PERGAMON

Safety Science 30 (1998) 235–247

SAFETY SCIENCE

Using system simulation to model the impact of human error in a maritime system

J.R. Harrald [a], T.A. Mazzuchi [a], J. Spahn [a], R. Van Dorp [a], J. Merrick [a], S. Shrestha [a], Martha Grabowski [b]

[a] School of Engineering and Applied Science, The George Washington University, Washington, D.C. 20522, USA

[b] Department of Decision Sciences and Engineering Systems, Rensselaer Polytechnic Institute, Troy, NY, 12180 USA

Abstract

Human error is cited as the predominant cause of transportation accidents. This paper describes the modeling of human error related accident event sequences in a risk assessment of maritime oil transportation in Prince William Sound, Alaska. The risk analysts were confronted with incomplete and misleading data that made it difficult to use theoretical frameworks. They were required, therefore, to make significant modeling assumptions in order to produce valid and useful results. A two stage human error framwork was developed for the Prince William Sound Risk Assessment based on interviews with maritime experts. Conditional probabilities implied by this framework were elicited from system experts (tanker masters, mates, engineers, and state pilots) and used within a dynamic simulation to produce the risk analysis base case results discussed. The ability to quantify the effectiveness of proposed risk reduction interventions aimed at reducing human and organizational error were limited by the level of detail described by the taxonomy of human error.

Keywords: Prince William Sound; Human error; Maritime accidents; Expert judgement; Risk assessment; Risk management

1. Human error and risk assessment

The *Torrey Canyon*, the *Argo Merchant*, the *Exxon Valdez*, the *Tuo Hai-Tenyo Maru*, the *Morris T. Berman*, the *Sea Empress* ... the sequence of internationally publicized oil spill producing maritime accidents caused by one or more obvious human errors continues. Human error is the primary cause of most transportation-related accidents according to all research studies and investigation reports. Prevention programs must, therefore, effectively reduce the incidence of human error. But where should these programs be targeted? How effective will they be? How

0925-7535/98/$—see front matter
PII: S0925-7535(98)00048-4

much do we know about the types and causes of human error that result in maritime accidents?

Risk assessment tools and techniques provide partial answers to these questions through reasonable quantitative estimates of the linkages in causal chains leading to an accident involving human error. Quantitative models, however, require very specific data that enable the description of the phenomena and relationships of interest. Note the circularity of these statements: risk analysis is used to predict the potential for accidents due to human error, but a knowledge of the linkages between human error and accidents is essential to building risk models. Analysts modeling human error in maritime risk assessments are confronted with misleading and incomplete data. Significant modeling assumptions and a great deal of skill and effort in obtaining relevant data are required to portray human error in a manner that produces useful and accurate results. Risk managers and other stakeholders must pay attention to the hidden assumptions and selected data that drives the risk models. The recently completed Prince William Sound Risk Assessment (PWSRA) (Harrold et al., 1996) used innovative techniques to capture and model human error related accident sequences. This paper discusses these techniques, the assumptions that were made in order to use incomplete data, the use of expert judgment as a source of system specific data, and the limitations imposed on the analysis by the inability to fully model human error.

The Prince William Sound (PWS) risk assessment project was a joint project of Det Norske Veritas (DNV), Rensselaer Polytechnic Institute (RPI), and The George Washington University (GWU). The project was directed by a steering committee comprised of the Prince William Sound Shipping Companies (ARCO, Sea River, British Petroleum, Chevron, and Tesoro Alaksa), the Prince William Sound Regional Citizens Advisory Committee (RCAC), the Alaska Department of Environmental Conservation (ADEC), and the US Coast Guard (USCG). The PWS risk assessment project had three primary objectives:

1. to identify and evaluate the risks of oil transportation in PWS,
2. to identify, evaluate, and rank proposed risk reduction measures, and
3. to develop a risk management plan and risk management tools that can be used to support a risk management program.

