PROMETHEUS, 2016
VOL. 34, NO. 2, 115–132
https://doi.org/10.1080/08109028.2017.1279873

Routledge
Taylor & Francis Group

RESEARCH PAPER

Safety in unpredictable complex systems – a framework for the analysis of safety derived from the nuclear power industry

Craig S. Webster

Centre for Medical and Health Sciences Education and Department of Anaesthesiology, School of Medicine, University of Auckland, Auckland, New Zealand

ABSTRACT

The increasing complexity and power of our technologies compels us to find new ways in which to conceptualise, understand and maintain their safety in the long term. Some complex technological industries have performed better than others in terms of applying sustained and systematic approaches to the maintenance of safety. The United States nuclear power industry can be seen as an ideal test-bed for the development of safety initiatives, being responsible for the control of potentially unpredictable technology that involves extraordinary forces and costs. This paper describes and formalises a framework for better understanding the safety of complex socio-technological systems, based on key events in the development of safety in the United States nuclear power industry. The framework comprises two components: (1) a state-space approach for better conceptualising system failures, the benefits of incident reporting and remedial safety initiatives; and (2) a set of milestones that can be used to assess the development of safety in socio-technological industries. Healthcare and the United States nuclear power industry both represent complex socio-technological systems with similar technical characteristics. However, safety strategies in healthcare have not kept pace with the increasing complexity of clinical practice, and there have been international calls for improvements in patient safety. The framework is applied to the analysis of safety in healthcare, demonstrating its utility as an alternative safety analogy in healthcare. Use of the framework indicates substantial scope for improvements in healthcare safety through major evidence-based system redesign. By lowering the threshold for the reporting of incident data to include accident precursors, it is possible to identify problem areas before patient harm occurs.

Introduction

As our technologies become increasingly complex and powerful, we need to consider the larger picture of how we maintain and monitor their safety in the long term (Plsek and Greenhalgh, 2001; Barach and Johnson, 2006; Leveson, 2011; Webster, 2012). The performance of high technology industries has been variable in terms of the application of

CONTACT Craig S. Webster c.webster@auckland.ac.nz
Accepting editor: Joanne Roberts

systematic approaches to the maintenance of safety. For example, safety remains of pressing concern in healthcare, largely because of the late adoption of systematic approaches to safety and the fact that advances in healthcare technology are rapidly outstripping traditional methods of maintaining safety (Department of Health, 2000; Institute of Medicine, 2000; Webster and Grieve, 2005; Healy, 2011; Vincent and Amalberti, 2016). By contrast, in terms of scale, cost and risk, the United States nuclear power industry provides an example of a relatively very safe, large-scale, socio-technological system in which safety has been a central consideration from its conception. Unlike in healthcare, systematic approaches to safety have been adopted and refined from the first years of the existence of the nuclear power industry, and over the course of its history (Clarfield and Wiecek, 1984; Perrow, 1984; Morone and Woodhouse, 1986; Sagan, 1993; Casey, 1998; Perin, 2005). I have previously used a novel framework, based on the US nuclear power industry, to analyse a failure in the delivery of safe anaesthesia for surgery in the operating room (Webster, 2005). This framework comprises a state–space approach for analysing accident sequences and a set of milestones for judging the development of safety in a socio-technological system.

One of the most prominent safety analogies in healthcare is aviation safety, which recently resulted in the development of the celebrated World Health Organisation surgical safety checklist – a checklist tool for use in the operating room to ensure that key steps in safety procedures are not omitted (Gawande, 2009). In a 2009 multinational study, use of the checklist was shown to reduce significantly postoperative adverse events in surgical patients (Haynes *et al.*, 2009). This checklist was, in part, inspired by the response of the United States aviation industry to the crash of a Boeing model 299 aircraft in 1935 (Gawande, 2009; Webster *et al.*, 2015). However, human beings are not aircraft, and the validity of such safety analogies tends to be taken at face value without formal consideration (Webster, 2002).

Therefore, in the following I aim to formalise a safety-analysis framework based on key events in the more recent history of the development of safety in the US nuclear power industry. The framework will allow the better conceptualisation of safety and the analysis of accidents in healthcare, and give a new perspective on many currently piecemeal safety initiatives in healthcare. Such a formalised framework will support the safety analogy of the nuclear power industry in healthcare as an alternative to the aviation analogy, with the potential to guide future safety initiatives in healthcare (Vincent and Amalberti, 2016).

