

Duffey, Romney

Atomic Energy of Canada Limited, Chalk River, Ontario, K0J 1J0 Canada

Managing and predicting risk, safety and stability in a challenging world

Keywords

accidents, errors, experience, learning, outcomes and risk

Abstract

It should be obvious that we must learn from our mistakes, so all of society, and ourselves, should have progressively safer, less risky systems and behaviors as we learn. Accidents are seemingly random in their occurrence, but in fact, this very apparent randomness is also containing information. The information we have researched and analysed covers nearly 200 years of knowledge from literally millions of multitudinous observations. The failure rate provides the expression for the probability of any outcomes, and the resulting curve is called the Human Bathtub. By quantifying the randomness, the uncertainty and the disorder, we have provided a new objective measure of “safety culture”, “organizational learning” and “engineering resilience”. We have linked individual learning and skill acquisition to the systematic risk reduction observed for entire systems with increasing experience. The results will be of interest to those interested and engaged in risk management, and in the social sciences where risk perception is important.

1. Learning from disasters and dangers

Plane crashes, automobile accidents, train derailments, medical malpractice, bridge collapses, stock market crashes, building fires, and product cost escalations – all have a common cause. It is the failings, foibles, and fallibility of us, the human drivers, the bankers, builders, doctors or pilots, the speculators, lawyers and regulators. We are all an integral and inseparable part of the transportation, manufacturing, financial, production and social systems in which we work, perform, drive, fly, design, invest, use or operate, or just have a job. Conversely, as opposed to our many failures and errors, we have successes and achievements. Avoiding losses, accidents and tragedies, reducing event rates and even the near-misses, we achieve success by learning, as we gain skill and knowledge by and from our accumulated *experience*. This learning from experience includes all the existing and pre-existing knowledge and skill acquisition for the system.

It should be obvious that we must learn from our mistakes, as intelligent creatures are supposed to do, correcting our errors and mending our ways, training ourselves, using our both good and bad experiences.

We correct our conscious and unconscious mental models of the world and our reaction to it, and hopefully shift our behavior as we gain knowledge, understanding and skill. So all of society, and ourselves, should have progressively safer, less risky systems and behaviors as we learn. Knowing what we know, we should be able to predict when an accident or tragedy will occur to us, and what to do and how to manage our risk exposure, or reduce our chance of harm, be it physical, fiscal or psychological. After all, we have had many centuries of learning, developing and using technological machines, the lawn mowers, cars, airplanes, rockets, trains, controls, and signals of our everyday world. As individuals in any system, and as individual human beings, we learn to survive, improve and gain knowledge.

The number of casualties, deaths or injuries – heart rending though they are – is purely yet another arbitrary statistic, at most a media moment and a highly personal tragedy. Some events in our lives, like medical errors, being stuck in a hotel elevator or a fire, experiencing a train derailment or a aircraft near-miss, are not due to anything under our real control, as we offer our bodies or ourselves up for

transport, diagnosis, treatment, practice and operations. Nothing we do is “risk free”, so as a society we have to tolerate some risk (a chance accident or outcome) and just hope it does not happen to us as an individual. But we do not know if it will, although we can simply make a statistical estimate that it is unlikely to happen to ourselves, since we are just one person amongst many risk-exposed people. We make or avoid such decisions every day.

Massive inquiries into the more spectacular events, like the UK Ladbroke Grove railway collision, the USA Three Mile Island reactor melt down, the Space Shuttle losses, the Buncefield fire, the Toulouse and Texas City explosions, all these and others all show the cause is due to an unexpected combination or sequence of human, management, operational, design and training mistakes. Once we know what happened, we can fix the engineering or design failures, and try to obviate the human ones by fines, rules and punishments, hoping in a very human way that it “will not happen again”.

Knowing what we know, what we have experienced up till now, is it actually possible to predict the future? The possibility of an accident, or even a repeat? The price of a product, and whether it will rise or fall? How likely is it that something will actually happen? When will we even see and measure it? Hidden in the seemingly impersonal data of the events that happen to “others”, behind the almost routine reporting of the statistics of deaths, injuries, accidents, costs, crashes and errors is the valuable and key information that we must not neglect, that we ignore at our peril. It is the learning or forgetting trends for both the number and the rate of events as a function of our experience. In predicting these events, and proceeding further in our analysis, we are guided and governed solely by data.

