.

Performance Shaping Factors (PSFs) are a logical way to reflect how "things" affect the different "base" rates of human errors. General information is available to help guide quantifications of base rates of some types of errors. For example, given normal conditions, an experienced person will mis-dial a telephone about once in one hundred times. However, given influences that can have sources in personal, organizations, procedures, equipment, and environments, this rate can be dramatically increased. PSFs attempt to capture these influences with multipliers on the base rates. Given a combination of stressors like distracting noise, poor lighting, a rush to dial, an unusual telephone, and an unusual dialing procedure, the rate probably rises to unity. If one wants to quantify the likelihoods of human errors in task performance, then such PSFs can be very useful.

The authors would like to thank Dr. Sielski for his detailed review of our paper, a very stimulating discussion of some of its critical aspects, and his thoughtful questions.