Predicting tragedies, accidents, errors and failures using a learning environment

Romney B. Duffey *

Atomic Energy of Canada Limited, Chalk River Laboratories, Chalk River, Ont., Canada K0J 1J0

Received 4 March 2005; accepted 30 July 2005
Available online 25 October 2005

Abstract

Failures in technological systems are dominated by the properties of the ‘human material’ component more than those due to the physical material. As we introduce new technology into our technological systems, we take the risk of failure that is dependent on our experience. We present the facts behind the integral contribution of human error to accidents and risk and provide the theoretical analysis of outcome rates based on the Learning Hypothesis. The importance of the accumulated experience on the learning rate, and the depth of experience on the distribution of outcomes is presented. The implications are given for predicting the outcomes, and also for the safety of the human material embedded and used in homo-technological systems (HTS).

© 2005 Elsevier Ltd. All rights reserved.

Keywords: Events; Human error; Failures; Predictions; Learning environment; Experience; Probability

1. Introduction

For technological systems, the usual focus is to define the failure rate and probabilities for the components and materials, adding safety and design margins plus back ups when needed. But, in practice, the ‘human material’ component is the major contribution to failures in all technological systems (> 60%), and for the observed tragedies, accidents, errors and events.

The human component is inextricably interwoven into the system via design, construction, quality, operation, management, training, maintenance, and regulation. We have no good testable theory or method for including human failures, and for preventing them. We have no good testable theory for predicting when events will occur and at what rate: we just observe them and fix the cause(s) afterwards case-by-case.

There is a rich history of technological failures, including the best technology available at the time. The mighty wooden warships Mary Rose and Vasa both sank due to design and operational errors, as well as the liner Titanic sinking and the Hindenburg airship burning. Major technological disasters have been the Chernobyl reactor burning, the Bhopal chemical release and the Three Mile Island reactor accident. Even the recent Indian Ocean tsunami deaths were due to inadequate warning technology and lack of engineered disaster mitigation systems. Our best materials used in technology have caused multiple recent events that were all human-designed failures, such as: the mode of the steel-framed World Trade Center collapse; the tire puncture and resultant fuel tank puncture causing the Concorde aircraft crash, and the insulation debris impacting and damaging the wing edge that lead to the Columbia shuttle loss. Everyday, more mundane events occur: auto accidents, train wrecks and derailments, vehicle rollovers, tire failures, murders, medical errors, aircraft crashes, oil spills, pipeline and chemical explosions, aircraft near-misses. They occur in all our best and latest homo-technological systems (HTS), and somewhere they all have the contribution from human mistakes in common.

The inability to predict, anticipate or prevent is because all accidents are stochastic in their occurrence. Like the spectacular crashes of the Space Shuttle and the supersonic Concorde, or everyday events like collisions driving your vehicle, they are due to an unexpected, unforeseen or unpredicted combination of human and technological system failures, working together in totally undetected ways, occurring at some random moment. Such was the same identical cause with the great Power Blackout of 2003 in the USA and Canada leaving millions literally in the dark.

Afterwards we want to know what happened and why. So massive and expensive inquiries are held into these crashes and events, like the Space Shuttle losses, and into other spectacular
events like the tragic UK Ladbroke Grove railway collision, the JCO nuclear criticality incident in Japan, and the melting core of the USA Three Mile Island reactor. The results invariably differ in detail, but always show the causes are due to an unexpected combination or sequence of human, management, operational, design and training mistakes. The major contribution and cause is invariably due to the human involvement, not just failure of the engineering material. Once we know the facts of what happened, reactively and collectively, we fix the engineering or design failures, and try to obviate the human ones.

We have invented and used HTS over the last two centuries, from the very start of the Industrial Revolution to today. We now have amassed literally many millions of data points on deaths and injuries, losses and damages, on real individual tragedies. We find that technological systems do not usually fail all by themselves after all we collectively did invent and build them. The major contributors are the errors caused by us, *Homo sapiens*, in the system design, construction, operation, maintenance, control, licensing and management. Be it a bridge collapse, an airplane near-miss, an operating room death, a milling machine accident, a train derailment, an automobile collision, a fall from a ladder, a pipeline fire, a chemical plant explosion, or a mineshaft collapse, we cannot separate out the human component of the mistakes easily. When looking for cause, or blame, or improvement, we have to consider the entire homo-technological system (HTS), of which we humans are an integral and essential part.

By hypothesizing initiating events and conducting hazard analyses, invoking safety and performance standards, and probabilistic risk and safety assessments, we implement design and protective features using safety measures and equipment. We also try to prevent future events by invoking laws, procedures, regulations and penalties for similar mistakes. By paying attention to features like the human–machine interface (the controls, the alarms, the layout, and the labeling) using human-centered design, we can also try to make the operation less risky, the diagnosis and prevention of failures more simple and intuitive, and avoid blunders in the operation. Automation can be introduced, taking the human out of the control and decision-making loop, although this itself can lead to issues of when, if ever, the human should intervene or take over.

