MANAGING AND MEASURING RISK IN TECHNOLOGICAL SYSTEMS

Romney B. Duffey
●
Atomic Energy of Canada Limited, Chalk River, ON, Canada
duffeyr@aecl.ca

John W. Saull
●
International Federation of Airworthiness, East Grinstead, UK
John@ifairworthy.fsnet.co.uk

Abstract
Safety Management is intended to create order out of disorder, to reduce the "information entropy", for the purpose of improved safety. Our purpose here and now is to try to introduce some predictability and insight into the risk or occurrence of apparently random events, where a general risk prediction we adopt a fundamental must be testable against the world's existing data. The risk management issues are clear, given the classic features of major human involvement and contribution to accidents, errors and outcomes occurring with modern technological systems. Prior incidents and prior knowledge and experience must be fully incorporated or learned from. If we do not know where we are on the learning curve, we also do not know the probability of such an event, and we have no objective measure of the "safety culture". Emphasis on defining and finding so-called "lack of safety culture" has resulted in an extensive and detailed study of the safety management and process safety of many global corporations. We utilize the concepts adopted in thermodynamics and Information Theory to establish the information entropy as a finite, physically based and useful measure of risk in technological systems. The results that we demonstrate show that the risk is dynamic, and can be utilized for management and predictive risk analysis purposes.

Keywords
Risk management; Information theory; Measurement; Safety culture; Prediction; Technological systems; Uncertainty; Probability

THE RISK PREDICTION PURPOSE

One simple worldview is that at least 90% of accidents, disasters and undesirable events (outcomes) are really due to management causes and issues, which we regard here as simply categorized as due to insufficient learning. Workers, organizations, corporations, investors and managers are all placed at risk from such events. To solve that problem, the attributes of a desired organizational "safety culture" have been defined and investigated in a number of ways, primarily based on structured surveys, interviews and questionnaires. The idea is to provide a qualitative measure or idea of how staff and management really feel and act about safety, which we regard here as some implied elimination of the error states. There are no equations and no theory: it is social science and psychometrics applied to safety.

Modern technological systems fail, sometimes with catastrophic consequences, sometimes just everyday injuries and deaths. The risk is given by the *probability* of failure, error or of any adverse outcome, and hence the *measure of risk is reflected in and by the uncertainty.* We have already examined the worldwide trends for outcomes (measured as accidents, errors and events) using data available for large complex technological systems with human involvement. We found and showed how all the data agreed with the learning theory when the accumulated experience is accounted for in Duffey and Saull (2002) [1]. Here, learning includes both positive and negative feedback, directly or indirectly, as a result of prior outcomes or experience gained, in both the organizational and individual contexts as in Ohlsson (1996) [2].

We introduce a measure of uncertainty to provide predictability and the needed insight into the risk or occurrence of these apparently random events. In seeking such a general risk measure, we adopt a fundamental theoretical approach that is and must be testable against the world's existing data.

THE MANAGEMENT CHALLENGE

Many solutions and major recommendations have been made for improving the safety management of process plants and their staff, e.g. BP Baker Panel Report [3]. Typically, the recommendations focus on how to minimize and manage the personal and business risk, paraphrasing and generalizing as follows:

- Corporate management must provide effective leadership on and establish appropriate goals for process safety;
- Establish an integrated and comprehensive process "safety management system" (SMS);
- Develop and implement a system to ensure that all management and managers, supervisors, workers, and contractors, possess an appropriate level of process safety knowledge and expertise;
- Develop a positive, trusting, and open process safety culture;
- Clearly define expectations and strengthen accountability for safety performance at all levels;
- Provide more effective and better coordinated process safety support for line management;
- Develop and implement, maintain an integrated set of leading and lagging performance indicators for more effectively monitoring safety performance;
- Establish and implement an effective system to audit process safety performance;
- Senior corporate officials should monitor the ongoing process safety performance;
- Use the lessons learned from past outcomes and events to become a recognized industry leader in process safety management.

