



Is big data risk assessment a novelty?

Paul Swuste

To cite this article: Paul Swuste (2016) Is big data risk assessment a novelty?, Safety and Reliability, 36:3, 134-152, DOI: [10.1080/09617353.2016.1252084](https://doi.org/10.1080/09617353.2016.1252084)

To link to this article: <https://doi.org/10.1080/09617353.2016.1252084>



© 2016 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 06 Dec 2016.



Submit your article to this journal [↗](#)



Article views: 1574



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 3 View citing articles [↗](#)



ARTICLE

OPEN ACCESS

Is big data risk assessment a novelty?

Paul Swuste

Safety Science Group, Delft University of Technology, Delft, The Netherlands

ABSTRACT

Objective: What metaphors, models and theories were developed in the safety science domain? And which research was based upon 'big data'?

Method: The study was confined to original articles and documents, written in English or Dutch from the period under consideration.

Results and conclusions: From the start of the 20th century, human error was a dominant explanation for causes of occupational accidents. Although external factors were seen as main contributors, it was not till after World War II when scenario analysis was conducted in detail. The main drivers were the upscaling of the process industry in this period, as well as the introduction of high hazardous industries, like aerospace and nuclear sector, and consequently disasters occurring in these sectors. Already from the beginning, big data research was no exception in the safety science domain. 'Big' in this context is defined by numbers.

KEYWORDS Safety science; history; high hazard industries; big data

Introduction

Big data is a fashionable term among scientists, marketers, forecasters and safety experts. With the current developments in computing power and automated analytical methods, vast amounts of data can be exploited and analysed to gain insight into almost anything, including risks, hazards and dangers. The rail sector is also generating massive volumes of data through all sorts of sensors and automated devices. Those in favour of big data point to its advantages; it provides other ways of thinking about and looking at data. No theories or models are required to gain new insights based upon correlations. 'Let data speak for itself' seems to be the motto. As mentioned by the RSSB, this creates a dilemma in the form of the causality-based approach as opposed to the correlation-based approach (RSSB, 2014).

CONTACT Paul Swuste p.h.j.j.swuste@tudelft.nl Safety Science Group, Delft University of Technology, Jaffalaan 5, 2628BX, Delft, The Netherlands

© 2016 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Data are simply defined as ‘raw facts’, signs or symbols, or observations which are unorganized and unprocessed and therefore have no meaning or value because they lack context and interpretation. To transform data into information, some sort of classification is needed. In the case of risk assessment, models and metaphors linked to major accident processes can provide such a classification, giving a context and an explanation for the data collected. If taken one step further in this hierarchy, information becomes transformed into knowledge but this requires validation. Knowledge, placed at the top of the knowledge pyramid, is based upon theories of accident causation, major or otherwise, thus facilitating sound prediction of future events.

In the Safety Science domain there are some concepts that are popular amongst scientists, managers, lawyers and laymen. It is commonly believed that safety and major accidents are related and are sometimes causally linked to the behaviour of front line operators and workers, safety culture and safety management systems. The focus on behaviour began in the early 20th century, the American Safety First Movement (Aldrich, 1997; Hoffman, 1909; Palmer, 1926; Swuste, van Gulijk, & Zwaard, 2010), being its first and powerful promotor. In 1919 the accident proneness theory of Greenwood & Woods provided a scientific basis for what was termed the individual hypothesis that was used to explain accident causation. Since the INSAG report following the Chernobyl disaster of 1986 (INSAG, 1986; Guldenmund, 2010) safety and culture have been closely linked and safety management has had two advocates. The first is the well-known Robens’ report (1972), which recommended that the technical control of hazards should be delegated to those who create them; to industry. The second pertained to the Piper Alpha disaster of 1988, which received scathing criticism in the Cullen report because of the quality of the safety management of both the mother company, Occidental, and the offshore safety regimes, like that of Piper Alpha (Cullen, 1990).

Despite its popularity, even today, the individual hypothesis was heavily criticised in the academic press just before and after World War II (Vernon, 1936; Hale & Hale, 1970; Hale & Hale, 1972). The main objections were that the low correlations between psychological test results and accident figures did not provide proof of causality and that the focus was on only one exclusive factor of the accident process, the psychological stability of victims. This discredited the individual hypothesis in the academic press. The comments made on the other two concepts of safety culture and safety management (systems) differed from those made on behaviour. So far, with the exception of a few case histories, no convincing scientific evidence has been produced to link these two concepts to safety, or to safety levels within companies. Similar remarks apply to the two other concepts of reliability and safety indicators. High Reliability Organisations (HRO) were

postulated in the late 1980s by Weick, Rochlin, La Porte, and Roberts (Weick, 1987; Roberts, 1988; Rochlin, La Porte, & Roberts, 1987). Despite the extremely hazardous operating conditions, these HRO's managed to function without major accidents, and to operate as effective learning organisations. Safety indicators, or more precisely the lack of safety indicators, was cited as one of the contributing factors in the BP-Texas disaster of 2005 (Baker, 2007).

The main reason why these concepts of behaviour, safety culture, safety management and safety indicators either have the status of belief, or at least of scientifically unproven links to safety, is because of their weak or absent connection with accident processes. To be more accurate, the link with accident scenarios, major or otherwise, has never been substantiated, although it is also acknowledged that conducting research to prove such relations would be extensive and difficult. HRO might be an exception since the model is based on just a small number of case studies in a few sectors, mainly aircraft involving carriers and air traffic control.

This chapter will present a bird's eye view of the scientific developments in the safety science domain. It will be restricted to developments in metaphors, models and theories on accident causation (Swuste et al., 2010, 2015; Swuste, Van Gulijk, Zwaard, & Oostendorp, 2014). The two research questions below will be central to this paper:

What metaphors, models and theories were developed in the safety science domain?

Which research was based upon 'big data'?