The involvement of all Trans Alaska Pipeline (TAPS) shippers, the RCAC, Alyeska, USCG, and the ADEC in the management of the project provided the study team with unique access to individuals and information, and ensured that all viewpoints are considered in the analysis.

The risk of an accident is defined as the product of the probability of occurrence of the accident and the consequences of that accident. An accident is an event that has adverse consequences (e.g. injury, loss of life, economic loss, and environmental damage). Seven accident types were considered in the PWS risk assessment: collision, powered grounding, drift grounding, foundering, structural failure, allision, and fire/explosion. An incident is defined as a triggering event, such as a human error or a mechanical failure that creates an unsafe condition that may result in an accident. As shown in Fig. 1, the assessment framework consists of a six stage causal

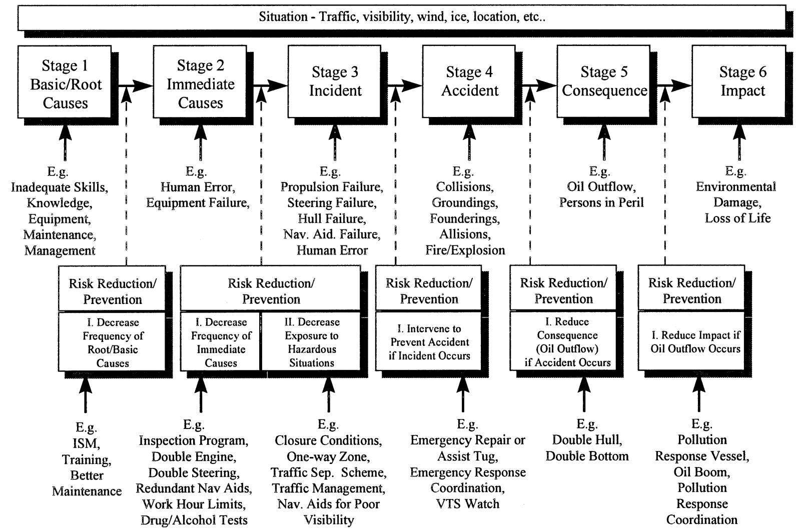

Fig. 1. Framework for maritime risk assessment and risk reduction interventions.

chain: root/basic causes, immediate causes, triggering incidents, accidents, consequences, and impacts. The combination of a triggering event and situational conditions (location, wind, weather) results in a hazard or the significant potential for an unwanted event. Pauchant and Mitroff (1992) recognized that accidents that occur in complex system are multiply determined by internal and external (situational) factors and suggested the term triggering event, rather than causal event, to describe the final stage of the accident chain.

Fig. 1 also illustrates that risk reduction interventions intervene at different points in this causal chain. Safety management programs, for example, prevent the occurrence of error. Closing the port or waterway (in PWS this means preventing transits through Hinchinbrook Entrance or Valdez Narrows) prevent exposure to a situational hazard, escort vessels prevent an incident from becoming an accident, and double hulls may prevent an oil spill if an accident occurs (but will not prevent the accident).

The results of a risk assessment provide the baseline for risk management. Risk management is the adoption of a strategy for controlling and reducing risk. The funding and adoption of specific risk reduction measures and the rejection of others as 'too expensive' or 'not cost effective' provides an operational definition of an acceptable level of risk. The sponsors of the PWS risk assessment were clearly motivated by the potential use of the risk models and risk assessment results for risk management. The primary value of the project was its evaluation of the effectiveness of proposed system interventions that could provide the basis for a risk management plan.

Estimates of the conditional probabilities that link the stages in the causal chain must be made in order to predict the risk of accidents due to human and organizational error. Unfortunately, these linkages are extremely difficult to establish and require assumptions and innovative uses of available data. The linkage is established when an uncorrected human error is allowed to effect the system. The creation of the ability to detect and correct human error before an accident occurs has the same effect as preventing the occurrence of the error in the first place. The ability of the system to 'capture' human error must also be considered when estimating the conditional probabilities in the accident chain.