2. Disasters, accidents and deaths: the learning hypothesis

Accidents and outcomes are seemingly random in their occurrence, like the crashes of the Space Shuttle or Concorde, the great North East Blackout, the Chernobyl reactor exploding, the auto crash down the road, the ding in the parking lot, the surgery that did not go as planned, the stock price that decreased not increased. They seemingly can occur as observed outcomes at any instant, without warning. But in fact, this very apparent randomness is also containing *information*. Since these events are due to a combination of human and technological system failures, working together in totally unexpected and/or undetected ways, occurring at some random moment, they actually exhibit overall the impact of our experience, personally and collectively with that

system, linking our very individual cognitive learning (that goes on inside our brains) with our entire social learning (inside our group, organization or even culture). This happens because we share and manage the results of our learning opportunities, and do that randomly also as we learn and unlearn what is right, wrong, harmful or risky to do.

The human rules governing the market and personal behavior, the tug and pull between supply and demand, the fine line or gray area between the safe and the unsafe, and the elusive social concepts and constructs of safety culture, risk aversion, and engineering system complexity are the stuff and content of our modern technical world, and which inform and colour the decisions we make every day.

Our own work [1], [2] has shown that what governs the outcomes we observe is the way humans behave and learn from their own and collective mistakes, as we are seemingly doomed to rediscover and make errors. Thus, the Learning Hypothesis states simply:

The rate of reduction of the error rate is proportional to the rate.

We introduce corrective and best practice procedures, training, rules, safety measures and management incentives to minimize the number and chance of mistakes. In cognitive psychology, performance and neurology there is a plethora of “models” for human decision making, psychological reasoning, mental learning, social behavior and decision making, including how we learn and behave under stress, make judgments, take actions, gain skill, plus how we might train, and how we may use our skill, rule and knowledge base to improve. All of this material and literature is largely empirical and qualitative. The results of repetitive tests on learning and pattern recognition by individual subjects empirically correlate and explain the improvement in skill and response with repetitive trials, just as surgeons and pilots gain skill and reduce errors from repeated practice, both in real life and in simulators. The result of the Learning Hypothesis is the Universal learning Curve (ULC) that is shown in Figure 1. As both systems and individuals (after all a system is just a collection of individuals), we progress and gain skill by going down the curve, transitioning from being a novice/learner to becoming an expert at sufficiently large experience, finally achieving a finite, non-zero outcome rate.

3. Proof of principle: validation of the universal learning curve

Contrary to many statements and perceptions, there is no shortage of data on human failings and mistakes – so we have deliberately cast a wide net [1], [2]. We have thus been able to make a

measurable and testable prediction about the *universal* rate at which mistakes (outcomes or errors) are made or will occur, and that rate is determined by our individual and collective accumulated experience. The information we have collected and validated uses officially and publicly available data that generally show a definite learning trend, and support the universal applicability. We know of no other validation exercise of such breadth and extent, and not only can the data now be normalized on the same chart, but can and does also show unambiguously that they all follow the same learning trend. Previously hidden in and by the yearly trends that are usually discussed and reported, what really matters is to adopt the relevant experience measure which is *not* usually calendar time. The information we have researched and analyzed covers nearly 200 years of knowledge from literally millions of multitudinous observations [1], [2]. The entire learning and probability datasets now include the following sources and outcomes, with the historical time spans shown in brackets:

- From official government and industry websites USA auto passenger deaths numbering 181,086 (1975-1996); 833,494 railway injuries (1975-1999); 31,770 coal mining deaths (1938-1998); some 8696 oil spills of over 1000 gallons each at sea (1961-2001); 3386 airline near-misses (1980-1997); 6784 recreational boating deaths (1960-1996); and some 3386 reported mid-air near-misses (1980-1997).
- More globally, the learning data set we have amassed now contains multiple technologies worldwide: coal and gold mining injuries in Australia, UK and South Africa (1969-1999); worldwide over 20 million pulmonary disease deaths (1840-1970); some 284,797 cataract operations in Canada (2001-2003); 283 infant heart surgeries in UK (1984-1995); a world-wide total of 1882 rocket launches (1962-2005); some 46,510 auto deaths (1981-1999); 247 pilot deaths in Australia (1990-2005); fatal crashes for over 210 million commercial airline flights (1970-2000); on UK railways some 1652 train derailments (1988-1999); 4568 danger signals passed (1994-2000); 1359 Canadian mid-air “air proximity” events (1989-1998); and the anti-missile interception and destruction of 3861 German V1 bombs (1944).
- Systems that do not show significant learning, as measured by decrease or declining loss and error trends include adverse medical events (errors), a sample of some 376,962 deaths in reported studies (1975-2005) of where work habits, traditional practices, legal issues, management and liability pressures, and patient confidentiality