Timeline of safety science theories, models and metaphors

Early days

Occupational safety became an item in the 19th century at a time when the United Kingdom led industrialisation with its great technical discoveries. Various British commissions reported on working hours in the textile industry and so that led to the start of social legislation in 1802. The installation of the British Factory Inspectorate, responsible for legislative supervision, dates from 1833.

From 1844 onwards, the Inspectorate was also bound by law to monitor safety in factories, such as various forms of machine and installation protection (Hale, 1978; Le Poole, 1865). In this period, occupational safety was turning into a professional field. Engineers started enclosing moving parts on machines and fencing off heights to improve safety techniques. The publications on occupational safety written at that time were very practical (Calder, 1899). These publications did not provide any theoretical analysis

on the causes of accidents. Implicitly it was assumed that heights and mechanical moving parts were causing accidents.

The period up until world war II

After a short while the United States followed the United Kingdom by adopting the above-mentioned national campaign of the Safety First Movement with such ploys as 'the safe road to happiness' poster and the Pittsburgh survey of 1906–1907. Occupational safety became a vehicle for efficient production and many initiatives were started in that period, like the formation of the National Safety Council, the Safety Museum, the professional 'Safety' journal, national safety congresses and safety medals for companies that exemplified best safety practices.

A whole range of books and publications were also published that dealt with practical safety issues for specific branches of industry and there were general reference books that addressed the managerial aspects of safety (Swuste et al., 2010).

The Pittsburgh survey (Kellogg, 1909; Eastman, 1910) was the first sociological survey in the United States on the living and working conditions of workers in the steel district of Allegheny County in Pennsylvania, US. The survey constituted the first extended analysis of occupational mortality and more than 520 fatal accidents were examined over a one year period. The results advocated the environmental hypothesis which focussed on the external causes of accidents, such as very long working hours, overcrowded workplaces, dangerous machines and the increased pressures of work and speed of production. The 520 examinations were the big data projects of their day.

The difference between the environmental and the individual hypothesis remained an active point of debate until after World War II. DeBlois, chairman of the safety committee of Dupont de Nemours was a strong advocate of the environmental hypothesis. His 1926 book stated that if similar accident scenarios were apparent in a company then that indicated that management was not taking safety seriously enough. So, repeated accidents were linked to mismanagement. He was not in favour of the Safety First Movement, risks and risk taking were considered to be an essential part of the process enabling people to learn.

The important contribution made by DeBlois pertained to his assumptions on accident causation and his general rules for prevention. Accidents should be seen as a consequence of a sequence of events which, either directly or on the long term, would cause harm and damage. For the first time, accident causation was viewed as a process guided by accident scenarios. Hazard was what formed the basis of any accident and hazard was equivalent to kinetic energy potential or could be of a mechanical, electrical

or chemical nature (DeBlois, 1926). Unfortunately, he had to conclude that there was insufficient data to support predictions of accident occurrences so his ideas remained theoretical.

It was different for another influential safety thinker. Heinrich was an advocate of the individual hypothesis. In the same period as DeBlois he published comprehensive ratios on accident costs, accident causes and accident mechanisms. The indirect costs of accidents were four times higher than the costs of compensation. Based on 12,000 randomly selected insurance records of his own Travel Insurance Company and 63,000 reports of factory owners he found that most accidents could have been prevented: 88% of all accidents were caused by unsafe acts on the part of workers. From 50,000 accident reports he established a fixed relationship between no injury accidents, minor and major injuries (Heinrich, 1927; Heinrich, 1928; Heinrich, 1929). In 1941, in the 2nd edition of his reference book on safety, the well-known domino metaphor appeared (Heinrich, 1941), with the removal of unsafe acts as a primary prevention strategy which could be simply effected by selecting and training workers (Figure 1). Heinrich may be seen as an early adaptor of big data and he used it to its full advantage: the massive amount of data in this investigations ensured that this theories would reverberate around the industry for close to half a decade.

In the United Kingdom the environmental hypothesis gained support. Vernon's (1936) reference book on safety addressed the influence of temperature, fatigue, speed of production, ventilation and alcohol consumption on safety (Swuste et al., 2010; Vernon, 1936). Again, this work was based on extensive datasets from investigations in factories.

During World War II operational research was developed, a mathematical and statistical approach aimed at solving military problems. After the war these techniques were applied in the private sector to support management decisions on, for instance, production planning in various branches of industry (Moore, 1949), and later also in reliability engineering.

The post-war period

The post-war period saw five other developments. The first was the influx of American physicians into the safety domain. They questioned the lack of progress surrounding safety research and accident prevention and went on to introduce the epidemiological triangle (Figure 2), a model that was very effective in the fight against cholera in the nineteenth century (Swuste et al., 2014; Gordon, 1949; Haddon, 1968).

Prevention was achieved by changing the corners of the triangle, or blocking their links. The second development was Heinrich's management of safety model (Heinrich, 1950), to ensure safe and efficient production. Thirdly, the Dutch physician Winsemius, addressed man-machine relations