2. Hidden assumptions required to model human error

The PWS risk assessment project team had to make four basic assumptions in order to model human error.

2.1. The '80-20' rule applies to PWS

Most studies of transportation related accidents have concluded that human errors cause approximately 80% of all accidents. However, the studies vary in their meaning of 'cause'—some studies count just human error as the immediate or triggering event, others include errors that occur further back in the causal chain. The significance of this assumption is that mechanical failures that are potential triggering events are more accurately captured by the reporting systems of the shipping companies and the coastguard than are human errors that have a similar potential. The frequency of occurrence of human error in the base-case system was, therefore, estimated based on the relatively complete mechanical failure data. A more equal ratio of mechanical versus human error triggering events may be reasonable in a sub system such as tanker operations in PWS which contains a high level of both internal vigilance (second officer and pilot on the bridge) and external vigilance (escort vessels, Vessel Traffic System (VTS). Changing this proportion of triggering events in the PWS risk models significantly changes the composition of predicted accidents. The proportion of drift groundings (caused by mechanical failures) to collisions and powered groundings (caused primarily by human errors) increases as the assumed ratio of human error to mechanical failure triggering incidents is decreased.

2.2. Historical accident and incident data accurately describes human error causes

Additional problems are encountered when historical accident data is used to decompose the general category of human error. This decomposition, however, must be made to examine the effects of potential risk reduction measures. Investigators can differentiate between different types of mechanical failure and can identify a mechanical failure chain. It is far more difficult, however, to differentiate between types of human error (was an error a bad decision or poor judgment?) and to untangle a human error causal chain (was the bad decision due to fatigue, to lack of knowledge, or stress?). The biases and practices of the investigator effect the quality and usability of the data recorded in accident databases. The PWS risk assessment team encountered great difficulty when trying to develop a useful decomposition of human error based on available historical data.

2.3. *Human error data from other domains is applicable to maritime risk analysis*

Human error is universal in the sense that all humans make errors. However, is it valid to use data from one environment (e.g. frequency of rule based mistakes in an industrial setting) to represent another (piloting a ship)? In the pws risk assessment, the project team used data from other environments to estimate frequency of incapacitation and inattention, but did not use non-maritime data to estimate performance related errors.

2.4. *The error capture effect of additional bridge personnel can be estimated*

Redundancy in people is not the same as mechanical redundancy where true redundancy can be achieved with independent parallel systems. If the failure rate of a single system in p_o, then the failure rate of a system consisting of two identical subsystems in parallel (e.g. a second radar) is $(p_o)^2$; if there are *n* identical systems the failure rate is $(p_o)^n$. If humans behaved like mechanical systems, human error could be virtually eliminated by adding extra persons to the bridge team! In the PWS risk assessment, the effect of additional officers on the bridge was estimated based on limited data and the personal judgment of the project team members.

3. Human error frameworks

The PWS risk assessment attempted to bridge the gap between the world of human factors research and maritime data. Reason (1990) in his Generic Error