Large-scale technological systems

A complex system is one with a sufficiently large number of possible component interactions such that accurately predicting the long-term safety and behaviour of the system from knowledge of its constituent parts is difficult or impossible. Socio-technological systems contain human operators or workers as vital components in the system's everyday function. Thus, the nuclear power and healthcare industries are both complex socio-technological systems. Yet healthcare is one of the last industries of this type to adopt a systematic approach to safety (Perrow, 1984; Morone and Woodhouse, 1986; Sagan, 1993; Barach and Small, 2000; Helmreich and Merritt, 2001; Chiles, 2001; Webster, 2005; Cook and Rasmussen, 2005; Pronovost and Hudson, 2012).

A landmark contribution to the understanding of complex technological systems remains Charles Perrow's *Normal Accidents Theory*, which has its origins in the analysis of the Three Mile Island Number 2 (TMI-2) nuclear power plant accident in 1979 (Perrow, 1984).

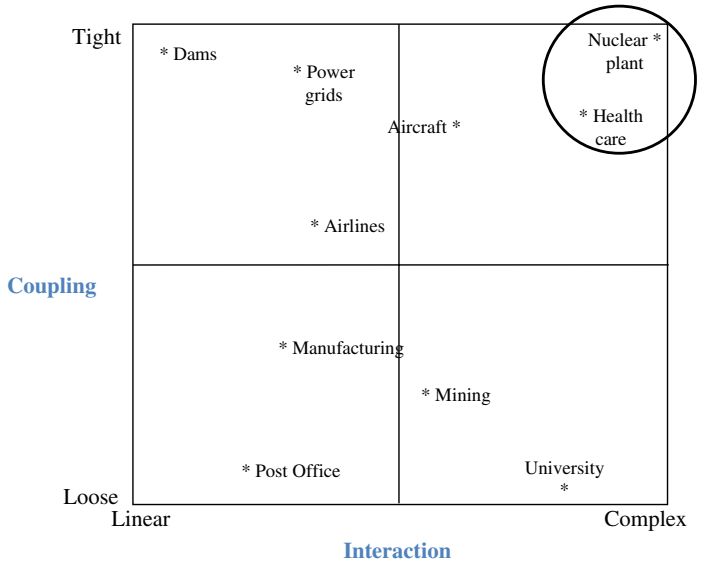

Figure 1. Interaction vs coupling space.
Note: The interaction/coupling space [adapted from Perrow (1984)], showing the placement of nuclear power plants and healthcare in the potentially most dangerous top right-hand quadrant, with various other industries and technologies for comparison.

Accidents in complex systems can be considered normal in the sense that they persist despite many safeguards, and do so through the unanticipated interaction of multiple failures. The complexity of the system both predisposes it to simultaneous multiple failures and masks the many potential ways in which such individual failures may interact in dangerous ways (Perrow, 1984). Perrow also describes how the characteristics of any system can be classified along two dimensions – interaction and coupling. On the interaction dimension, a task or process can be said to have *complex* interaction among parts if there are many alternative sub-tasks at any point in its completion, or *linear* if it consists of a set of fixed steps carried out in a rigid sequence. The coupling dimension describes the extent to which an action in the task or process is related to its consequences. A system is *tightly* coupled if consequences occur immediately after an action – hence tightly coupled systems result in more accidents because minor mistakes can become serious accidents before they can be corrected. A *loosely* coupled system is more forgiving of error or mistakes as it allows greater opportunity for an error to be corrected in time to avoid serious consequences (Weick, 1976). These two dimensions form Perrow's interaction/coupling space with which human activities can be classified (Perrow, 1984).

Nuclear power plants are structurally and functionally complex (Helmreich and Merritt, 2001; Webster, 2002) and are rated as the most highly complex and tightly coupled technology in Perrow's interaction/coupling space. Therefore, they fall into Perrow's potentially most dangerous top right-hand quadrant (Perrow, 1984) (Figure 1). Healthcare is a mixture of loosely and tightly coupled processes, but as a technological endeavour, it also falls into the same potentially most dangerous top right-hand quadrant of the interaction/coupling space. Technologies in this quadrant require sophisticated and systematic approaches to the maintenance of safety. The commercial aviation industry has often been used as an example of a safe system from which healthcare can learn, but it is worth noting that in terms of the