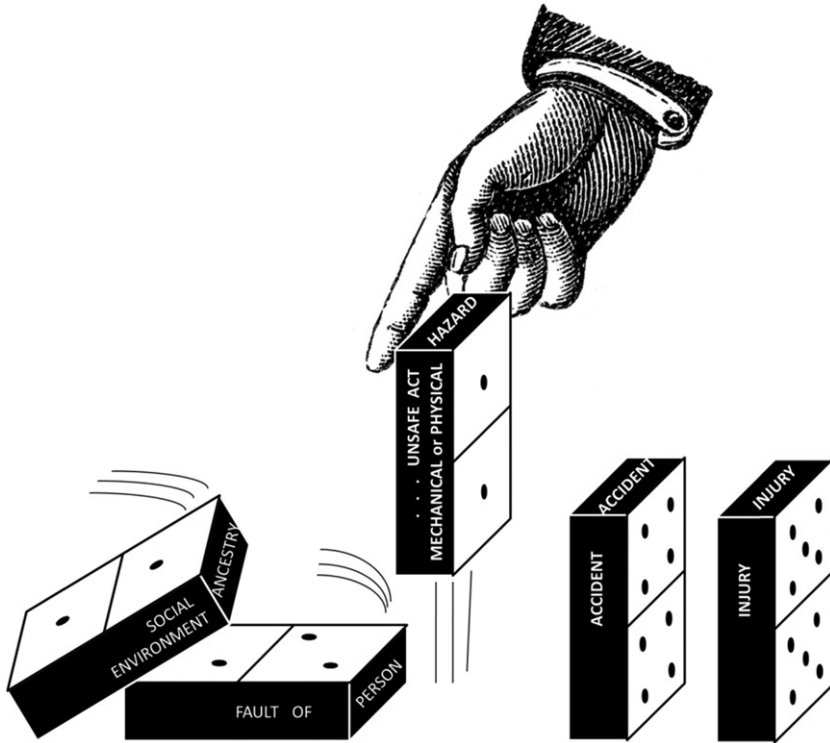


Figure 1. Safety dominoes (after Heinrich, 1941).

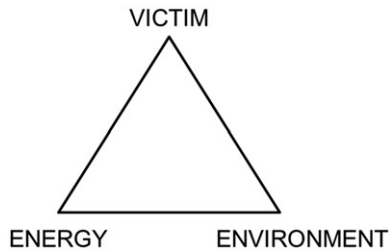


Figure 2. The epidemiological triangle.

in his research into 1300 accidents at the former Dutch steel works Hoogovens. The huge amount of data allowed him to develop a theory that postulated that human behaviour and unsafe acts are response reactions on the part of workers during process disturbances; such behaviour and acts were a consequence of context and not a cause of accidents. He was the father of 'task dynamics theory' (Swuste et al., 2014; Winsemius, 1951). The fifth development related to the introduction of the concept of a psychological climate and was based on 5 years of lost time accident data

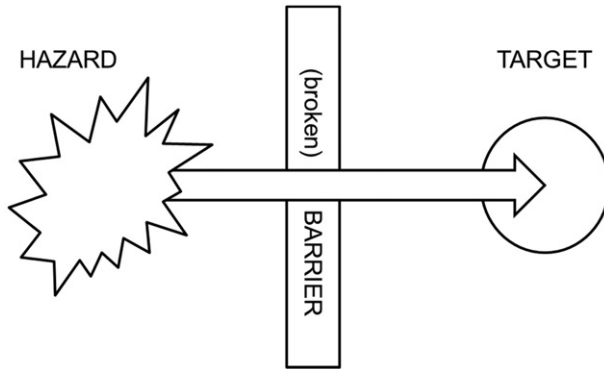


Figure 3. The Hazard-Barrier-Target model.

involving 7,100 workers (Keeman, Kerr, & Sherman, 1951). Huge datasets and dedicated research yielded reliability engineering. The focus of safety science had shifted to quality control and the reliability of electronics (Saleh & Marais, 2006). Ten years later, the well-known hazard-barrier-target or hazard-barriers-accident model was introduced (Figure 3).

Following the DeBlois notion of hazard being energy, barriers came to be viewed as physical entities stopping or reducing the energy flow of the accident scenario (Gibson, 1964; Haddon, 1962). The model was a logical extension of the epidemiological triangle. The term 'target', implied that there were additional effects, apart from injuries. Bird introduced the damage triangle which was similar to Heinrich's accident mechanism ratios (Bird & Germain, 1966), only with different numbers.

At the same time in the military domain operations were becoming increasingly complex and the traditional fly-fix-fly approach, which had until then been customary in engineering, became obsolete.

The same was true of the process industry, where a massive upscaling of processes had increased complexity, and consequently also the accompanying risks. A movement was initiated to increase system reliability. Safety techniques were developed, mainly originating from the military domain, Loss Prevention started in the process industry, and became Reliability Engineering in what was applied aviation and the nuclear sector (Swuste et al., 2014). With Loss Prevention and Reliability Engineering a probabilistic approach had entered the safety domain. In the following period, the 1970s, safety became a hot item.

Disasters in the process and the nuclear industry received ample attention in the media in Western countries. Public resistance to industries and companies that could not control their processes grew, leading to disasters and environmental pollution. In scientific safety literature the term 'safety

management' was introduced together with safety audits (Petersen, 1971), concepts such as loose and tightly coupled processes (Reeves and Turner, 1972), and organizational culture (Turner, 1971). Organisational culture preceded the construct of safety culture, which was developed later. As had already been mentioned by DeBlois, the notion was clear that major accidents had multiple causes as illustrated in the pre-bowtie diagram of Nielsen (Nielsen, 1971) (Figure 4). These causes were not necessarily technical factors. Both in the United States and in the United Kingdom, attention was drawn to managerial and organisational factors as aspects of major accident scenarios.

Johnson, when conducting safety research in the nuclear industry, formulated a comprehensive definition of accident scenarios (Johnson, 1970):

'An accident is the result of a complex series of events, related to energy transfer, failing barriers, and control systems, causing faults, errors, unsafe acts, and unsafe conditions and changes in process and organisational conditions'.

Also the Management Oversight Risk Tree (MORT) derived from the nuclear industry, pointed to changes and errors made by supervisors, and to managerial and planning issues as preconditions for employee error (Johnson, 1973). A few years later, came the British researcher Turner who studied a substantial dataset in an aggregated form (84 governmental reports on major accidents in civil aviation, trains, ships and mines) introduced the concept of disaster incubation time: the notion that the result of mechanisms could blind organizations to weak disaster scenario signals (Turner, 1978).