Management generally wants to do what is right, and Regulators particularly seem to like this type of approach, as it attacks the management failings in a hopefully non-threatening and constructive way. Safety culture surveys are aimed at the attitudes, beliefs, practices and norms that hopefully characterize a pro-active approach to improving safety. But the adage "you can only manage what you can measure" means there must still be an *objective* measure of risk. What we present here is to enable such general recommendations to become a specific safety reality.

What we propose and develop are the validated means, tools and methods for management to use to: manage risks; prioritise work and recommendations; objectively measure and report the state of learning and the "culture"; and provide the company a rational approach to try to actively *predict* progress and outcomes. In that sense, what we propose is to move away from reliance on qualitative surveys to special emphasis on the quantitative measurement of learning using the knowledge gained from experience.

WHAT WE MUST PREDICT

We manage the risk, but only if we include the human element. We have shown how all outcomes develop in phases from a string or confluence of factors too complex to predict but always avoidable. We now know that a universal learning curve (ULC) exists and we can utilize that to predict outcome rates and track our progress as we improve, based on the known probability of an outcome. We have shown that [4, 5] the risk probability is given by the classic result:

$$p(\varepsilon) \equiv F(\varepsilon) = 1 - e^{-\int \lambda d\varepsilon} \quad (1)$$

where, from the Learning Hypothesis [1] at a given experience, ε, the failure rate, $\lambda(\varepsilon)$ naturally includes the human element as given by:

$$\lambda(\varepsilon) = \lambda_m + (\lambda_0 - \lambda_m) \exp - k(\varepsilon - \varepsilon_0) \quad (2)$$

We suggest, at least for the present, that it is practically *impossible* to try to describe all the nuances, permutations and possibilities behind human decision-making. Instead, we treat the homo-technological system (HTS) as an integral system. We base our analysis on the Learning Hypothesis, invoking the inseparability of the human and the technological system. Using the data, we invoke and use experience as the correct measure of integrated learning and decision-making opportunity; and we demonstrate that the HTS reliability and outcome probabilities are dynamic, simply because of learning.

The basic and sole assumption that we make every time and everywhere is the "learning hypothesis" as a physical model for human behaviour when coupled to any system. Simply and directly, we postulate that humans learn from their mistakes (outcomes) as experience is gained. So, the rate of reduction of outcome rate (observed in the technology or activity as accidents, errors and events) is proportional to the rate of the outcomes that are occurring.

That learning occurs is implicitly obvious, and the reduction in risk must affect the outcome rate directly. To set the scene, let us make it clear that the probability of error is quite universal, and can affect anyone and everyone in a homo-technological system (HTS). There are clear examples of highly skilled well-trained operators, fully equipped with warning and automated systems still making fundamental errors as an inseparable part of the technological system.

THE RISK PROBABILITY AND THE RATE OF ERRORS

Given the outcome rate, now we need to determine the outcome (error) probability, or the chance of failure. The hazard function is equivalent to the *failure or outcome rate* at any experience, $\lambda(\varepsilon)$, being the relative rate of change in the reliability, R, with experience, $1/R(\varepsilon)\ (dR(\varepsilon)/d\varepsilon)$. The *CDF or outcome fraction*, $F(\varepsilon)$, is just the observed frequency of prior outcomes, the ratio n/N, where we have recorded n, out of a total possible of N outcomes. The *frequency of prior outcomes* is identical to the observed *cumulative prior probability*, $p(\varepsilon)$, and hence is the CDF, so $F(\varepsilon) = p(\varepsilon) = (n/N) = 1 - R(\varepsilon)$, where $R(\varepsilon)$ is the *reliability,* 1-n/N, a probability measure of how many outcomes or failures did *not* occur out of the total.

The *future (or Posterior) probability*, p(P) is proportional to the Prior probability, $p(\varepsilon)$ times the Likelihood, p(L), of future outcomes. The chance of an outcome in any small observation interval, is the PDF $f(\varepsilon)$, which is just the rate of change of the failure or outcome fraction with experience, $dp(\varepsilon)/d\varepsilon$. The *Likelihood,* p(L) is the ratio, $f(\varepsilon)/F(\varepsilon)$, being the probability that an outcome will occur in some interval of experience, the PDF, to the total probability of occurrence, the CDF; and we can write the PDF as related to the failure rate integrated between limits from the beginning with no experience up to any experience, ε, as in equation (1).