2. A new approach

We need to understand how we must and do learn from our mistakes and improve our safety as we go, so we may anticipate and prevent using prior knowledge. Since their occurrence is random, not predictable, and we may observe an overall rate of events, deaths or injuries, we cannot predict when any one event will occur or how much reduction in outcomes a particular safety measure will provide. So we end up being reactive, in our laws, our punishment, our redesigns, our recalls, and in our training and our knowledge of failure. The simplest hypothesis we can make is the Learning Hypothesis governing human behavior in any and all HTS (for a full discussion see Duffey and Saull) [1]. For any HTS, the basic and sole assumption that we make is to adopt the ‘Learning Hypothesis’ as a physical model for human behavior when coupled to and in any technology (an HTS).

Simply and directly, we postulate that humans learn from their mistakes (outcomes) as experience is gained. We must translate our thinking from the usual paradigm of time to that needed for experience with an HTS. As humans, we observe and record events in the frame of calendar time, often on a daily, yearly or fiscal basis. But technological systems, failures and human learning depend on the experience accumulated. Only in the special case of constant, unchanging systems are the two observation frames equal. We must translate our thinking and our observation frame from ‘time passing’ to ‘experience space’ so we can construct and utilize a ‘learning environment’.

The answer seems to be that we just have to learn. In a sense this is a Darwinian extension, since those who do not learn would, without help or a forgiving society, be eliminated and not survive in the technological world. Learning as we go, we literally ‘klutz’ our way through modern society. Although we make errors all the time, as we move from being novices to acquiring expertise, we should expect to reduce our errors, or at least not make the same ones.

Assuming only that the rate at which errors or accidents reduce with increasing experience, $d\lambda/dx$, is proportional to the rate, $\lambda$, at which errors or accidents are occurring, we have the Minimum Error Rate Eq., or MERE:

$$\frac{d\lambda}{dx} \propto \lambda$$  \(\text{(1)}\)  

or

$$\frac{d\lambda}{dx} = -k(\lambda - \lambda_m)$$  \(\text{(2)}\)

where $k$ is the learning rate constant. The solution to this Eq. is, for an initial rate $\lambda_0$ at $x = 0$, and a minimum rate $\lambda_m$ at $x \rightarrow \infty$, given by the exponential universal learning curve (ULC):

$$E^* = \exp -KN^*$$  \(\text{(3)}\)

where $E^* = (1 - \lambda/\lambda_0)(1 - \lambda_m/\lambda_0)$ and $N^*$ is the non-dimensional accumulated experience, $\epsilon/\epsilon_M$. Thus, hopefully, we should descend a ‘learning curve’ for any and all technological systems, like the path shown in Fig. 1, where our rate of making mistakes decreases as we learn from experience.

The failure rate as a function of any accumulated experience, $\lambda(\epsilon)$, is then obtained by straightforward integration of Eq. (1) from some initial rate and experience, $\epsilon_0$, out to the present or future with any larger experience. Thus, the error or failure rate, $\lambda$, decreases exponentially with increasing experience, $\epsilon$, as given by the MERE.

We need to relate this failure rate to the recorded accident, event and error data for outcomes, say as numbers killed per unit experience. We have the proportionality that the failure rate $\lambda(\epsilon) \propto A$, the observed outcome rate, measured as counts of accidents, deaths, outcomes or errors. Now the failure rate is proportional to $(dn/dx)$, the observed instantaneous outcome rate with changing experience, so:

$$\lambda(\epsilon) = \{(1/(N-n))(dn/d\epsilon)\}$$  \(\text{(4)}\)
agencies, administrations, inquiries, safety authorities, industries to use the multitudinous outcome data reported by the many measurable, recorded or observed outcomes, but simply allows that result does not pretend or assume that some errors have no errors on the way.

Fig. 1. The Learning Hypothesis— as we learn we descend the curve decreasing

falling away towards

Plotting the error rate data, $A = IR = \lambda$, versus the accumulated experience, $\varepsilon$, we should expect to also observe an exponential decline if we are learning, starting out with a rate, $\lambda_0$, and falling away towards $\lambda_m$, the minimum contribution from human error. Evidently, a constant rate, $CR$, is when $\lambda \sim n/\varepsilon$.

All humans and engineers, in particular, would like to believe they learn from their mistakes. The past rate of learning determines our trajectory on the learning path, and thus how fast we can descend the learning curve determines where we are on the learning curve. If changes are due to feedback on learning from our mistakes and if we have not sustained a learning environment, an increase in rate signifies forgetting.