The continuing series of major accidents in the 1980s had a stimulating effect on safety research. In that period safety engineers, risk researchers, psychologists, and sociologists continued to develop new models, metaphors and theories. In occupational safety, the importance of the safety climate was stressed (Zohar, 1980) thereby revitalizing similar concepts from the 1950s. Also in line with Winsemius, from the same period, the importance of process disturbances as causes of accidents was articulated in Swedish studies (Kjellén, 1984).

Simultaneously, Kaplan, and the Reactor Safety Study WASH-1400 developed a method to try to estimate risks based on failure data which was now gathered on a huge scale in some industries. This information could be used in the new risk formula, the risk triplet, which combined major accident scenarios with the deterministic approach and the probabilistic approach (Rasmussen, 1975; Kaplan & Garrick, 1981):

$$R = \{ \langle s_i, p_i, x_i \rangle \}, \quad i = 1, 2, \dots, N.$$

where s_i is a scenario identification or description; p_i is the probability of

that scenario; and x_i is the consequence or evaluation measure of that scenario, i.e. the measure of damage.

Another point was the rapid changes in the organisation of labour within big companies. With their increasing complexity, and with the automation of activities and process, the role of front line operators and workers had changed dramatically. Automation had already started in the 1960s. Instead of operating machines, workers' activities were reduced to controlling processes, and only interfering during abnormal conditions.

There was little understanding of human functioning in such complex technological systems. In the 1980s Rasmussen and Reason developed the skill-rule-knowledge theory (Rasmussen, 1982; Rasmussen, 1987), while Singleton addressed the man-machine interface. If operators were faced with a high degree of complexity in terms of equipment failure or other abnormal conditions then the design of the man-machine interface had to be supportive for the operator and the task that was expected of him or her (Singleton, 1984).

Sociologists had a different focus, they looked not so much at human interaction with technology but rather at indicators of major accidents, either within the organisation, as Turner did in the late 1970s, or at the technology itself. Perrow took a similar approach as Turner by analysing hundreds of accident reports in the process industry, in air and sea space, dams, mines, weapons and recombinant DNA research. He first came to the conclusion that 'great events may have small beginnings'. In the 1980s he finally developed the 'normal accidents' theory (Perrow, 1984).

Major accidents in the production and service industries were not ascribable to individual employees or to their motives but could instead be traced back to two indicators of production systems: the degree of coupling and the complexity of the interaction (Figure 5). These features were responsible for the inevitability of major accidents as was reflected in the naming of the theory.

Coupling is a technical term which refers to the presence of a buffer or space between two elements in a system and to the degree of variability between process steps. In a tightly coupled system there is no buffer and the process steps have to follow a predetermined sequence. A fault in one system element or a process failure will propagate to all following elements affording limited options for correction and recovery. As with coupling, interaction also has two levels: linear and complex interactions. In engineering terms, complex interaction is reflected in what are known as common mode functions, where one system element will steer two or more subsequent system elements.

Common modes can spread the consequences of faults, or process disturbances, like an ink blot. Especially sectors in the top segment of Figure 5 are vulnerable, because of their tight coupling or their complex interactions

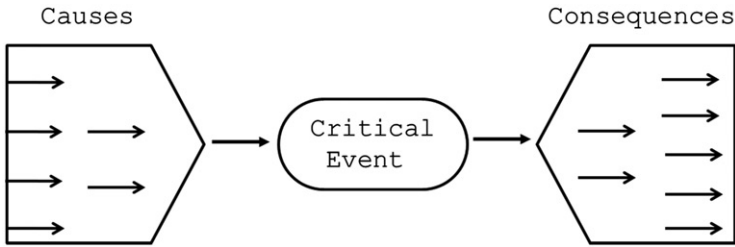


Figure 4. Conceptualization of the bow-tie for mathematical modelling (after Nielsen, 1971).

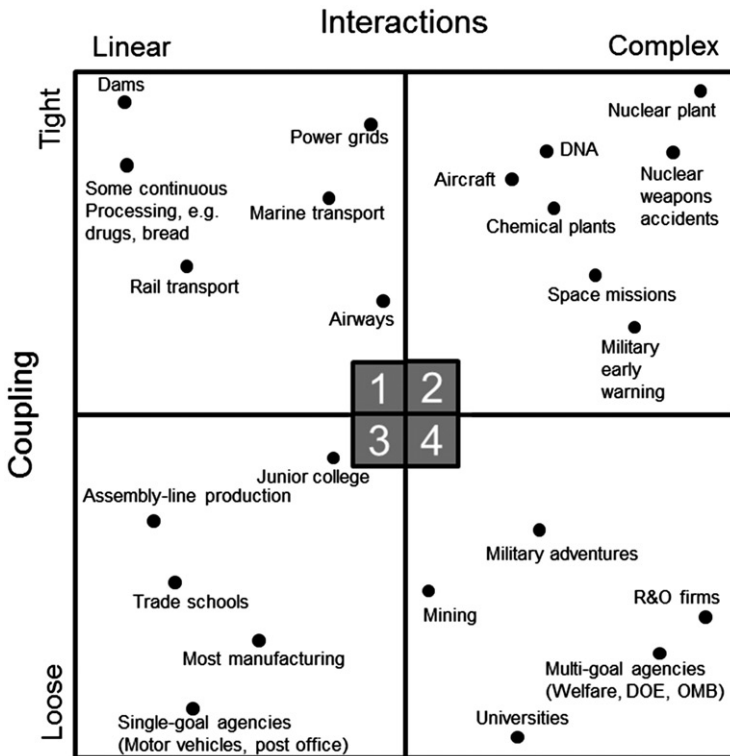


Figure 5. Normal Accidents Theory (After Perrow, 1984).

or for both reasons. This prediction was also confirmed by Le Coze's comparison of major disasters in the 1970s and the 1980s and during the first decade of the twenty-first century (Le Coze, 2013) (Table 1).