We can also determine the maximum and minimum risk likelihoods, which are useful to know, by differentiating the resulting probability expression. The result shows how the risk rate systematically varies with experience and that the most likely trend is indeed given by the learning curve. In other words, we learn as we gain experience, and then reach a region of essentially no decrease, in rate or in probability, and hence in likelihood. It is easy to obtain the first decrease in rates or probabilities but harder to proceed any lower. This is exactly what is observed in transport, manufacturing, medical, industrial and other accident, death and injury data [1].

From the analysis of many millions of data points that include human error in the outcomes, we have been able to derive the key quantities that dominate current technological systems. These now include commercial air, road, ship and rail transport accidents; near-misses and events; chemical, nuclear and industrial injuries; mining injuries and manufacturing defects; general aviation events; medical misadministration and misdiagnoses; pressure vessel and piping component failures; and office paperwork and quality management systems.

From all these data, and many more, we have estimated the minimum failure rate or error interval, initial rate, λ_0, of $1/\varepsilon$, or a typical initial error interval, initial rate, at small experience of about one per 20,000 to 30,000 hours ($\lambda_0 \sim 5.10^{-5}$ per hour of experience); and a minimum attainable rate, λ_m, at large experience, ε, of about one per 100,000 to 200,000 hours ($\lambda_m \sim 5.10^{-6}$ per hour of experience);
The learning rate constant for the ULC, $k \sim 3$, is derived from the fit of a mass of available data worldwide for accidents, injuries, events, near misses and misadministrations. The following numerical dynamic form for the risk rate is our "best" available estimate from equation (2), adopting $\lambda_0 = (n/\varepsilon)$, with $n = 1$ for the initial outcome [1,2],

$$\lambda = 5.10^{-6} + (1/\varepsilon - 5.10^{-6})\, e^{-3\varepsilon} \qquad (3)$$

The risk rate, λ, can be evaluated numerically, as well as the probability, p(ε), and the differential PDF, f(ε). The result of these calculations is shown in Figure 1, where $\varepsilon \equiv \tau$ units in order to represent the accumulated experience scale.

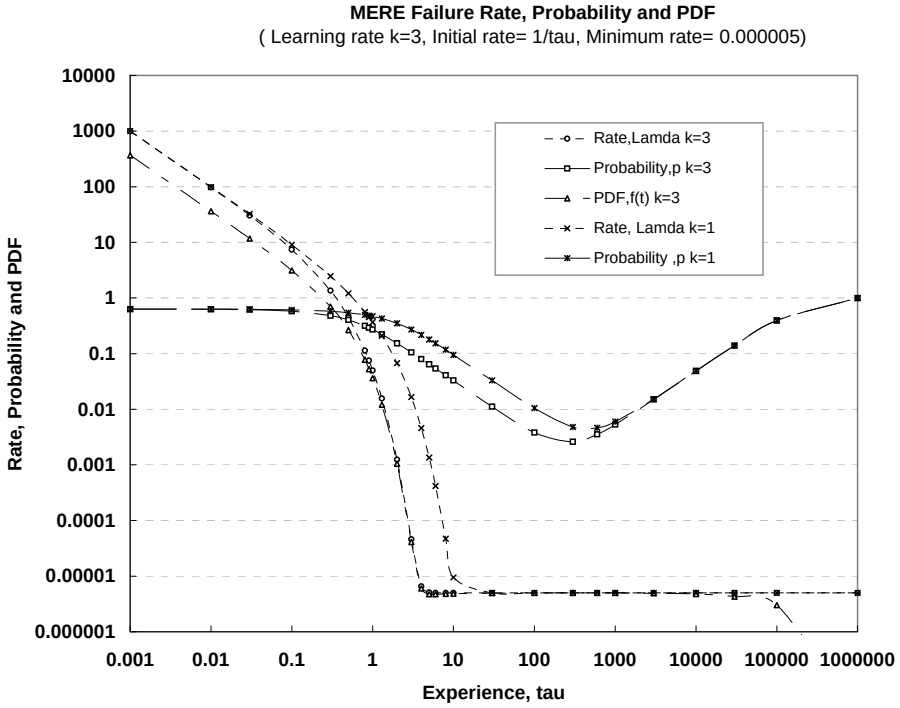

Figure 1. The best risk estimate with learning.