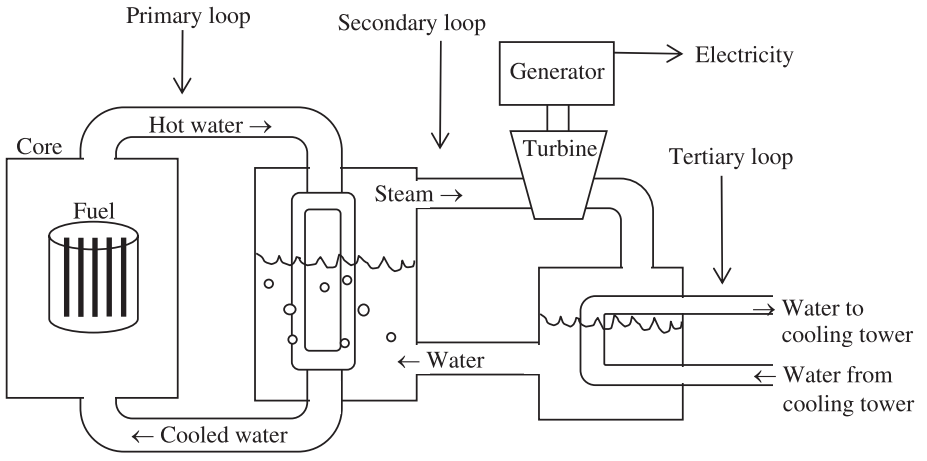

Figure 2. Principal components of the most common type of nuclear reactor.
Note: A conceptual diagram of the principal components of the most common type of nuclear reactor used for electricity generation (the pressurised water reactor or PWR – not to scale). The core contains uranium fuel and the primary loop contains high-pressure radioactive water. Both are isolated inside a special containment building.

important system characteristics of interaction and coupling, healthcare is more closely co-located with nuclear power in the Perrow space than with airlines or aircraft, suggesting that the nuclear power industry is a better comparator for safety analogies (Hunt, 1988; Helmreich and Merritt, 2001; Webster, 2002; Kapur *et al.*, 2016).

The system characteristics of complexity, tight coupling and the potential for large-scale disaster make nuclear power plants a kind of worst-case test-bed for the development of safety strategies. The US nuclear power industry has undergone significant cultural change and system redesign in the pursuit of safety – spending vastly more time and money on safety than many other high technology industries, including healthcare. The nuclear power industry has also made early use of the reporting of incidents and accidents in order to avoid such events in the future, and the use of training in simulators to improve the response to such events when they do occur (Schlager, 1994). A wealth of documentation and systems analyses exist for many cases of significant system failure in the nuclear power industry. In other industries, episodes of systems failure are relatively under-documented, and what documentation there is tends to be blame-centred, thus directing attention away from the underlying factors which persistently predispose human error and system failure. Therefore, the development of the nuclear power industry is instructive in understanding the nature of safety in complex, socio-technological systems in general, and in identifying safety lessons for other industries with relatively under-developed safety cultures.

Nuclear power in the 1950s – the age of the friendly atom

After the Second World War, Americans were ready to embrace the peaceful use of nuclear power and to enter the new atomic age. Conceptually at least, the first nuclear reactors designed for electricity generation appeared to be relatively straightforward devices. Light water reactors (LWR) were the first class of reactors to be built for commercial electricity generation in the US, so called because they used ordinary water for coolant. An LWR

uses uranium as a heat source to generate steam, which in turn drives turbines to generate electricity. Hence, to many engineers in the 1950s, reactor plans did not look too different from familiar steam boiler technology. The attraction of uranium is that it yields a great deal of heat from a small amount of fuel – and so, in this sense, is much more efficient than other fuels, such as coal or oil. The most common type of LWR to be built, and which remains the most common in service today, is the pressurised water reactor (PWR). These comprise three independent loops of piping, all filled with water or steam. The first loop runs directly through the reactor core, where the uranium fuel heats the water under high pressure (also making it radioactive). This first loop then provides the heat needed to generate steam in the second closed loop of piping via a heat exchange device. The steam produced drives turbines that in turn generate electricity. The third loop of piping uses another heat exchange to remove excess heat from the system, via a river, ocean or cooling tower (Figure 2). As the first and second loops of piping are closed, radiation from the core should never leave the main reactor chamber (which is well shielded from the rest of the reactor facility and the surrounding environment). The amount of heat produced by the core depends on the amount of nuclear fission taking place. Inserting control rods into the core absorbs neutrons and slows the rate of fission and heat production. Removing control rods speeds up fission and heat production. The rate of fission in a nuclear reactor must always remain well controlled – too much heat from an uncontrolled fission reaction can lead to steam explosions, the melting of the reactor core itself, and the release of deadly radiation. Even with all control rods inserted in a PWR, the core remains very hot, and so must always be filled with coolant to prevent it from overheating and potentially melting, a so-called 'nuclear meltdown'.