From the late 1980s and the 1990s onwards, another theory emerged from extensive research carried out at many different Shell locations around the world. It drew attention to the organisational, human and technical factors of accident processes. First these factors were labelled 'resident pathogens' (Reason, 1987), thereby building on Turner's concept of the

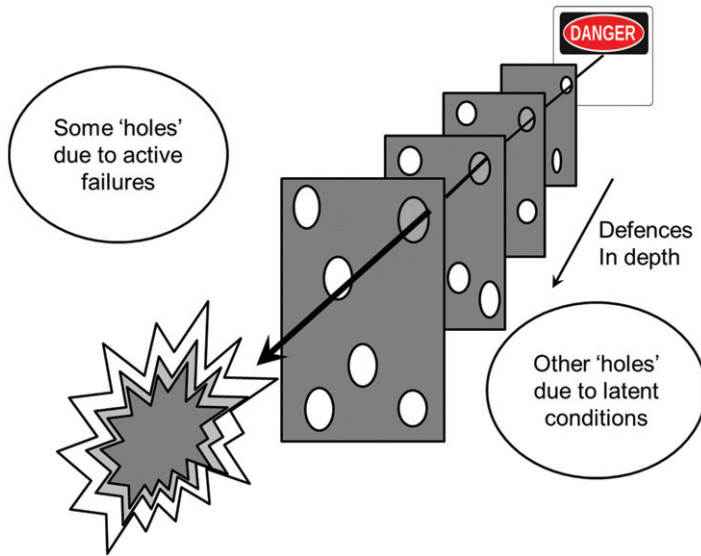


Figure 6. Swiss cheese (after Groeneweg, 1992).

incubation period of major accidents. As in a human body, technological systems also bear the seeds of their own destruction. Major accidents and disasters are rarely caused by just one single factor. At any one time, system errors, human error and other unforeseen conditions come together, while none of these factors alone could cause a disaster. In later publications resistant pathogens was replaced by the broader term, latent failures.

These latent failures, induced by decisions, were considered to mark the starting point of the accident process.

The psychology background of the researchers is apparent from the presence of Heinrich's psychological precursors and unsafe acts. (Wagenaar, Groeneweg, Hudson, & Reason, 1994). This finally led to the Tripod theory, which still is very popular in quite a few countries, and to the corresponding Swiss cheese metaphor (Groeneweg, 1992) (Figure 6).

Latent factors had by then been replaced by active failures and latent conditions. The new terms created some confusion but it was generally accepted that active failure was a consequences and not a cause of accidents. Like weak signals, latent conditions also almost served as a black box of accident causation. In Tripod these conditions were called '*basic risk factors*', and they were based on the results drawn from thousands of respondents (Table 2). In that sense, the investigation generated its own huge data information source.

The final two contributions to the safety science domain discussed in this chapter are the bowtie metaphor, published by Visser (1995) and the drift to danger model of Rasmussen (1997).

Table 1. Major accidents, a déjà vu (after Le Coze, 2013).

High-risk industries	Period	
	1970–1980s	2000–2010s
Nuclear	Chernobyl, 1986	Fukushima, 2011
Offshore drilling	Piper Alpha, 1988	Deepwater Horizon, 2010
Fuel storage	Port Edouard Heriot, 1987	Buncefield, 2005
Aerospace	Challenger, 1986	Colombia, 2003
Aviation	Tenerife, 1977	Rio Paris, 2009
Petrochemical	Flixborough, 1974, Bhopal, 1984	Toulouse, 2001, Texas City, 2005
Railway	Clapham Junction, 1988	Ladbroke Grove, 1999
Maritime I	Zeebrugge, 1987	Costa Concordia, 2012
Maritime II	Exxon Valdez, 1989	Erika, 2003
Air traffic management	Zagreb, 1976	Umberlingen, 2002

Table 2. Basic risk factors (after Groeneweg, 1992).

1	Design, poor design installation, equipment, tools
2	Hardware, deficiencies in quality of equipment, tools
3	Error enforcing conditions
4	Maintenance, inadequate management
5	Defences, absent, inadequate protection
6	Procedures, deficiencies in quality, workability
7	Housekeeping, poor housekeeping
8	Training, deficiencies in knowledge and skills
9	Incompatible goals, conflicting requirements
10	Communication, relevant information does not reach recipients
11	Organisation, deficiencies in structure

The bowtie metaphor represents the relationships between scenarios as shown by the arrows going from left to right for the barriers and, for the management factors, the vertical arrows (Figure 7).

The central event positioned in the centre of the metaphor depicts a state in which energy (hazard) has become uncontrollable. Managerial factors relate to the acquisition, maintenance and, more generally, to the quality of the barriers. The metaphor has a concealed time factor. It can be a long time, similar to the incubation period of Turner, before a hazard reaches the central event state. Once uncontrollable, scenarios will generally unroll very quickly in direction their ultimate consequences. In the model of Rasmussen (1997), the latent conditions have a different origin.

Rasmussen emphasizes the dynamics of the decision making of stakeholders pushing for faster, cheaper and more efficient production. The pace of change of technology is very fast, and is represented in many domains, like transport, shipping, energy, manufacturing and the process industry. This pace of change is much faster than the pace of change in management structures (Rasmussen, 1997). It is said that ‘a second generation of management is applied to a fifth generation of technology’. An even longer lag in response to change is found in legislation and regulation, where prescriptive legislation has been replaced by performance-based legislation.