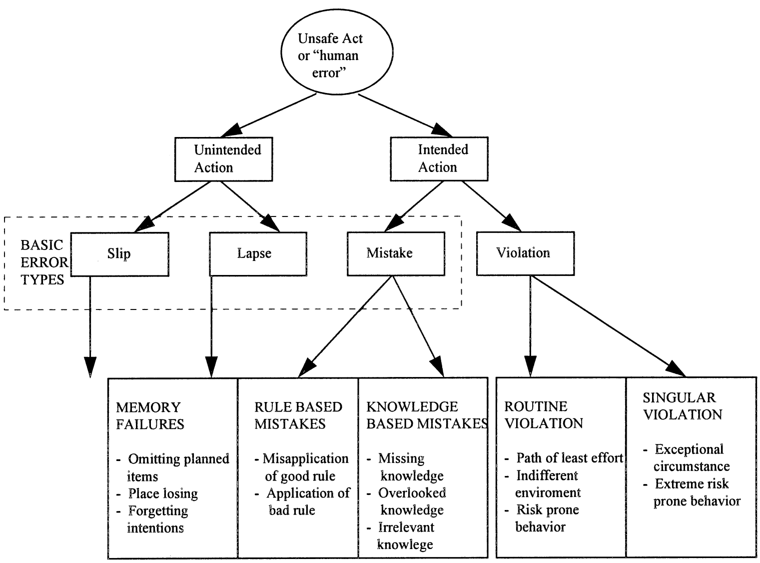

Fig. 2. Behavioral decomposition of unsafe acts (human error). Source: Reason, Silverman, Berenji.

Modeling System (GEMS) differentiates between skill based, rule based, and knowledge based errors. Fig. 2 shows a structuring of unsafe acts based on Reason (1990) as presented by Berenji (1997). This framework is, essentially, an attempt to functionally decompose generic human error into logical, mutually exclusive categories. Swain (1978) and others have proposed similar taxonomies. Swain and Gutlman (1983) also defined the term performance shaping factors to describe those organizational conditions that contribute to human error. These performance shaping factors include: inadequate work space and work layout, poor environmental conditions, inadequate human engineering design, inadequate training and job aids, and poor supervision. Boniface and Bea (1996) linked the concept of performance shaping factors to Reason's human error framework and developed a tool for analyzing maritime accidents. The impact of organizational culture on the incidence of human error has been studied by Perrow (1984, normal accidents), Roberts (1990) and Sagan (1994), high reliability organizations), and Pauchant and Mitroff (1992, crisis prone organizations).

Unfortunately, classifying an error using these theoretical frameworks requires an understanding of intent and prior personal history that is only available in the most comprehensive investigations. A theoretical framework is of little use in a risk assessment unless there is relevant data that supports it. Data is not recorded in accident databases in a form compatible with these theoretical constructs, making it very difficult to utilize the results of human factors research in risk assessment.

Two modeling techniques were used in the PWS risk assessment to capture both the impacts of human error and the potential effects of human and organizational error related risk reduction interventions. DNV analysts used a fault tree approach, used successfully in risk assessments relating to North Sea vessel traffic and offshore platforms. DNV found that accident statistics supported the estimation of acts of omission due to absence, injury, or sleep and those due to impairment that typically fall outside of the Reason–Berenji taxonomy. However, accident data did not, support the estimate (or subdivision) of substandard human performance that was used to describe all errors leading to unintended or intended incorrect performance of tasks. The GWU team developed a system simulation/expert judgment based modeling technique based on prior risk assessment projects in the State of Washington and in the Port of New Orleans. The following section describes how human error was treated by the system simulation approach.

4. The system simulation

The dynamic system simulation methodology developed by GWU is based on two assumptions:

1. risk is a dynamic property of the maritime system, and
2. the judgment of the experts that have a deep understanding of the system provide a more accurate basis for the calculation of risk than does incomplete and misleading data.

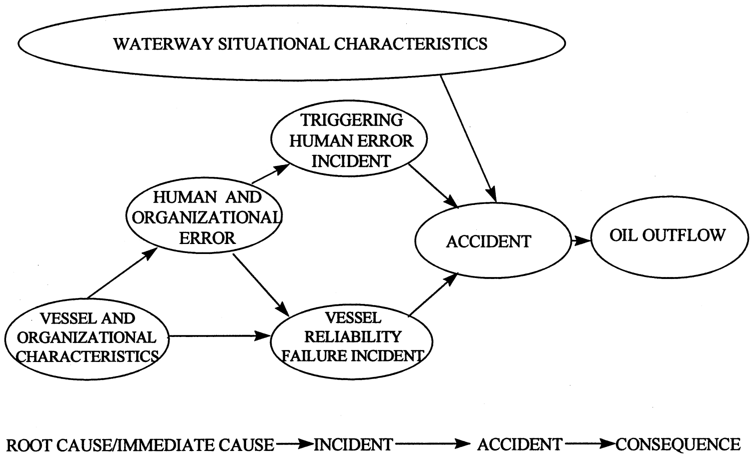

Fig. 3. Illustration of conditional relationships used in PWS system simulation.