The development of safety strategies in nuclear power generation

The potential hazards of nuclear reactors made the industry very different from others, such as the chemical industry, in that trial-and-error approaches to safety regulation could not be used in the commercial sector and only rarely in the military (Anonymous, nda; Morone and Woodhouse, 1986; McKeown, 2003). Wide safety margins and apparently fail-safe approaches had to be employed from the beginning. However, a paucity of experimental data made it difficult to define just what was safe. The earliest approaches to safety regulation, therefore, fell back on the small amount of experience and few data that were available, much of which was related to the safety practices established in the Manhattan Project.[1] Thus, the first experimental reactors designed for electricity generation were isolated from populated areas by two concentric clear zones (McKeown, 2003). The first was completely unpopulated and under direct control of the Atomic Energy Commission (AEC), and the second was populated by no more than 10,000 people.

In terms of fatalities, the most serious nuclear power plant accident in the United States occurred in an early experimental military reactor called the Stationary Low-Power Reactor Number One (SL-1) (Anonymous, ndb). This plant was constructed in the late 1950s in the remote area of Idaho Falls, thus employing the isolation safety strategy (Anonymous ndb; McKeown, 2003). SL-1 was a type of LWR called a 'boiling water reactor' because the primary loop was not pressurised, allowing the water in it to boil during power production. On 3 January 1961, during a manual phase of a maintenance operation, control rods were withdrawn too far from the reactor core. This immediately led to a huge spike in heat

production within the core, causing a steam explosion which breached the reactor vessel, sprayed radioactive coolant throughout the reactor chamber, and expelled control rods from the top of the reactor at high speed. All three operators present at the time of the accident were killed, including one man who was impaled by a control rod and pinned to the ceiling of the reactor chamber. Radiation levels in the men's bodies were so high that some body parts had to be disposed of in a nuclear waste site rather than in a cemetery. The remainder of their bodies were buried in lead lined coffins. Despite the deadly nature of the accident, the safety strategy of building the plant within an isolation zone worked in that no one in the surrounding areas was affected by radiation.

An additional safety mechanism devised in the late 1940s was the reactor containment vessel. This consisted of a huge gas-tight sphere constructed over the entire reactor facility (or at least the primary loop) and designed to contain radioactive material in the event of a reactor explosion. At about the same time, the US Navy adapted the first experimental, land-based reactors for use in submarines. Such safety strategies as reactor isolation and bulky containment vessels were impossible inside a cramped submarine, and so new preventative safety strategies were developed. These comprised building reactors to withstand the 'worst credible' operating circumstances, rather than the average or expected circumstances. Components were designed to withstand higher pressures and temperatures than were ever likely to be encountered and backup mechanisms were built into the control and shutdown systems, thus employing the safety principle of redundancy. So, if one control or shutdown system failed, another independent system could be operated.

Nuclear reactors close to home

With the establishment of civilian nuclear power generation, a desire to build reactors closer to the population centres they served soon emerged. Large population centres consume the most electric power, and transporting power over long distances is inefficient. Partly to achieve this, by the late 1950s the AEC required that all new nuclear reactors employ both containment safety strategies and the preventative strategies developed in the Navy's submarine reactors. The specifications of reactors were therefore upgraded conservatively to allow wider safety margins in terms of expected pressures and temperatures. More redundancy in backup systems was added, including multiple, independent cooling systems and shutdown systems. Engineers attempted to anticipate malfunctions and sequences of malfunctions, and designed automatic emergency systems that would trigger and attempt to stabilise the reactor when dangerous states occurred. All safety mechanisms were designed to contain or prevent the release of radioactive material in the case of the 'maximum credible accident' (even if such an accident destroyed the reactor core).