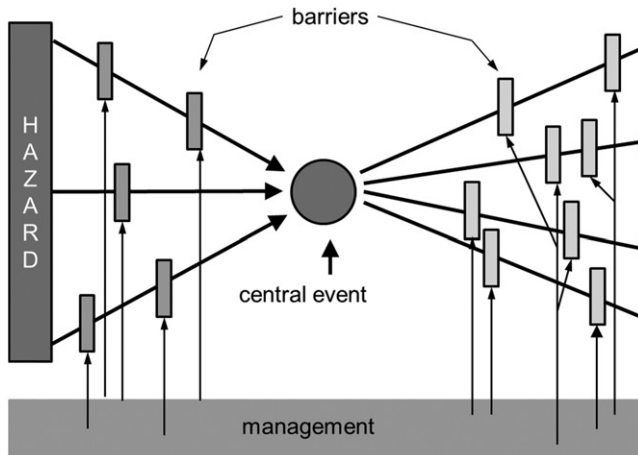


Figure 7. Bow-tie as a metaphor (after Visser, 1995).

This has prompted public concern, the worry being that it is too loose and not easily enforceable. Company documents are increasingly becoming the primary source for inspections. This places a heavy burden on the competence and knowledge of controllers and regulators. These lag time differences create a problem, notably in high-energy-high-risk industries, where pressure on cost-effectiveness dominates. It can land a system in a situation where it has strayed from its safety envelope. When the boundaries of the safety envelope are reached, the system drifts towards danger. This explains why investigations into serious accidents from the point of view of acts, events and errors are not very useful but should instead be directed towards research into decision making and to integrating the knowledge and the context of such decision. Risk management should be focussed on understanding the dynamics of the safety of processes and the need for stakeholders to determine the boundaries and gain insight through feedback control, into when a state of 'drift to danger' occurs (Svedung & Rasmussen, 2002).

Discussion and conclusions

This chapter gives a bird's eye view of the history of more than 150 years of safety science developments. The most notable developments are shown in Table 3. The overview stops in the late 1990s. A major development after that period was resilience engineering (Hollnagel et al., 2006). Only in the scientific community was resilience seen as being very similar to high reliability, a development that started 20 years earlier (Hale & Heijer, 2004).

Safety Science used to be a domain with many different research disciplines, ranging from political science, law and economics, to sociology,

Table 3. Significant developments in safety science till 1990.

1844	Safety by shielding (UK)
1906	Safety First movement (US)
1909	External causes theory (US)
1919	Accident prone theory (UK)
1926	Hazard = energy model (US)
1927	1:4 accident cost model (US)
1928	88:10:2 cause model (US)
1929	1:29:300 accident precursor model (US)
1935	External factors theory (UK)
1941	Domino metaphor (US)
1949	Epidemiology triangle (US)
1950	Safety management ladder (US)
1951	Task dynamics (NL)
1951	Psychological climate (US)
1957	Reliability engineering
1961	Safety barriers (US)
1963	Formal methods HAZOP, FT & FMEA
1964	Loss prevention (UK)
1966	Iceberg metaphor (US)
1967	Man-machine interface (UK)
1971	Organizational culture (UK)
1971	Safety audits (US)
1971	Disturbed information (UK)
1971	Bowtie math model (DK)
1973	MORT model (US)
1978	Incubation and weak signals (UK)
1980	Safety climate (ISR)
1981	Process disturbances (SE)
1981	Risk triplet (US)
1982	Skill-rule-knowledge (DK)
1984	Normal accidents (US)
1985	Inherent safe design (UK)
1986	Safety culture model (SSSR)
1987	Resident pathogens metaphor (UK)
1987	High-reliability-organization (us)
1992	Basic risk factors (NL)
1994	Impossible accidents (NL)
1997	Swiss cheese model (UK)
1997	Drift to danger (DK)
1998	Bowtie metaphor (NL)

Table 4: Rolt's scenarios (Rolt, 1955/1976).

1	Double line collisions
2	Blow-ups and breakdowns
3	Bridge failures – storm and tempest
4	Other men's responsibilities – permanent way faults and runaway locomotives
5	Single line collisions
6	High speed derailments
7	Stray wagons and breakaways
8	Signalmen's errors
9	Driver's errors
10	How much automation?

management and organisations, psychology, ergonomics and engineering. All those different disciplines only rarely developed common discussion podia, or shared research projects. It was only in the mid-1970s that the first academic safety science groups were formed at universities, starting in Germany, Belgium and the United Kingdom and, towards the end of the 1970s, followed by the Netherlands (Hale & Kroes, 1997). As an independent discipline Safety Science is rather young which explains the relative weakness of its theories. The theories themselves are solid enough to analyse major accidents; it is just that retrospective research has its known pitfalls, and there is always the bias of hindsight. The theories developed are not able to anticipate major accidents which still take us by surprise, not only in the process industry but also in rail transport, aviation, the nuclear sector and in other high-energy-high-risk industries. Apparently too many variables are involved.

In the past, the volume of safety research and surveys was vast and drew on huge data sets from different sources. Heinrich, Eastman, Winsemius, Turner and Groeneweg, to name but a few, were some of the exponents. Their big data might not fit the present definition of high velocity, but it certainly complies with high volume and great variety. These examples show that big data is not an entirely new concept in Safety Science.

The theories presented can provide a classification, a necessary structure and can help in the interpretation of the results derived from big data analysis. Such classification is crucial, because 'theory-free' correlations, in combination with big data analyses, will not give any insight into relations, any understanding of why correlations will change over time or any form of data bias, as mentioned by RSSB. Historically the rail sector has his own big data pioneer in the form of Rolt who collected and analysed 125 years of accident data gathered from the Railways Inspection Department, starting in 1840 (Rolt, 1955; Rolt 1976) (Table 4). His book listed general scenarios, presented as chapters. Some of these scenarios, relating for instance to signalmen's errors and stray wagons, have already been superseded by time but some still seem relevant today. According to DeBlois, this is a sign of bad management since big data could be directed towards exploring the conditions of such recurrent events.

Disclosure statement

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of this article.

