The Quantification of Resilience:
Learning Environments and Managing Risk

Romney B. Duffey
Atomic Energy of Canada Limited, Chalk River, ON K0J 1J0, Canada
duffeyr@aecl.ca

Abstract. We define and quantify the abstract concept of Resilience Engineering (RE). The Hollnagel's definition of resilience is some intrinsic ability of an organization to maintain or regain a dynamically stable state when responding to upset and stress. We propose and demonstrate a formal quantification of "resilience" and organizational stability based precisely on, and linking RE directly to the "established risk management approaches (that) are based on hindsight and emphasise error tabulation and calculation of failure probabilities… to anticipate the changing shape of risk before damage occurs".

Our new concepts are consistent with Ilya Prigogine's theory of how a non-equilibrium and statistically fluctuating system can evolve towards an ordered state in the physical world; and with Rudolph Giuliani's leadership principle of "relentless preparation" in dealing with the unexpected in organizations. We quantify the objective measures of safety using the very uncertainty and complexity of what has already happened and been observed, and determine the probability of present and future occurrences as learning occurs. The new organizational risk stability criterion that quantifies resilience thus emerges naturally, and can be compared to actual data.

The emergence of order from disorder characterizes risk management and organizational stability, and hence system resilience, as a result of learning.

1 MANAGING THE RISK OF THE UNKNOWN

Our previous work has established the existence or universal learning trends at both the system and individual level, and has demonstrated and validated that approach using all the available and relevant data (Duffey and Saull, 2002, 2008).

We quantify here the intrinsic and learned stability arising from those many instances in life, where both managing and taking risks requires planning for unknown disasters and outcomes, including events such as terrorist attacks, explosions and fires (Giuliani and Kurson, 2002; BP, 2005, 2007; and Barthelemy, 2001. We must understand and take into account both what we know and what we do not know about the risk. We must expect the unexpected, and anticipate the unanticipated, and be able to respond. This dilemma is crystallized in decision theory and analysis, when we must determine our response, decisions and actions based on both what we do and do not expect to happen. The need to counter the actions of terrorists and extremists is one recent example of risk caused by the unknown and of extreme organizational stress. After the attacks on the World Trade Center, Mayor Rudolph Giuliani of New York made it clear that collectively they did not expect to be attacked in that way: it was an unknown and unanticipated outcome. But because of what Guiliani characterizes as "relentless preparation" the emergency services (fire, police, ambulance, security, treatment, transport etc.,) with the direction of the Mayor were able to stitch together an effective emergency response and command structure from the remnants or pieces of prior anticipated (or known) events, capabilities, planning, exercises and knowledge.

Such ability to cope and provide order out of chaos also at the heart of remaining "organizationally stable" against unexpected and large stress and unexpected occurrences. This same question and issue of system stability under stress also has direct application to the subjective (and somewhat topical) concept of "resilience engineering", where "*…resilience is the intrinsic ability of an organization (system) to maintain or regain a dynamically stable state, which allows it to continue operation after a major mishap and/or the presence of a continuous stress"*(Hollnagel et al, 2006). But "resilience" has not been actually measured or quantified anywhere: it is simply a desirable property. We develop here the numerical and objective criterion that is precisely applicable to the quantification of "resilience" and organizational stability, hence incidentally unifying that empirical concept with the general theory and practice of managing risk through learning. This criterion is also relevant to "crisis management" policies and procedures, and emergency response centres in major corporations, facilities and industries.

In this paper, derived from (Duffey and Saull, 2008), by a physical analogy, the approach links the emergence of learning in human organizations and entities with recent ideas of the emergence of order and structure from chaos in the physical sciences.

2 ORDER AND DISORDER IN PHYSICAL AND MANAGEMENT SYSTEMS

As eloquently suggested by Ilya Prigogine and his co-workers (Kondepudi and Prigogine, 1998), a non-equilibrium and statistically fluctuating system can evolve towards an ordered state. Paradoxically, the fluctuations at the unobserved microscopic atomic and molecular level that characterize the non-equilibrium themselves provide the *necessary* opportunity for structure and order to emerge as distributions at the observed macroscopic level. Non-equilibrium entropy characterizes the creation of structure from disorder in HTS as well as for the purely physical and chemical systems discussed in Prigogine's work. We assert the analogy between these two apparently vastly different fields, namely physico-chemical and homo-technological systems.