It is evident that for k>0 the probability is a classic "bathtub" shape, being just under near unity at the start (Figure 1), and then falling with the lowering of error rates with increasing experience. After falling to a low of about one in a hundred "chance" due to learning, it rises when the experience is $\varepsilon > 1000$ tau units, and becomes a near certainty again by a million tau units of experience as failures re-accumulate, since $\lambda_m \sim 5.10^{-6}$ per experience tau unit. The importance of learning is evident, since without learning there is no achievable minimum, which is the goal of management.

Our maximum risk is dominated by our inexperience at first, and then by lack of learning, and decreasing our risk rate largely depends on attaining experience. Our most likely risk rate is extremely sensitive to our learning rate, or k value, for a given experience.

So, as might be logically expected, the *maximum likelihood for outcomes occurs at or near the initial event rate when we are least experienced.* This is also a common sense check on our results: *we are most at risk at the very beginning.* Therefore, as could have been expected, the most likely and the least risks are reduced only by attaining increasing experience and with increased learning rates.
This approach to reduce and manage risk should come as no surprise to those in the education community, and in executive and line management positions. *A learning environment has the least risk.*

ORGANIZATIONAL LEARNING AND SAFETY CULTURE: THE "H-FACTOR"

Having examined the data and methodologies used to establish SMS, let us now return to the definition of "safety culture", which is where we started, and how it can be quantified.

Recall that the desiderata for the creation of a "safety culture", coupled to an organizational structure, places unending emphasis on safety at every level. But there is always a probability of error, a near- miss or of an event from which we must learn. We propose and prefer the use of the term and the objective of sustaining a "Learning Environment", where mistakes, outcomes and errors are used as learning vehicles to improve, and we can now define why that is true. We can manage and quantify safety effectively tracking and analyzing outcomes, using the trends to guide our needed organizational behaviors.

In the Statistical Error State Theory (SEST) [5] we found the variation in outcomes varied exponentially with depth of experience. Also the degree of order attained in a HTS was defined by "information entropy", or H-factor, the summation being a function of the probabilities of error state occupation.

The H-factor is well known in statistical mechanics where it is called the "uncertainty function", e.g., Greiner et al. (1997) [6], and in Information Theory where it is called the Shannon "information entropy", e.g., Pierce (1980) [7]. It has some key properties, namely: "as a fundamental measure of the predictability of a random event, which also enables intercomparison between different kinds of events". The H-factor is an objective measure of the *uncertainty*, and hence of the risk. This property is exactly what we would require to assess a SMS's effectiveness in reducing outcomes; and in assessing the risk in any given "safety culture".

In addition, the H-factor has the useful and necessary properties that for equally possible outcomes, $p(P) \sim 1/N$, and the (Laplace-Bernoulli) uniform prior presents the largest uncertainty, as we would expect. For a "sure thing" the H-factor is independent of the probability; and also satisfies the condition of additive probabilities for independent events. Its obvious application to safety management measurement is however *totally new* as presented here in Duffey and Saull [8], and arises quite naturally from the need for management to create order from disorder, and reduce uncertainty.

In terms of probabilities based on the frequency of microstate occupation, $n_i = p_i N_j$ and using Stirling's approximation we have the classic result for the Information Entropy:

$$H_j = - \Sigma p_i \ln p_i \qquad (4)$$

and the *maximum uncertainty value occurs for a uniform distribution of outcomes*. The corresponding probability of occupation as a function of experience:

$$p_i = p_0 \exp(\alpha - \beta\varepsilon_i) \qquad (5)$$

We note that since we have observed the outcomes, the usual normalization condition for all the N_j outcomes to exist is, summing the probabilities over all the j observation ranges, $\Sigma_j p_i = 1$. For the probability distribution of a continuous random variable, we can transform the sum to an integral. This normalization says simply that whatever outcomes happened must occur. *The risk always exists, somewhere in observational space*.