constraints all defy openness and error reduction. Also a nearly constant rate of one per thousand shipping-years pervades the thousands of global shipping losses at sea (for 1800-1950 and 1972-1997) where the continuing influence and reliance on the human element overrides massive changes in technology and the robustness of ship design. In each case, large efforts are underway nationally to understand the lack of learning trends and reduce the event rates, for many commercial, perception and genuine safety reasons.

- Data for individual actions (as opposed to system outcomes) are available from many thousands of individual human subject task and learning trials in the psychological literature (1930-2000). These have established the rate of skill acquisition via Laws of Practice, which we have shown are consistent with the ULC, and also the reduction is response time with repeated trials. Individual surgical skills do however improve with practice, as shown by data for cataract and heart surgeries. In addition, the system learning behavior mirrors that of the individuals within. The predicted probability of operator error agrees with about 900 published French nuclear plants events (1997-1998); and with 55 or more simulator tests conducted for nuclear power plants (2003). Recovery actions for power restoration for 148 power losses at over 100 US nuclear power plants from 1986-2004 are also in agreement, as well as the power blackout repairs probability for New York Queen’s Borough that affected (i.e., disconnected) some 175,000 customers over a period of several days.

These are the cases for which data exists – and they form a compelling picture of the learning effect on error, accident and event rates (*Figure 1*).

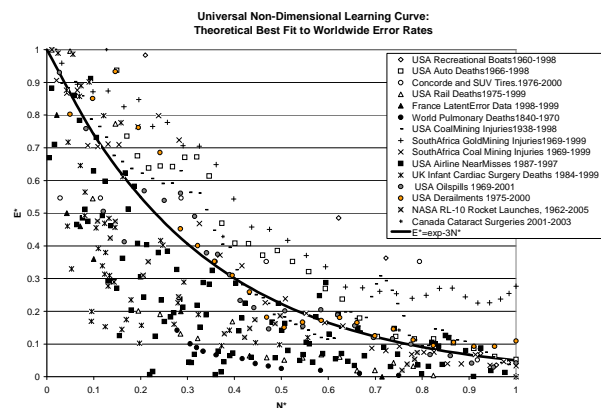


Figure 1. The Universal Learning Curve.

In addition to selecting the relevant and known measure for experience or risk exposure, the two

“free” parameters in the model derived from the data are important, and are:

1. The learning rate constant, k , which impacts the shape of the exponential, and a “universal” or average value value of $k \sim 3$ seems to be the best fit to the meta-analysis of all data, but individual sets may of course have slightly superior fits using slightly different values.
2. The minimum error rate, λ_m , which is the baseline or lowest attainable value, which is about one per 200,000 experience-hours as derived from the commercial aircraft fatal accidents, mid-air near-misses, boiler failures, licensing paperwork errors, and auto accidents for experienced (older) drivers.

With these two facts derived from observation and tested against data, we are able to make predictions.

4. The probability of failure or success: the human bathtub

The observed, prior or past data allow us to calculate the probability of any outcome as a function of our accumulated experience. The failure rate provides the expression for the probability of any outcomes, and the resulting curve is called the Human Bathtub because of its shape (*Figure 2*).

Basically, the probability starts out high when we are a novice with little experience and there is an almost equal chance of making a mistake or not (i.e., about 50:50). As we learn and gain experience, we descend into the bottom exponentially, depending on the rate of learning. Generally, this is nearly a constant factor, so the indicator of our risk probability varies with experience until we reach the minimum probability (of about one in a thousand). Ultimately, we climb out of the bathtub, the probability increasing again, due to the non-zero minimum rate. Eventually at sufficiently large experience, and despite being an expert, we are doomed to have an outcome (a probability of unity) since the total risk exposure (experience) interval is now very large.