New complexity and unanswerable questions

From the early 1960s, plans for increasingly large nuclear reactors were before the AEC for approval. By 1966, the US had 15 nuclear power plants running, nine more under construction and 22 on order (Perrow, 1984; Chiles, 2001). The AEC was concerned that containment and prevention strategies were no longer sufficient to deal with the greater temperatures and pressures associated with the new, larger reactor cores, many of which were six times larger than anything previously built. In addition, attempts to strengthen containment and

prevention mechanisms required the anticipation of possible failures and malfunctions in even greater detail than before. The precise conditions involved in many 'credible accidents' were simply unknown. In addition, multi-system failures began to be considered seriously for the first time. What if supposedly independent reactor systems were not independent after all and both failed simultaneously? What if such an event occurred when the reactor core was in a vulnerable state? What if existing safety systems had serious failure modes that had not been considered or encountered? How could reactor designers anticipate the unanticipated? Questions such as these began to be seen as endless and unanswerable. It was becoming clear that knowledge of the reactor's parts was not sufficient to predict its emergent behaviour. The complexity of reactor design was now such that it was unclear whether many 'unlikely' circumstances were actually potential 'credible accidents'.

Construction problems and everyday behaviours in safety-critical systems

Even once a reactor plan had been approved, the larger, more complex reactors were more difficult to build to specification than had been anticipated. A 1979 report identified 35 nuclear plants in operation with 'significant differences' between the design specifications and the way they were built (Perrow, 1984). Construction problems occur in every field of industry, and so in this respect they are nothing new. The problem arises when everyday flaws occur in extraordinary safety-critical systems (such as nuclear power plants), because here they have a significant potential to contribute to catastrophic failure. The same is true of everyday behaviours. For example, in 1978 a worker in the Rancho Secco 1 reactor in California was changing a light bulb in a control panel when he accidentally dropped the bulb. This caused a short circuit inside the panel which triggered a shutdown of the reactor. However, because some sensors were lost because of the short circuit, controllers could not fully assess the state of the reactor. This led to rapid cooling of the core, an event that carries a risk of the core vessel cracking with subsequent release of radioactive coolant (Perrow, 1984).

Apparently trivial events leading to highly nontrivial consequences are direct results of a system that is both tightly coupled and has complex interaction among its parts. In this context, let us now consider the most significant, studied and costly reactor failure in US history, that at Three Mile Island, in order to demonstrate the ability of the state–space approach to conceptualise system failures and their remediation (Perrow, 1984; Chiles, 2001).

The Three Mile Island accident[2]

The Three Mile Island Number 2 nuclear reactor (TMI-2) cost more than US$700 million to build and began operation on 30 December 1978. Just three months later, on 28 March 1979, a partial melting of the reactor core reduced the plant to scrap. The clean-up was a dangerous and complex process that took 11 years, removed 150 metric tons of radioactive debris from the containment building, and cost a further US$973 million (Schlager, 1994).

The accident began just after 4am while TMI-2 was operating at full capacity, generating 7 million horsepower, enough to supply electricity to a city of 400,000 people (Perrow, 1984; Schlager, 1994; Chiles, 2001). To avoid electricity loss, independent systems in nuclear plants allow routine maintenance to be carried out while reactors are running. Before the accident, a maintenance crew had isolated a set of pipes in the secondary loop and had opened them for cleaning (see Figure 2). During this process, water seeped into part of the pneumatic control

system that opens and closes valves throughout the plant. This small variation from normal conditions triggered an automatic safety system that immediately shut down power generation and the entire secondary loop, cutting off cooling water to the primary loop. The reactor core and primary loop then began to overheat, causing a safety system automatically to push all control rods into the core to shut down fission and heat production. However, residual heat produced by the core, even once shut down, remained sufficient to generate electricity for 18,000 homes (Perrow, 1984). In the absence of sufficient cooling, the core continued to overheat. To relieve the building pressure in the reactor core, a pilot-operated relief valve (PORV) then automatically activated, venting very hot radioactive water into a storage tank inside the reactor containment building. After reducing pressure, the same automatic system told the PORV to close. At this point, another emergency cooling system was triggered, injecting high-pressure coolant directly into the reactor core at a rate of 1000 gallons a minute.

Operators lost in the system – the scramble for control

At this point, only about two minutes had passed since the beginning of the accident sequence and operators were scrambling to understand and control the rapidly-evolving crisis. A two-tone warning horn continued to sound in the control room, and more than 100 alarm lights had lit up on control panels (Schlager, 1994). One vital piece of information the operators did not have was that the PORV had not closed as it should have within the first two minutes. A light on the control panel indicated that it had closed, but – as was discovered later – this meant only that the computer had *sent the command* to close the PORV. In reality the PORV remained stuck open and was draining much-needed coolant out of the overheated reactor core at a rate of 220 gallons a minute.