In this view, illustrated in Fig. 3, the attributes of a vessel, and the characteristics of the vessels owner and operator are predictor's of the likelihood that the vessel will experience a mechanical failure or human error. The situational attributes of the waterway (waterway configuration, location, traffic density, weather, current, etc) influence the probability that incident will become an accident (or a 'near miss'). In the language of probability, the system simulation is based on conditional probabilities: the probability that an incident will occur is conditioned upon the vessel; the probability that an accident will occur is conditioned upon both the situation (the state of the system) and the occurrence of a triggering incident. Note that Fig. 3 models only a portion of the causal chain illustrated in Fig. 1: root causes and impacts are not part of the simulation model. The oil outflows resulting from accidents were predicted based on an oil outflow model developed by DNV.

The dynamic risk assessment process, shown in Fig. 3, required four distinct steps:

1. The relative probability of a vessel reliability failure or human/organizational error would occur on each vessel in the Alaskan fleet was calculated based on paired comparison judgments elicited from expert questionnaires.
2. The relative probability that an error or a failure occurring on a tanker would result in an accident was calculated for a set of different situational conditions based on paired comparison judgments elicited from expert questionnaires. Four categories of vessel reliability failures were defined based on the most common technological (non-human) causes of maritime accidents: propulsion failures, steering failures, electrical power failures, and structural failures.
3. The frequency of occurrence of each situational condition was determined based on actual weather, ice, visibility, and traffic data.
4. The frequency of occurrence of each accident type was calculated and calibrated against actual incident and accident data.

The attributes used to describe tankers and situations in the PWS system simulation and the critical values and conditions for each attribute, were developed from discussions with the PWS Risk Assessment Steering Committee, and interviews with representatives of the diverse groups of experts and stakeholders in PWS. The project team members made multiple ship rides in PWS, underway observations and interviews were conducted on each tanker in the PWS calling fleet, each type of escort vessel, and several other vessel types (state ferries, tour boats, and fishing boats). The experts, not the analysts, established the domain of the expert elicitation tools.

Table 1 describes five types of basic human and organizational errors that were defined as the primary causes of human error based on the USCG (1995) Prevention Through People Report. In the GWU model, these five types of basic human and organizational errors, termed vessel operational error 1 (VOE1) in the GWU model, are predicted by the values of attributes describing vessels and organizations. The conditional probabilities used to make these predictions are determined by expert judgment.

Table 1
Vessel organization and human errors (VOE1)

Human/organizational error classification	Description
1 Diminished ability	Physical, mental, motivational or emotional conditions that degrade performance
2 Hazardous shipboard environment	Poor ergonomic design, poor maintenance, or poor vessel housekeeping
3 Lack of knowledge, skills, or experience	Lack of general professional knowledge, ship slpecific knowledge, knowledge of role responsibility, or language skills
4 Poor management practices	Poor supervision, faulty management of resources, inadequate policies and procedures
5 Faulty perceptions or understanding	Inability to correctly perceive or understand external environment

VOE1, vessel operation error 1.

Table 2
Vessel operational error classification

Vessel operational error classification	Description
1 Poor decision making	Navigational or ship handling error due to failure to obtain, use or understand critical information
2 Poor judgment	Ignoring potential risks, excess speed, passing to close, etc.
3 Lack of knowledge	Inaccurate knowledge of position and situation, inability to use navigational equipment and aids
4 Poor communication	Confusing or misunderstood communicationwithin bridge team, or between vessel and VTS

VTS, ??.