In our technological world we have the same randomness and disorder that must exist for order to appear and for learning patterns to occur. Managers, executives, employees, procedures, training and individual skill acquisition intend to achieve the creation of order in any HTS or corporate organization from the natural disorder. The system learns how to behave and learn macroscopically (externally and organizationally), when in fact it is a collection of a myriad of microscopic (internally and individually) fluctuating and unpredictable interactions (e.g., in discussions, meetings, rules, procedures, communications, training, one-on-ones, coffee breaks, lunch groups, hallway gatherings, rumour mills….). In practice in the “real” world, such multitudinous random and informal learning opportunities exist in addition to the purely formal and official ones.

We have already shown that the objective measure of uncertainty and hence of risk, is also a measure of the degree of order or learning, being the information (or learning) entropy, H (Duffey and Saull, 2008) and also determines response time during learning. The Universal Law of Practice emerges. This H-factor is well known in statistical physics, thermodynamics and information theory as a measure of the “missing information” and is called the “uncertainty function” (see e.g. W. Greiner et al, 1997). 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”. This property is exactly what we would require to assess effectiveness in reducing and managing outcomes.

In addition, the H-factor has the useful and necessary property of a uniform prior being the largest uncertainty, as we would expect, and also satisfies the condition of additive probabilities for independent events. Its obvious application to safety management *measurement i*s however totally new as presented here, and arises quite naturally from the need for management to create order from disorder. In terms of probabilities based on the frequency of error state occupation, $n_i = p_i N_j$, we have the classic result for the Information Entropy:

$H_j = - \Sigma p_i \ln p_i$

and the *maximum value occurs for a uniform distribution of outcomes*. Interestingly, this is also equivalent to the Bayes-Laplace result, when the posterior probability, $p(P) \sim 1/N$, for a uniform risk. The occupancy number as a function of depth of experience, ε_i gives the corresponding probability distribution:

$p_i = p_0 \exp(\alpha - \beta\varepsilon_i)$, where α and β are constants.

Summing the probabilities over all the j observation ranges, $\Sigma_j p_i = 1$, which normalization says simply that whatever happens outcomes must occur. *The risk always exists, somewhere in observational space*.

In practice, the probability of distribution on occupation is approximated by a fit to all the available outcome data given by (see Duffey and Saull, 2008):

$p_i = p_0 \exp - aN^*$,

where, a, is a constant, and N*, the non-dimensional measure of the depth of experience, $\varepsilon/\varepsilon_M$.

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 in any j^{th} observation interval is also given by the integral (Duffey & Saull, 2008):

$$H_j = - \int p_i \ln p_i \, dp$$
$$= p_i^2 (1/4 - ½ \ln p_i)$$

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

$$H_j = ½ \{p_0 e^{-aN^*}\}^2 \{aN^* + ½\}$$

where, on a relative basis, $p_0 = 1$, and then $H \rightarrow 0.25$ as experience decreases. This parameter, H_j, is an objective measure of the uncertainty, and hence of the *risk for any system*.

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. The range chosen varies around the "best" value of $a = 3.5$, which is as derived in (Duffey and Saull, 2008) from aircraft near-miss and auto death data. The relative value of the information entropy H-factor at any experience depth is a direct measure of the 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 and response, as reflected in internal organizational and communication aspects, just as it would in a crisis.

3 THE STABILITY CRITERION

The stability condition adopted for molecular systems simply represents how they respond dynamically to thermodynamic entropy changes. Thus, for stability the incremental change in the thermodynamic entropy, dS, must be negative (Kondepudi and Prigogine, 1998), so in any *time* increment:

$$dS/dt \leq 0$$

Now, we have the Information Entropy, not the thermodynamic entropy, but both are a measure of the system disorder. We define our equivalent organizational risk stability criterion (ORSC) from the fact that the incremental change in risk, as measured by changes in the information entropy, H, with changes in probability must reduce. In any *experience* increment we must have:

$$dH/dp \leq 0$$

where,

$$dH/dp = - p \ln p$$

There is a parallel requirement for convergence in iterative computational learning machines, which is termed "empirical risk minimization". A necessary condition for convergence is for the equivalent computational entropy measure to vanish in the limit of many observations or samples, meaning as $n \rightarrow \infty$ (Vlapnik, 2000). So the equivalent numerical convergence theorem or requirement is such that:

$\mathrm{Lim}_{n\to\infty}$ (H/n) = 0

We postulate that this requirement is equivalent and analogous to Prigogine's approach for physical and chemical systems, being an incremental stability condition for *any* system (i.e., where order is emerging from chaos). For incremental changes, and since p = n/N, assuming a sufficient numbers of outcomes,

dH/dp → 0, as n → ∞

From (Duffey and Saull, 2008) in any observation interval for any and all homo-technological systems, the outcome probability distribution varies with depth of experience nearly as,

$p \approx p_0\, e^{-\beta\varepsilon}$

Since by definition dH/dp = - p ln p, then,

$dH/dp = p_0\, e^{-\beta\varepsilon}(\beta\varepsilon - \ln p_0)$

where, as usual, the probability, $p_0 = n_0/N_j$, is the ratio of relative irreducible zeroth error state number to the total observed, and β, is the learning shape parameter. We assert that for any observed HTS to be stable the necessary stability condition (ORSC) is given by evaluating the condition for convergence. The above stability inequality requires that:

$\beta\varepsilon - \ln p_0 \rightarrow 0$

Physically, we can say that the ratio of error reduction by learning to the number of irreducible errors, at any moment of experience, must be such that for large outcome numbers, n, simply:

$\{\beta\varepsilon/\ln p_0\} \leq 1$

Applying this stability condition to "resilience" further, we examine the limits of rare and unexpected events, which are particularly challenging from the management perspective.

4 RARE AND UNEXPECTED EVENTS

Major events, like the 9/11 terrorist attacks in New York (Giuliani & Kurson, 2002), or the Texas City refinery explosion (BP, 2005 and U.S. CSB, 2007), are unexpected and hopefully rare in their occurrence. They threaten and disrupt normality and how well the consequences as well as the event itself are managed are the measure of past and present learning. For the first or rare event n~1, since it has hopefully not been observed before and the posterior probability p(P) may be approximated by the pdf, f(ε), where for rare events, f(ε) ≡ n/ε, where n = 1. We assume that the zeroth and irreducible probability, $p_0 \sim (1/\varepsilon)$, at whatever experience it was observed and the stability condition becomes:

$dH/dp = e^{-\beta\varepsilon}(\beta\varepsilon - \ln(1/\varepsilon))/\varepsilon$

Hence for convergence, the ORSC becomes, $\{\beta\varepsilon/\ln(1/\varepsilon)\} \leq 1$

Expanding the logarithmic term for the relevant case of small experience when, $0<\varepsilon \leq 2$, and retaining the first term only, experience being small, the stability condition is

$\beta\varepsilon\ (1-\varepsilon+\ldots) \leq 1$

or, very approximately, when we have little experience and $\varepsilon << 1$,

$\beta \leq 1/\varepsilon$

Thus physically and organizationally, the required learning rate constant should be such that very nearly, $\beta \leq 1/\varepsilon$, in order to maintain at least a constant level of risk, or resilience against risk of "organizational" collapse. *So for a "stable" resilient organization, naturally the risk variation with learning must remain at or below what it was at the very beginning.* This is not surprising: indeed it is to be expected, and requires vigilance, attention, learning, management, commitment and measurement of the prior trends.

5 ORGANIZATIONAL RISK STABILITY

This new stability ratio ($\beta\varepsilon/\ln p_0$) is called the Organizational Risk Stability Number (ORSN), also incidentally representing the quantification of "resilience". For stability and/or convergence the ORSN *must* have a value that is less than unity.

To demonstrate the principle, we use real data for a simple and well-defined data set, in this particular case for coal mining fatalities. The variation of the number of fatalities, F, with depth of experience (accumulated millions of tons mined, Mt) was used to provide the working estimate for the learning exponent, β; and the value for p_0 was calculated to be ~ 0.046 from the numbers of fatalities. Figure 1 shows the resulting ORSC below the plot of the instantaneous rate, IR. It can be seen clearly that the estimated value of $\{\beta\varepsilon/\ln p_0\} \leq 1$ always, and is stable for this purely illustrative example.