This certainty of an outcome is not a reason to be depressed or fatalistic – we can defer reaching this certainty by having a sufficiently low minimum rate, and making changes to our system(s) if and when we detect or measure the increase occurring.

The available data again support this trend, but only when there is a meaningful measure available or recorded for experience. Demonstrated examples of the relevant experience measures include:

- a) numbers of flights for commercial airline fatal accidents (1970-2000);
- b) amount of oil shipped by sea for oil spills (1973-2000);
- c) time into the transient for nuclear plant events (1997-1998);

- d) launch and burn time for rocket failures (1962-2005); and
- e) restoration/recovery time for nuclear plant power losses (1986-2004).

The resulting Human Bathtub curve is compared to these data in *Figure 2*, from which it is self-evident that the data and theory are in accord. Since the probability of success is the complement or opposite of failure, the probability of success is automatically reclaimed.

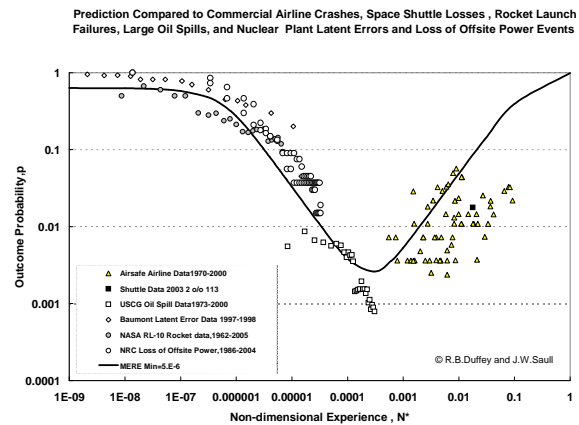


Figure 2. The Human Bathtub

5. Predicting the future and the unknown: using missing information

Engineers, regulators, professional bodies, legislators and consultants use a myriad of techniques to continuously improve our technological system designs, ease of operation, safety of performance and the margins to component and system failure. 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 actions. 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.

In the fiscal world, various risk management techniques are adopted to minimize risk exposure, like varying the portfolio of products, not betting on one outcome or product being successful, or building in margins. The risk that an investor is prepared to take is then estimated by balancing the exposure (the downside) against the gain (the upside), just like trying to make a rational decision about the future when it is uncertain and unknown. This has been called in financial circles as the “Black Swan”,

which occurrence probability we have shown is in fact calculable on the basis of sampling the known and unknown outcomes [2].

This whole issue of personal risk and predicting when and where an event will occur cannot be simply solved using such human-centered rule, regulation, recommendation and design practices to make items and technology safer and simpler to use. It also cannot be solved by inquiries into cause of failure; or meting out punishment for failure; or legal and fiscal assignment of responsibility. These are all case-by-case, item-by-item, after-the-fact, without general application. Instead, we need to understand how we must and do learn from our mistakes and improve our safety and performance as we go, so we may anticipate and prevent using prior knowledge.

With a firm basis, we can now make predictions, using Bayesian reasoning and approaches. The formula can be expressed symbolically in words as: *Probability of a future outcome (Posterior, $p(P)$) is given by the historical or known probability (Prior, $p(\varepsilon)$) times the conditional chance of the event occurring (Likelihood, $p(L)$) given what has happened before.*

The Human Bathtub shown in Figure 2 provides a unique estimate of the prior, $p(\varepsilon)$, whereas the likelihood is given by the chance that the next event will occur in the next increment of experience, so that $p(P) \sim 1/\varepsilon$, or the simple rate at which outcomes are occurring randomly anyway as we gain experience or are risk exposed.

Now, because the outcomes and events (like the crashes of cars, planes, trains and stock markets) occur randomly, the number of possible ways they could occur or be observed as given by the information entropy, H . Classically, it can be shown that mathematically, H , is given by the summation of the known prior probability distribution (or symbolically, $H = -\sum p \ln p$). This quantity is a unique measure of uncertainty, and relates to the “missing information” and is uniquely related to the depth of experience that we as individuals and collective systems accumulate. In essence, and in a very real way, the information entropy, or H -factor, is both a measure of the random firing of neurons as we establish skill and knowledge patterns in our (plastic) brains, but also reflects and measures the random interactions and learning behavior of the myriad of individuals in any technological, corporate or working operation. Since the H -factor is a function of depth of experience, it exhibits the required learning trends and systematic reduction with increasing experience. In effect, we must have disorder (random happenings in the systems and neuron firings in the brains in that system), so that order, knowledge acquisition, skill and learning

emerge. This is analogous to the emergence of order in chemical and physical systems, and applies both to our individual brains and to our collective organizations.