The causal classes in Table 1 do not make sense to a mariner as triggering incident types. The five VOE1 error types were re-classified in operational terms, based on the PTP results, to the four primary types of human or organizational error triggering events (termed vessel operational error 2, VOE2, in the GW model): poor decision making, poor judgment, lack of knowledge, and poor communications. Note that VOE1 and VOE2 are different mappings of same set of events; if a VOE1 occurs, a VOE2 must also occur. Table 2 provides a re-framing of the error types into the 4 classes that describe the type of operational error that could be used as the basis for expert elicitation.

Both the classes in Table 1 or those in Table 2 could in principle, categorize any human error. Both schemes were used in the expert judgment questionnaires. Table 1 was used to elicit the likelihood an error based on ship and organizational attributes. Table 2 was used in questionnaires that elicited the likelihood of an accident based on the occurrence of an error under specified situational conditions. The expert questionnaires were developed in a way that experts could visualize and answer and that the responses could be quantified for subsequent use in the simulation. Experts can explain how risky different situations are when the situations are well defined at an elemental level. They cannot estimate the frequency of occurrence of rare events such as collisions and groundings. The elicitation methodology assumed that experts in the system deal with situational risk every day, and possesses a great deal of knowledge that, when quantified, can be used to estimate the baseline risk of the system and the effectiveness of risk reduction measures.

The relationship between the two categories is shown in Table 3. If data could be found to support the calculation of the distribution of errors in Table 3, the power of the simulation model to evaluate intervention measures would be greatly enhanced. Unfortunately, this distribution could not be done based on available data, and errors were assumed to be evenly distributed (e.g. one third of all errors due to diminished ability were assumed to be poor decisions, one third were cases of poor judgment, and one third were incidents of poor communications).

The conditional probabilities developed from expert judgment were incorporated into the system simulation. The simulation, shown in Fig. 4, modeled the occurrence of situations based on actual data and calculated the probability of occurrence of

Table 3
Relationship between human error causal classes and human error event type classes

Error casual class	Poor decision making	Poor judgment	Lack of situational knowledge	Poor communications
Diminished ability	XXX	XXX		XXX
Hazardous shipboard environment	XXX			XXX
Lack of knowledge, skills, for experience	XXX		XXX	XXX
Poor management practices	XXX	XXX	XXX	XXX
Faulty perceptions or understanding	XXX		XXX	XXX

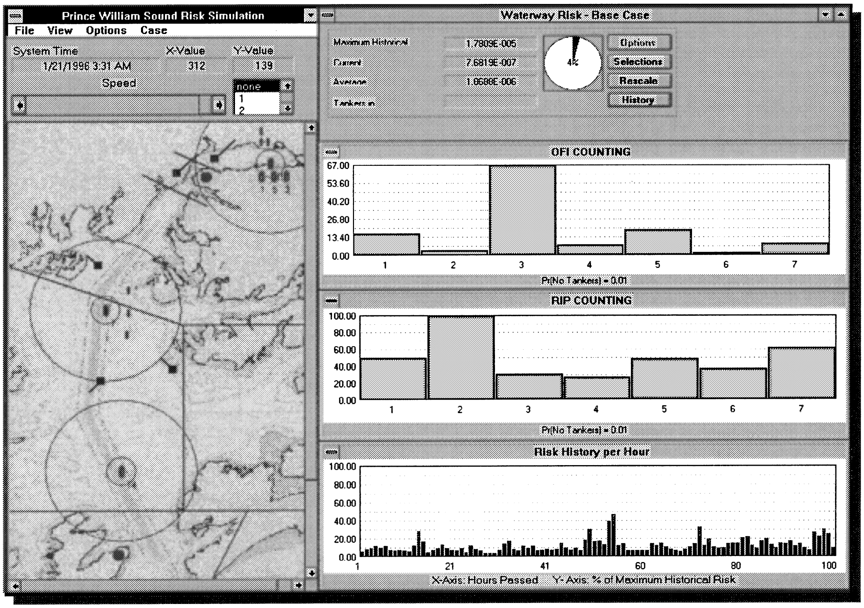

Fig. 4. Simulation analysis tool.

incidents and accidents based on expert judgment. The presence of/absence of internal vigilance is defined by situational parameters (location, presence or absence of escorts). The simulation of incidents and situational conditions made the evaluation of risk reduction measures possible.