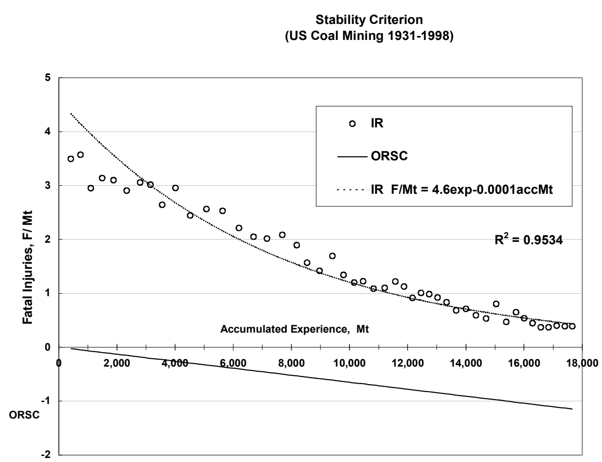

Fig. 1. Example stability criterion calculation

6 CONCLUSIONS: QUANTIFYING RISK AND RESILIENCE

Our major objective is to provide both quantification and *prediction* of risk, which also means turning qualitative concepts such as “safety culture”, “safety management system” and “resilience” into actually measurable and manageable trends and quantities.

The quantification of the degree of order in a system, and hence the effectiveness of management, learning and skill acquisition, is the information entropy, a fundamental quantity in the physical and human world. By analogy between stability in physical and mathematical learning systems, we derived a new criterion (the ORSC) that uniquely determines the (organizational) stability for any system. This then also quantifies the desirable ability of any system to respond to crisis, stress or unexpected events, including rare events.

For the case of “engineering resilience” the stability condition is derived from the necessary condition for the existence of increasing order. Self-evidently, stability requires that organizational learning be such as to maintain a constant or reducing risk or outcome probability. Hence, we have illustrated the obvious importance of learning influences on effective emergency preparedness and management response. We have formally shown the same requirement also holds for rare and/or unknown events. The concept of “resilience” is entirely synonymous with the presence and pursuit of learning, demonstrated in and by the entire system.

REFERENCES

BP (2005). Fatal Accident Investigation Report: Isomerization Unit Explosion, Final Report, Texas City, Texas, USA, December 9th (known as “The Mogford Report”) plus 26 Appendices, *www.bp.com*.

BP (2007). The Report of the BP U.S. Refineries Independent Safety Review Panel, (the “Baker Panel”) January, www.bp.com.

Barthelemy, F. (2001). Report of the General Inspectorate for the Environment, Affair No. IGE/01/034. “Accident on the 21st of September 2001 at a factory belonging to the Grande Paroisse Company in Toulouse, produced jointly with the Explosives Inspectorate and with help from the INERIS”, Inspection Générale de l'Environnement, Paris, France, 24 October.

Duffey, R.B. & Saull, J.W. (2002). *Know the Risk*. Boston, USA: First Edition, Butterworth and Heinemann.

Duffey, R.B. & Saull, J.W. (2008). Managing and Predicting Technological Risk and Human Reliability : A New Learning Curve Theory. *Proc. ImechE. Vol. 222* Part O : J. Risk and Reliability pp. 245-254.

Duffey, R.B. & Saull, J.W. (2008). *Managing Risk: the Human Element*. Chichester, UK: John Wiley & Sons Ltd.

Giuliani, R.W. & Kurson, K. (2002). *Leadership*. Chapter 3. New York: Hyperion.

Grenier, W., Neise, L. & Stocker, H. (1997). *Thermodynamics and Statistical Mechanics,* pp.150-151. New York: Springer.

Hollnagel, E, Woods, D.W., & Leveson, N. (Editors) (2006). *Resilience Engineering: Concepts and Precepts*. Chapter 1, Resilience - the Challenge of the Unstable, pp.9-16.

Aldershot, England: Ashgate Publishing.

Kondepudi, D. & Prigogine, I. (1998). *Modern Thermodynamics: From Heat Engines to Dissipative Structures*. New York: John Wiley & Sons.

U.S. Chemical Safety & Hazard Investigation Board (CSB) (2007). Investigation Report No. 2005-04-I-TX. Refinery Explosion and Fire, BP, Texas City, March 23, 2005.

Vlapnik, V.N. (2000). *The Nature of Statistical Learning Theory*, pp. 43-45. New York: 2nd edition, Springer-Verlag.