A general accident theory, like the Learning Hypothesis, should be testable, able to provide real predictions and explain actual error data.

### 3. Results, predictions and implications

Analysis of failure rates from many HTS in the modern world shows the importance of the rate of learning. The ‘learning hypothesis’—that we learn from our mistakes—allows a new determination of the dynamic human error rate in HTS, derived from all the available world data. Key data sources include US Bureau of Transportation Statistics, TSB Canada, US NTSB, US Railroad Association, US FAA, DOT NHTSA, UK CAA and ex-DETR; Australia’s outstanding ATSB; plus data from the EU, Australia, New Zealand, Norway, accident encyclopedias (such as Berman for ships, Hocking for aircraft, etc.), and many private, industry associations and commercial sources. Many other government safety licensing and investigative organizations have relevant data files and reports (US OSHA, ILO, US CSB, US NRC, USCG, US DoD, etc…). One of the major challenges and a major success is bringing all this information and veritable data mine together using a coherent system of analysis, and thus provide a unified and testable result.

From this simple learning hypothesis idea we have found that the rate of outcomes with experience (the error or failure rate) decreases exponentially with increasing experience, $\varepsilon$, as given by the MERE. If the failure (error) rate were constant, it would behave like population extinction. The failure rate solution leads directly to the ULC, which describes the learning path. It has become almost dogma that for continuous improvement (read ‘learning’) to occur, we must have a challenging and aggressive goal (read ‘zero events, accidents, or non-conformances’). What we argue, and what the data say, is that we all really need to know:

- where we are on the universal learning curve,
- how we can practically adopt a useful measure for our experience,
- how fast we improve (learn) or not, compared to others,
- what actual reduction in events our measures are achieving numerically,
- how close are we to the nirvana of the lowest possible rate,
- what goals should management set and expect,
- how we should improve our systems, safety and learning environment,
- how we should measure and report outcomes, and
- the probability of failure and of the next event.

One other key consequence arising from the world’s data over the last two centuries of observation is that ‘zero defects’ is actually unattainable. This residual finite rate to the non-removable component is due to human error and chance. That means a minimum rate, despite our best efforts, is observed about 1 event per 100,000 to 200,000 experience hours (an outcome frequency of $5 \times 10^{-6}$ per experience hour). This is the lowest that the safest HTS attain (see Table 1). For

### Table 1
Comparison and summary of derived minimum failure frequencies

| Event Type          | Minimum Frequency: One in Approximately … (hours) | Data source(s)                  |
|---------------------|--------------------------------------------------|----------------------------------|
| Fatal air crash     | 220,000                                          | World airlines 1970–2000         |
| Air event           | 20,000                                           | UK events 1990–1998              |
| Mid-air near-miss   | 200,000                                          | USA, UK & Canada 1987–1999       |
| Fatal train crash   | 20,000                                           | US Railroads 1975–1999           |
| Boating accident    | 350,000                                          | US data 1960–1996                |
| Ship sinking        | 350,000                                          | World shipping 1972–1997         |
| Auto accident       | 6000                                             | US data 1966–1998                |
| Auto fatalities     | 130,000                                          | ATSB data 1989–2002              |
| Industrial injury   | 30,000                                           | US data 1998                     |
| ACSI injury         | 200,000                                          | UK and US data 1970–1997         |
| Medical error       | 25,000                                           | US data 1983–1993                |
| Major Blackout      | 106,000                                          | US-Canada data 1966–2003         |
| Tire failure death  | 25,000                                           | US data 1991–2000                |
example, data from the failure of (material) pressure vessels in service also give a lowest rate of one per 175,000 hours. These are actually material objects, and are just another HTS, although we are used to thinking of them as purely mechanical devices. But they fail due to errors of operation, inspection, maintenance and design.

The presence of the ULC explains the trends noted with the hypothesis of potential risk compensation; and that the world accident data is largely invariant between societies and over time.

Best estimate values and relationships can be derived for both the HTS error rate and for the probability of failure in the system. Moreover, for perhaps the first time, we are in a position to be predictive about the rate and probability of failure, in any HTS simply based on our learning history.

### 4. The probability of failure

There are two main approaches or theories for error correction. In the first one, mistakes happen because we are human. Therefore, errors are natural, and we need them in order to learn. Improvement occurs through and by learning and so we need a ‘learning environment’, this is the theme of this paper. In the second, mistakes are avoidable because we are human. So someone or something is to blame. Therefore, corrective actions and precautions will ensure it cannot ‘ever happen again’, and we need more and better procedures, training, laws and rules.

These two theories represent: trial and error by learning versus the precautionary principle of zero risk.