In practice, the probability of occupation according to the SEST is approximated by a fit to the available outcome data [8] given by:

$$p_i = p_0 \exp - aN^*, \qquad (6)$$

where, a, is a constant, and N*, the non-dimensional measure of the depth of experience, $\varepsilon/\varepsilon_M$. Thus, for our continuous probability function, we can evaluate the (so-called grand) partition function, and write the probability of error state occupancy as:

$$p_i = p_0 \exp(-aN^*)/\int_0^\infty p_0 \exp(-aN^*) \quad (7)$$

or,

$$p_i = (a/\varepsilon_M) \exp(-aN^*), \quad (8)$$

and hence the probability decreases as the learning rate and experience depth increases. Since the outcomes are represented by a continuous random variable learning curve, the Information Entropy, H, in any j^{th} observation interval is also given by the integral:

$$H_j = -\int p_i \ln p_i \, dp$$
$$= p_i^2 (1/4 - \tfrac{1}{2} \ln p_i) \quad (9)$$

So, substituting in the expression for the Information Entropy, H, which we term the *"H-factor"*:

$$H_j = \tfrac{1}{2} \{p_0 e^{-aN^*}\}^2 \{aN^* + \tfrac{1}{2}) \quad (10)$$

where, on a relative basis, $p_0 = 1$, and then $H \rightarrow 0.25$ as experience decreases as $N^* \rightarrow 0$. This parameter is a measure of the uncertainty, and hence of the risk.

As either the learning rate or depth of experience increases ($N^* \uparrow$ or $a\uparrow$), or the zeroth order occupancy decreases ($p_0 \downarrow$), so does the value of the H-factor decline, meaning a more uniform distribution and increased order. We illustrate the variation in the relative information entropy, H, with non-dimensional experience depth, N*, in Figure 3, taking the zeroth probability as unity ($p_0 = 1$) for a range of learning rates. The range chosen varies around the "best" value of a = 3.5, which is as derived from the USA aircraft near-miss and Australian auto death data so that:

$$P_i = p_0 e^{-3.5N^*} \quad (11)$$

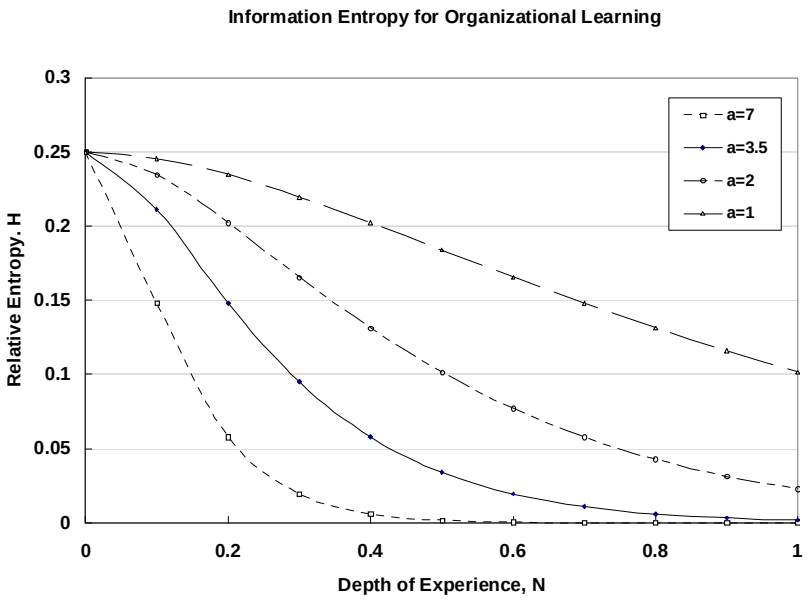

Figure 3: Organizational Learning and Experience

Clearly, the relative value of the information entropy H-factor at any experience depth is a direct measure of any cultural aspect of modern technologies called "organizational learning". This terminology is meant to describe the attributes of a HTS, and its ability to respond effectively to the demands for continuous improvement, as reflected in internal organizational and communication aspects.

The faster decline and decrease in the H-factor with increasing depth of experience and increasing learning constant is a reflection of increasing HTS organizational order. This is to be expected, and makes sense: it is

exactly what safety management intends. This relational effect is also exactly what we mean by maintaining a Learning Environment, and has been derived from the new SEST analysis.