Another instrument incorrectly indicated that the water level in the reactor core was reaching dangerously high levels (the operators assumed this was because of the high-pressure injection of coolant). This also fitted with the idea that the PORV was closed. If core pressure gets too high, the primary loop can burst, releasing tens of thousands of gallons of radioactive coolant and allowing the level of coolant in the core to sink so low that the uranium fuel becomes uncovered. Such a loss-of-coolant accident is the most feared in any nuclear reactor: a full-scale meltdown is likely to follow in which tons of molten uranium can melt through the bottom of the reactor vessel. When the molten uranium hits the spilt coolant on the floor of the containment building, a huge radioactive steam explosion can occur, which could blow the top off the containment building and spread radioactive material for miles. The operators, fearing the water level in the core was too high, followed standard procedure and cut back on the high-pressure injection of coolant. Operators believed they were avoiding a possible loss-of-coolant accident by reducing pressure in the core. In fact, they were already part way to a different kind of loss-of-coolant accident, as coolant continued to be lost through the open PORV. Cutting back on the high-pressure injection of coolant actually exacerbated this problem. As a result, the uranium fuel became uncovered, reached a temperature of over 5000 degrees Fahrenheit, and began to melt.

'Fresh eyes' enter the room

At just after 6am, reactor operator Brian Mehler was called into the plant for his shift an hour early to assist with the crisis. He found 50 operators, engineers and supervisors crowded

into the control room and trying to make sense of the control panels. After about half an hour, Mehler tried manually closing a valve to the PORV, suspecting it might still be open. Minutes later the reactor began to behave in a less mysterious manner. Over the next 11 hours, coolant levels were restored and the core was cooled. However, by this time, about half the core's fuel had melted. Mehler claimed only that he had 'brought a fresh pair of eyes into the room'. Subsequent analysis suggested that the core may have been as little as half an hour away from complete meltdown (Perrow, 1984; Schlager, 1994). The surrounding area had not been evacuated, as authorities did not want to cause 'unnecessary panic'. Had events led to failure of the containment building, it is likely that substantial loss of life would have followed.

Implications from disaster

It is beyond argument that TMI-2 was a financial disaster. The electric power the reactor generated in its short operating life earned only a tiny fraction of the construction and clean-up costs of US$1.7 billion. Perhaps surprisingly, however, some commentators claimed the events at TMI-2 demonstrated that nuclear reactors were well designed and that safety systems worked (Chiles, 2001). They claimed that even in the extreme circumstances of the TMI-2 failure, and despite the mess inside the reactor containment building, a complete meltdown or core breach had not occurred. The public begged to differ. Many people felt that 30 minutes and a thick concrete dome comprised too narrow a safety margin with which to be separated from a nuclear meltdown. It was easy to imagine more extreme circumstances in which a large-scale disaster would have resulted, with substantial loss of life. The TMI-2 accident was a major factor in a dramatic retreat from the building of new nuclear power plants by US utilities (Schlager, 1994). The age of the friendly atom had come to a dramatic end.

Although these events have cemented a negative view of the nuclear power industry for much of the public, it is worth considering that the number of deaths stemming from accidents in the US nuclear power industry is vanishingly small compared with the estimated 44,000–98,000 preventable deaths which occur every year in patients undergoing healthcare in the US (Institute of Medicine, 2000). Even estimates of the total death toll from the world's worst nuclear power plant accident (at Chernobyl in 1986) suggest that this disaster has led to fewer deaths than occur annually in those undergoing healthcare in the US (Anonymous, ndc).[3]

Fatigue and human performance

TMI-2 is one of a number of notorious industrial accidents to have begun in the early hours of the morning during a nadir in the circadian cycle of human performance (Reason, 1990; Gander *et al.*, 2008; Webster *et al.*, 2015). It seems likely that fatigue and compromised cognitive abilities contributed to the initial events that started the TMI-2 accident sequence and to the subsequent inability of night shift personnel to diagnose correctly the state of the reactor. In particular, fatigue may have made night shift personnel more likely to suffer from confirmation bias in their interpretation of control panels. This group of experts believed the PORV was closed, despite some control panels indicating the opposite. Therefore, the experts believed an incorrect instrument reading indicating the core pressure was dangerously high,