The failure rate solution of the MERE explicitly defines the probability of human error as a function of experience. The initial rate is given simply in the limit of rare events by $\lambda(\varepsilon) = \lambda_0 = 1/\epsilon$ for the very first or initial outcome.

Now as large experience is accumulated, eventually we have an outcome since $p(\varepsilon) \to 1$. The certainty of ultimately having an event is independent of how much learning that has occurred, and the experience at when it occurs only depends on the attainable minimum rate (i.e., $p = 1$ when $\varepsilon \to N/\lambda_m$). Thus, we are indeed doomed to failure (the ‘normal accident’ pessimism), but the event is most likely to occur later if we are learning and not forgetting, or we have attained the lowest rate (the ‘learning hypothesis’ optimism).

Thus, by using good practices and achieving a true learning environment, we can effectively defer the chance of an accident, but not indefinitely. Moreover, by watching our experience and monitoring our rate, understand and predict when we are climbing up the curve.

### 5. The statistics of prediction

We know that we can only describe the average statistical behavior of physical systems. The resulting indeterminacy of outcomes is well recognized in physics and materials sciences. Using the ideas of Boltzmann’s statistical mechanics we can reconcile the stochastic occurrence of events with the systematic effect of learning. The measure of disorder in the observed outcomes is the ‘information entropy’, $H$. All error configurations or distributions of outcomes (microstates) are equally likely, and hence can occur in any observation interval as a distribution of outcomes with total experience. The outcomes microstates are observed stochastically but are also a systematic function of experience. The total number of outcomes (errors and failures), $N = \sum n_i$, are conserved for the observation range, and the total experience, $\varepsilon = \sum n_i \varepsilon_i$, in the observation range is finite and conserved for the observed outcomes. Since the observed distribution of the microstates with experience that exists is, on average, the most likely, then we have the usual result of an exponential distribution of the number of outcomes with increasing depth of experience.

In terms of the rate of outcomes, we find for a given experience interval the outcomes occurring at a rate:

$$n_i / \varepsilon_i \approx (n_0 / \varepsilon) + (n_0 / \varepsilon) e^{-\beta}$$
whereas the MERE suggests the outcome rate decreases as experience is gained,

\[ \lambda = \lambda_m + \lambda_0 e^{-kr} \]  

(11)

Hence, we presume that the effective learning rate equivalence is, theoretically for the distributions to merge or be equivalent

\[ k = \beta \]  

(12)

Thus, the new statistical theory has indeed reconciled the random occurrences of outcomes (as microstate distributions) observed as a result of unpredictable circumstances, with the systematic and predictable trend simultaneously imposed by learning with increasing accumulated experience.

### 6. Observations and conclusions

The Four Steps to predicting events are:

1. Step 1 providing a learning environment and culture;
2. Step 2 measuring where you are on the ULC;
3. Step 3 predicting the future rate;
4. Step 4 attaining the minimum as quickly as possible

The learning hypothesis is entirely consistent with the precepts of science and engineering. In fact, we close with a key quotation by a pioneer thinker Henry Petrowski, as quoted in ‘Search for Safety’ p.83 by Aaron Wildavsky, Transaction Publishers, USA, 1991:

I believe that the concept of failure-mechanical and structural failure—is central to understanding engineering, (which) has as its first and foremost objective the obviation of failure. Thus the colossal disasters that do occur are ultimately failures of design, but the lessons learned … can do more to advance engineering knowledge than all the successful machines and structures in the world. Indeed, failures appear to be inevitable in the wake of prolonged success, which encourages lower margins of safety. Failures in turn lead to greater safety margins, and hence new periods of success. To understand what engineering is and what engineers do is to understand how failures can happen and how they can contribute more than successes to advance technology.

The results clearly demonstrate that the human error probability is dynamic, and may be predicted using the learning hypothesis. The future probability estimate is once again derivable from its unchanged prior value, based on learning, and thus the past frequency predicts the future probability.

### 7. Conclusions

- We can make future failure rate predictions
  - Using a model enables average predictions to be made and confidence levels to be estimated, based on theoretical grounds and not on arbitrary fits to the data or extrapolations.
- We need a Learning Environment
  - A sustained ‘learning environment’ is essential to the process of the reduction of human errors, not just a system of fines, blame and punishments, since zero is not attainable. An open reports system with confidential reporting is vital.
- We need mistakes and failures to advance
  - The importance of human error in engineering failures is supported by the world data, for all know modern technological systems from the Industrial Age over the last 200 years, including often large and significant changes in technology.
- We need more and better data and analysis
  - The hope is to encourage detailed analysis that will lead to more and better ways to record, track, analyze, prevent and predict such engineering failures, losses, accidents and events.

### References

[1] R.B. Duffey, J.W. Saull, Know the Risk, first ed., Butterworth and Heinemann, Boston, MA, 2002.