Before this discovery, all one could conduct were semi-empirical, qualitative and highly subjective comparative surveys of "organizational attitudes and beliefs". These would review opinions and attitudes to management, training and personnel systems without having a truly objective measure.

MANAGING INFORMATION AND SAFETY CULTURE

We can now argue that this purely theoretical concept of degree of order, the H factor, is actually a true practical and quantitative measure of the elusive quality termed "safety culture" by sociologists, human factors experts and industrial psychologists. Safety culture is therefore a reflection of the degree of order, and reduced uncertainty, attained in and by any HTS; and creating order is equivalent to reducing the risk probability of any outcome.

As we stated in the very beginning of this paper, it is management's expressed aim and intent in any technological system to create order from disorder, which it can only achieve by decreasing the information entropy. Unfortunately, most safety managers who are trained in traditional industrial safety methods, and corporate officers familiar to the world of business and accounting decisions and risks, would not recognize the concept of entropy, let alone information entropy, if they saw it. However, it is so simple to communicate the concept of the learning hypothesis and the impact on organizational learning, that it should be possible to obtain the management buy-in needed to adopt this approach to assess risk and safety.

Equally important to this quantification is the realization that this H-factor uses the actual outcomes as an explicit function of organizational learning and management effectiveness. We indeed do and can "manage what we can measure". This is simply common sense.

CONCLUSIONS: IMPLICATIONS FOR RISK MANAGEMENT AND PREDICTION

The implications of using this new approach for estimating risk are profound. This new probability estimate is based on the failure rate describing the ULC, which is derived from the Learning Hypothesis; and utilizes the validation from the outcome data of the world's homo-technological systems. For the first time, we are also able to make predictions of the probability of errors and outcomes for any assumed experience interval in any homo-technological system.

In addition the results implies a finite lower bound probability of based on the best calculations and all the available data. Analysis of failure rates due to human error and the rate of learning allow a new determination of the risk due to dynamic human error in technological systems, consistent with and derived from the available world data. The basis for the analysis is the "learning hypothesis" that humans learn from experience, and consequently the accumulated experience defines the failure rate.
The extension of the concept to "safety culture" shows this risk can be interpreted as uncertainty, and that uncertainty can be quantified by the Information Entropy, or H-factor. Management wish to emphasize "safety culture", which actually corresponds to a sustained Learning environment, managing risk creates order, reduces uncertainty and ensures predictability. Based on our theory and practical data, we have shown how to quantify order, reduce uncertainty and predict risk. The risk probability is based on experience.

The results demonstrate that the risk is dynamic, and that it may be predicted using the learning hypothesis and the minimum failure rate, and can be utilized for predictive risk management purposes.

REFERENCES

[1] Duffey, R.B. and Saull, J.W. (2002). Know the Risk, First Edition, Butterworth and Heinemann, Boston, USA.

[2] Ohlsson, S. (1996). Learning from Performance Errors, Psychological Review, Vol. 103, No. 2, 241-262.
[3] Report of the BP US Refineries Independent Safety Review Panel, (2007). The Baker Panel Report, available: www.bp.com.
[4] Duffey, R.B. and Saull, J.W. (2006). The Human Bathtub: Safety and Risk Predictions Including the Dynamic Probability of Operator Errors, Proceedings 14th International Conference on Nuclear Engineering (ICONE14), Paper No. 89476, Miami, Florida, July 17-20.
[5] Duffey, R.B. and Saull, J.W. (2005). The Probability of System Failure and Human Error and the Influence of Depth of Experience, Proceedings of International Topical Meeting on Probabilistic Safety Analysis (PSA'05), San Francisco, CA, September 11-15.
[6] Greiner, W., Neise, L. and Stocker, H. (1997). Thermodynamics and Statistical Mechanics, Springer, New York, pp. 150-151.
[7] Pierce, J.R. (1980). An Introduction to Information Theory Symbols, Signals and Noise, second Revised Edition, Dover Publications, Inc., N.Y.
[8] Duffey, R.B. and Saull, J.W. (unpublished). Managing Risk.