5. Decription of project results

The project developed a range of products that provided a basis for recommendations for the effective measurement, monitoring, and management of risk in PWS. These products were delivered in four sets: (1) a detailed description of the current system and of current system hazards, (2) an evaluation of the current or baseline system risk, (3) a description of risk reduction measures, and (4) an evaluation of risk reduction interventions. (Note: The final specific quantitative results of the risk assessment are contained in the *Prince William Sound Risk Assessment Study*, December 1966.)

As stated above, the primary motivation for the PWS risk assessment was to develop a risk management plan that would create a process of continued risk reduction. The steering committee and the study team developed a list of 117 potential risk reduction measures. In order to test the risk reduction measures, they had to be converted to a form consistent with modeling parameters. The intended effects of the risk reduction measures on the system had to be identified before the appropriate modeling changes could be determined. Fig. 1 provides a six-stage framework based on the concept of the causal chain, and was used as a basis for this re-classification of risk measures. Risk interventions can effect the system by influencing stages in the causal chain in one or more of the following six ways:

1. Decrease frequency of root or basic cause events.
2. Decrease frequency of immediate cause (triggering) events.
3. Decrease exposure to hazardous situations.
4. Intervene to prevent an accident if an incident (error or failure in hazardous situation) occurs.
5. Reduce consequences (oil outflows in the PWS case) if an accident occurs.
6. Reduce the impact of consequences (ameliorate impact of oil spills in PWSRA case).

Category 6, reducing the impact of an oil-spill once it occurs, was beyond the scope the PWS risk assessment.

Human error prevention interventions affect Stages 1 and 2. Stage 3 interventions preclude the occurrence of human error in hazardous situations by preventing the exposure to these situations (e.g. port closure conditions).

Table 4 compares the relative risk reductions that could be obtained through the implementation of interventions at three stages in the causal chain as predicted by the system simulation. Note that these percentage risk reductions refer to the base case PWS risk, and address only the risk of accidents involving tankers. Case 1 is the minimization of root cause errors and substandard conditions through safety management programs and increased vessel reliability. Case 2 is error capture through increased internal and external vigilance. Case 3 is hazard exposure reduction through increased traffic management and more stringent closure conditions. Notice that Case 3 actually increases the risk of accident for inbound vessels and increases the potential oil outflows in the system. This counter intuitive effect is due to the increased traffic congestion caused by interventions in the traffic flow through the imposition of closure conditions during high winds at Hinchinbrook Entrance and Valdez Narrows.

6. Limitations on the analysis of human error reducing interventions

Accident data that is collected in the maritime domain does not provide the level of detail necessary for a causal risk analysis. The system simulation used a complex framework to describe human error, but ability to exploit the potential explanatory power of this model was limited. Although the framework was based on the USCG PTP breakdown of human error, data could not be found to support two critical connections:

1. A reasonable estimate of the distribution of human errors among the four specific classes of triggering events (VOE2) could not be determined, and therefore

Table 4
Comparison of risk reduction impacts of system interventions

Effect of measure	System simulation Case 1	System simulation Case 2	System simulation Case 3
Expected accident frequency outbound	−17%	−15%	−28%
Expected accident frequency inbound	−21%	−9%	+6%
Expected outflow total	−22%	−13%	+13%

a uniform distribution was assumed. This assumption negated the effort expended estimating the relationships between specific classes of triggering errors and specific accident types through the use of exert judgement questionnaires

2. Data could not be found to link the incidence of basic human and organizational errors (VOE1) to the triggering events (VOE2). As stated above, a uniform distribution was assumed. This assumption masks differences in the effect of these basic errors that may be very significant. Errors due to poor management practice, for example, may result in triggering errors that are predominantly errors in judgment.