By quantifying the randomness, the uncertainty and the disorder, using the H -factor, we incidentally have provided a new objective measure of those illusive but desirable attributes and concepts of “safety culture”, “organizational learning” and “engineering resilience”, which previously were simply subjective desiderata. But now for the first time we can be both objective and predictive, and compare to actual data. We have shown that the systematic data trends agree with the depth of experience concept, at least for the cases with data for commercial aircraft near-misses in the USA and UK, and automobile fatalities in Australia. The conformation of the use of the H -factor to elucidate learning trends and relative performance improvements has been achieved recently for four very disparate industries (see Figure 3):

- safety performance indicators from North Sea oil and gas offshore rigs;
- train derailments in the UK;
- underground coal mining injuries in the USA; and
- commercial airline near-misses in the USA.

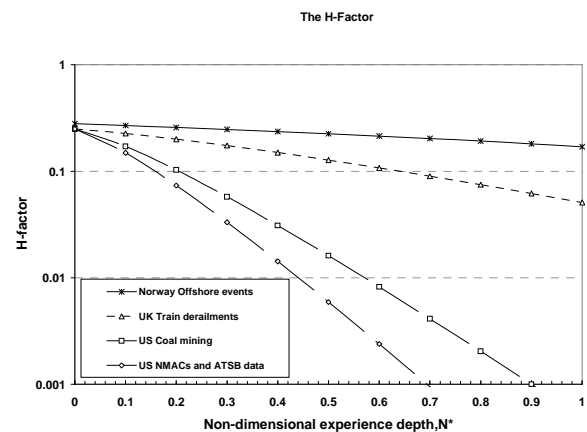


Figure 3. The Information H -factor comparisons

Using this H -factor as the single measure, the relative risk ranking is clearly in the order shown of actual and apparently decreasing risk (a) to (d). The implications of this type of plot or representation are very interesting. Our perception of risk seemingly follows these comparative trends also as some are regarded as more risky or dangerous than others, but note the very different learning behaviors, or lesser decreases with the increase in the non-dimensional depth of experience parameter, N^* . Apparently, both the slope (learning trend) and the H -Factor

magnitude (occurrence numbers) may be affecting our *perception* of relative risk. We thus look for, expect and appreciate learning and risk reduction with experience. But we also expect reductions in the number of events/deaths; so are apparently also more fearful of the larger numbers of events/deaths reported for some activities/systems than for others. So this new *H-factor* idea that we propose here is not in conflict, or at best is not rejected by these key example cases, and coincides with or reflects our risk perception. *We can suggest that humans can and do somehow quantify the basis for our relative risk perceptions based on our expectation of learning due to the number and distribution of outcomes.* These represent our collective and individual learning patterns and responses. Other factors, like fear, risk aversion or risk taking behavior, and personal risk exposure obviously may have an influence, and be stronger in some cases than others.

6. Conclusions: managing the future risk

We have derived the trends for the variation of risk as we gain experience and correct our mistakes. The rate and probability of events, outcomes, disasters are both a specific function and depend on our accumulated experience, as we learn and gain skill. The Learning Hypothesis is verified by extensive data comparisons. We have demonstrated the principle of a dependence of the distribution of the number of observed outcomes on the depth of experience, and that this trend is possibly reflected in our perception of risk.

We have linked individual learning and skill acquisition to the systematic risk reduction observed for entire systems with increasing experience. Of most importance to the present discussion is providing the basis for analyzing trends, and also provides a basis for understanding data trends and making future predictions. The results will be of interest to those interested and engaged in risk management, and in the social sciences where risk perception is important.

References

- [1] Duffey, R.B. & Saull, J.W. (2002). *Know the Risk*. First Edition, Butterworth and Heinemann, Boston, USA.
- [2] Duffey, R.B. & Saull, J.W. (2008). *Managing Risk: The Human Element*. John Wiley & Sons Ltd., West Sussex, UK.