Notes on contributor

Paul Swuste, is an associate professor of the Safety Science Group of the Delft University of Technology, The Netherlands. He has a MSc degree in

Biochemistry from the University in Leyden (1978) and finished his Ph.D. thesis 'Occupational Hazards and Solutions' in 1996. From 1980 onwards he is working at the Safety Science Group publishing frequently both nationally and internationally on results of research on occupational safety, hygiene, medicine, and process safety. He is member of various scientific committees.

References

- Aldrich, M. (1997) *Safety First: Technology, labour and business in the building of American safety 1870–1939*. Baltimore: John Hopkins University Press.
- Baker, J. (2007). The report the BP US refineries independent safety review panel. Washington: US chemical Safety Board.
- Bird, F., & Germain, G. (1966). *Damage control, a new horizon in accident prevention and cost improvement*. New York: American Management Association, The Comet Press.
- Calder, J. (1899). *Prevention of factory accidents. Being an account of manufacturing industry and accident and a practical guide to the law on safe-guarding, safe-working, and safe-construction of factory machinery, plant and premises*. London: Longmans, Green and Co.
- Cullen, W. (1990). The public inquiry into the Piper Alpha disaster. London: Department of Energy, Her Majesty's Stationary Office.
- DeBlois, L. (1926). *Industrial safety organisation for executives and safety engineer*. New York: McGraw-Hill Book Company.
- Eastman, C. (1910). *Work-accidents and the law, the Pittsburgh survey*. New York: Charities Publications Committee.
- Gibson, J. (1964). *The contribution of experimental psychology to the formulation of the problem of safety – A letter for basic research. Behavioural Approaches to Accident Research. Association for the Aid of Crippled Children, New York*. In Haddon, W., Suchman, E., & Klein, D. (Eds.), *Accident research, methods and approaches* (pp. 77–89). New York: Harper & Row.
- Gordon, J. (1949). The epidemiology of accidents. *American Journal of Public Health*, 39, 504–515. doi: 10.2105/AJPH.39.4.504.
- Greenwood, M., & Wood, H. (1919). The incidence of industrial accidents upon individuals with special reference to multiple accidents. Industrial Fatigue Board, report nr 4. London: Her Majesty's Stationary Office.
- Groeneweg, J. (1992). *Controlling the controllable, the management of safety (Doctoral thesis, Rijksuniversiteit Leiden)*. Leiden: DWSO Press.
- Guldenmund, F. (2010). *Understanding and exploring safety culture (Doctoral Thesis, Delft University of Technology, The Netherlands)*. Oisterwijk: BoxPress.
- Haddon, W. (1963). A note concerning accident theory and research with special reference to motor vehicle accidents. *Annals of the New York Academy of Science*, 107, 635–646. doi: 10.1111/j.1749-6632.1963.tb13307.x.
- Haddon, W. (1968). The changing approach to the epidemiology, prevention, and amelioration or trauma: the transition to approaches etiologically based rather than descriptive. *American Journal of Public Health*, 58, 1431–1438. doi: 10.2105/AJPH.58.8.1431.
- Hale, A. (1978). *The role of government inspectors of factories with particular reference to their training needs (Doctoral thesis, University of Aston, Birmingham)*.

- Hale, A., & Hale, M. (1970). Accidents in perspective. *Occupational Psychology*, 44, 115–122.
- Hale, A., & Hale, M. (1972). A review of the industrial accident research literature of the National Institute of Industrial Psychology. London: Her Majesty's Stationery Office.
- Hale, A., & Heijer, T. (2004). Defining and assessing resilience in risk control and safety management systems: Some thoughts from transport studies and chemical plants. Copenhagen, Denmark: First International Reliability Engineering Technology Workshop.
- Hale, A., & Kroes, J. (1997). 10 years of the chair of safety science. *Safety Science*, 26, 3–19. doi: 10.1016/S0925-7535(97)00024-6.
- Heinrich, H. (1927). The incidental cost of accidents. *National Safety News*, 17, 18–20.
- Heinrich, H. (1928). The origin of accidents. The apparent cause is not always the basic cause. *National Safety News*, 18, 9–13.
- Heinrich, H. (1929). The foundation of a major injury. *National Safety News*, 19, 9–11.
- Heinrich, H. (1941). *Industrial accident prevention, a scientific approach* (2nd ed.). London: McGraw-Hill Book Company.
- Heinrich, H. (1950). *Industrial accident prevention, a scientific approach* (3rd ed.). New York: McGraw Hill Book Company.
- Hoffman, F. (1909). Industrial accidents and industrial diseases. *Publications of the American Statistical Association*, 11, 567–603. doi: 10.2307/2964905.
- Hollnagel, E., Woods, D., & Leveson, N. (2006). *Resilience engineering: Concepts and precepts*. Ashgate: Aldershot, UK.
- INSAG. (1986). International Nuclear Safety Advisory Group, Report 75-INSAG-1 on the post-accident review meeting on the Chernobyl accident, Safety Series, Vienna.
- Johnson, W. (1970). *New Approaches to safety in industry*. London: Industrial and Commercial Techniques LTD.
- Johnson, W. (1973). The management oversight and risk tree – MORT, including systems developed by the Idaho Operations Office and Aerojet Nuclear, report – SAN 821-2/UC-41 of the US Atomic Energy Commission. Germantown: Division of Operational Safety.
- Kaplan, S., & Garrick, J. (1981). On the quantitative definition of risk. *Risk Analysis*, 1, 11–27. doi: 10.1111/j.1539-6924.1981.tb01350.x.
- Keeman, V., Kerr, W., & Sherman, W. (1951). Psychological climate and accidents in automotive industry. *Journal of Applied Psychology*, 31, 108–111. doi: 10.1037/h0053560.
- Kellogg, P. (1909). *The Pittsburgh Survey*. New York: Charities Publications Committee.
- Kjellén, U. (1984). The role of deviations in accident causation. *Journal of Occupational Accidents*, 6, 117–126.
- Le Coze, J. (2013). New models for new times. An anti-dualist move. *Safety Science*, 59, 200–218. doi: 10.1016/j.ssci.2013.05.010.
- Le Poole, S. (1865). Ongelukken in de fabriek (Accidents in the factory). *De Economist*, 14, 449–457. doi: 10.1007/BF02202382.
- Moore, P. (1949). *Basic Operational Research*. New York: Pitman Publishing.
- Nielsen, D. (1971). The cause/consequence diagram method as a basis for quantitative accident analysis. Report Risø-M-1374 of the Danish Atomic Energy Commission, Risø.