The inability to decompose human error in the causal models limited the ability to examine risk reduction measures. The simulation could, for example, assume a reduction in the frequency of bad decision type human errors based on the availability of improved navigational information. It exceeded the capability of the model, however, to differentiate between the value of better navigational charts and the value of real time tide, current and ice data. Similarly, the effects of a reduction in human error due to improved professional and general knowledge could be tested, but the difference between improved bridge team training and the inclusion of the state pilot in team training could not be measured.

7. Conclusions

The movement of tankers through PWS is a complex and dynamic process. The PWS risk assessment provides a comprehensive and unique set of models that predict the effects of human error and assess the potential effects of proposed risk reduction measures. The limitations imposed by the inability to decompose human error types and causes, however, limited the ability to measure the impact of specific risk measures intended to reduce or to capture human error. It is unlikely that accident databases will ever provide the type of data required to establish these linkages. Several companies, however, have incident reporting systems that are starting to capture more complete descriptions of human error in 'near miss' situations. The availability of this proprietary data for risk analysis will be critical to further advances.

The expansion of maritime simulators provides another opportunity for the capture of descriptive human error data. Berenji (1997) has demonstrated that aviation simulator training sessions can be used to create human error databases.

The use of simulator's has three obvious advantages: (1) the occurrence of human error in simulated hazardous situations can be observed, (2) trained observers can provide uniformity in data collection, and (3) participants can be questioned to confirm the types of errors made.

The support of the organizations represented on the PWS risk assessment steering committee ensured access to maritime experts and extensive organizational data. The results of the PWS risk assessment contribute to our understanding of maritime human error. The study also helps to define the progress that must be made in collecting maritime human error data and in relating the results of human factors research to the maritime domain.

Acknowledgements

This paper is based on research conducted by The George Washington University and Rensselaer Polytechnic Insitute members of the Prince William Sound Risk Assessment Team. References are made to work performed by Det Norske Veritas, the other PWS project team members (Erling Saebo, Emil Dahle, Tim Fowler, John Acomb, and Marianne Hauso). The opinions expressed in the paper are those of the authors and do not reflect those of GWU, DNV, RPI, or the PWRSA Steering Committee.

References

Harrald, J.R., Saebo, E., Grabowski, M., 1996. Prince William Sound Risk Assessment Study. Det Norske Veritas, The George Washington University, and Rensselaer Polytechnic Institute, USA.

Berenji, M., 1997. A Methodology for Capturing and Analyzing Human Error in Aviation. Unpublished Doctoral Dissertation. The George Washington University, Washington, D.C.

Boniface, D.E., Bea, R.G., 1996 Assessing the Risks of and Countermeasures for Human and Organizational Error. Proceeding of the Society for Naval Architects and Marine Engineers 1996 Annual Conference.

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

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

Reason, J., 1990. Human Error. Cambridge University Press, Cambridge, UK.

Roberts, K.H., 1990. Some Characteristics of One Type of High Reliability Organization. Organization Science 1 (2), 160–176.

Sagan, S.D., 1994. Toward a Political Theory of Organizational Reliability. Journal of Contingencies and Crisis Management 2 (4), 228–240.

Swain, A. 1978. Accident Sequence Evaluation Procedure Program NUREG/CR–4772, National Technical Information Service (NTIS), US. Dept of Commerce, Washinton D.C.

Swain, A. and Gutlman, H.E., 1983. Handbook of human reliability analysis with emphasis on nuclear power plant applicatons. NUREG/CRI278F National Technical Interaction Service. U.S. Dept of Commerce: Washington D.C.

U.S. Coast Guard, 1995. Prevention Through People. Washington, D.C.