- Palmer, L. (1926). The history of the Safety Movement. *Annals of the American Academy of Political and Social Sciences*, 123, 9–19. doi: 10.1177/000271622612300104.
- Perrow, C. (1984). *Normal accidents. living with high-risk technologies*. New York: BasicBooks.
- Petersen, D. (1971). *Techniques of safety management*. New York: McGraw-Hill Book Company.
- Rasmussen, N. (1975). Reactor safety study, an assessment of accident risks in the US commercial nuclear power plants, report WASH-1400, NUREG-75/014 by the Nuclear regulatory Commission, Rockville.
- Rasmussen, J. (1982). Human errors, a taxonomy for describing human malfunctioning. *Journal of Occupational Accidents*, 4, 311–333. doi: 10.1016/0376-6349(82)90041-4.
- Rasmussen, J. (1987). Reasons, causes, and human error. In: J. Rasmussen, K. Duncan, & J. Leplat (Eds.), *New technology and human error*. (pp. 293–301) Chichester: Wiley.
- Rasmussen, J. (1997). Risk management in a dynamic society: a modelling problem. *Safety Science*, 27, 183–213. doi: 10.1016/S0925-7535(97)00052-0.
- Reason, J. (1987). The Chernobyl errors. *Bulletin of the British Psychological Society*, 40, 201–206.
- Reeves, T., & Turner, A. (1972). Theory of organisation and behaviour in batch production factories. *Administrative Science Quarterly*, 17, 81–98. doi: 10.2307/2392097.
- Robens, A. (1972). *Committee on safety and health at work, Report of the Committee 1970–1972*. London: Her Majesty's Stationery Office.
- Roberts, K. (1988). Some characteristics of one type of high reliability organization. *Organization Science*, 1, 160–176. doi: 10.1287/orsc.1.2.160.
- Rochlin, G., La Porte, T., & Roberts, K. (1987). The self-designing high reliability organisation: Aircraft carrier flight operation at sea. *Naval War College Review*, 40, 76–90.
- Rolt, L. (1955). *Red for danger, a history of railway accidents and railway safety precautions*. London: Sutton Publishing Limited.
- Rolt, L. (1976). *Red for danger*. London: Pan Books.
- RSSB. (2014). *Big data analytics*. Report (S183) prepared for CIO forum, Rail Safety and Standard Board, London.
- Saleh, J., & Marais, K. (2006). Highlights from the early (and pre-)history of reliability engineering. *Reliability Engineering and System Safety*, 91, 249–256. doi: 10.1016/j.res.2005.01.003.
- Singleton, W. (1984). Future trends in accident research in European countries. *Journal of Occupational Accidents*, 6, 3–12. doi: 10.1016/0376-6349(84)90024-5.
- Swuste, P., van Gulijk, C., & Zwaard, W. (2010). Safety metaphors and theories a review of the occupational safety literature of the US UK and the Netherlands, till the first part of the 20th century. *Safety Science*, 48, 1000–1018. doi: 10.1016/j.ssci.2010.01.020.
- Swuste, P., Van Gulijk, C., Zwaard, W., & Oostendorp, Y. (2014). Occupational safety theories, models and metaphors in the three decades since WO II, in the United States, Britain and the Netherlands: a literature review. *Safety Science*, 62, 16–27. doi: 10.1016/j.ssci.2013.07.015.
- Swuste, P., Van Gulijk, C., Zwaard, W., Lemkowitz, S., Oostendorp, Y., & Groeneweg, J. (2015). Developments in the safety domain and in safety management from the 1970s till the 1979 near disaster at Three Mile Island. *Safety Science*, 86, 10–26. doi: 10.1016/j.ssci.2016.01.022.

- Svedung, I., & Rasmussen, J. (2002). Graphic presentation of accident scenarios mapping system structure. *Safety Science*, *44*, 397–417. doi: 10.1016/S0925-7535(00)00036-9.
- Turner, B. (1971). *Exploring the industrial subculture*. London: The MacMillan Press LTD.
- Turner, B. (1978). *Man-made disasters*. Oxford: Butterworth-Heinemann.
- Vernon, H. (1936). *Accidents and their prevention*. Cambridge: University Press.
- Visser, J. (1995). Managing safety in the oil industry. The way ahead. Proceedings of Loss Prevention and Safety Promotion in the Process Industries, June 6–9, Antwerp.
- Wagenaar, W., Groeneweg, J., Hudson, P., & Reason, J. (1994). Promoting safety in the oil industry. *Ergonomics*, *37*, 1999–2013. doi: 10.1080/00140139408964963.
- Weick, K. (1987). Organisational culture as a source of high reliability. *California Management Review*, *29*, 112–127. doi: 10.2307/41165243.
- Winsemius, W. (1951) De psychologie van het ongevalsgebeuren. Verhandeling van het Instituut voor Praeventieve Geneeskunde, Doctoral thesis (The psychology of the accidents. Discourses of the Institute of Preventive Medicine), Kroese, Leiden.
- Zohar, D. (1980). Safety climate in industrial organizations: theoretical and applied implications. *Journal of Applied Psychology*, *65*, 96–102. doi: 10.1037//0021-9010.65.1.96.