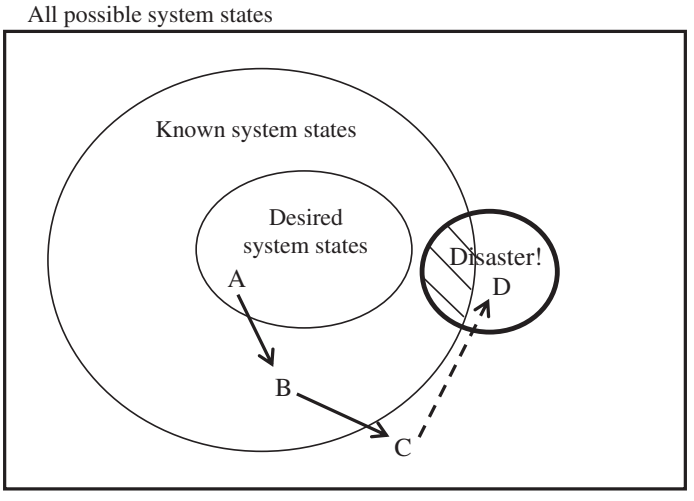

Figure 3. State-space diagram for an accident in a generic complex system.
Note: Known system states are a sub-set of the larger set of all possible system states. Some known states or 'credible accidents' lead to disaster, and only these can be specifically guarded against with safety sub-systems and procedures (the hatched area above). However, a probably larger set of unknown system states also leads to disaster, but these cannot be guarded against specifically because the causal mechanisms involved are unknown. (Figure reproduced with permission from John Wiley and Sons.)

as this was consistent with their existing incorrect diagnosis. Their actions (which actually made matters worse) were the standard procedures for the circumstances they perceived.

Brian Mehler was less likely to be suffering from confirmation bias when he brought a fresh pair of eyes into the room. He was not suffering from fatigue, had not been present from the start of the accident sequence, and so had no fixed diagnosis of the reactor in mind. Although he was aware of standard procedures, he also knew that these had not worked. Not only was the reactor in an 'off-normal' state (that is, a state different to any routine operational state), but it was also so far off-normal that it was behaving strangely in ways which did not match state sequences in any standard procedures. In these circumstances, it was unclear which standard procedures, if any, were appropriate to move the reactor back to a stable state. Following the reactor's operating rules had failed because there was no known rule for this set of circumstances. Mehler, therefore, resorted to reasoning from first principles to resolve the situation, something considerably more effortful than following an existing rule, and something of which the night shift operators, because of fatigue and fixation, would have been much less capable (Merry and McCall Smith, 2001; Gander *et al.*, 2008). The TMI-2 accident made it clear that even the experts' understanding of plant behaviour was far from complete. Despite exhaustive attempts to anticipate failure modes and credible accidents, many important emergent properties and system interactions remained hidden until they precipitated unexpected 'off-normal' reactor behaviour.

A state–space approach to complex systems

System failure in any complex system can be understood in terms of a state-space approach. Such an approach can also be used to understand and illustrate the benefits of incident reporting and better system design. In any complex system, the set of all possible system

states is very large and much larger than the sub-set of known system states (Figure 3). Desired states (e.g. where a reactor is safely generating electricity) are a sub-set within the set of known system states. Some *known* system states lead to disaster. Only this relatively small sub-set of states or 'credible accidents' can be specifically guarded against with the use of safety sub-systems and procedures (the hatched area in Figure 3). However, a probably larger sub-set of *unknown* system states can also lead to disaster. These pathways are much more difficult to guard against because the causal mechanisms involved are simply unknown, and this represents a blind spot in system safety.

During the TMI-2 accident, the night-shift operators were aware that the reactor was off-normal, but believed its state remained within the boundaries of known states (that is, they believed the reactor's state had migrated from A to B in Figure 3). Their attempts to move back to a desired state, therefore, made use of standard procedures. In fact, the reactor's state had migrated all the way to point C and was possibly within 30 minutes of attaining state D. Migration to D from either A or B is difficult by design because of the nuclear plant's extensive safety sub-systems. Migration from C to D however, is via the system safety blind spot, which bypasses safety sub-systems. Moving the reactor back to a desired state from anywhere within the set of unknown reactor states requires at least some degree of reasoning from first principles because there can be no specific rules for dealing with unknown states.

The role of incident reporting and better system design

Incident reporting, or the reporting of the details of events where things have gone wrong so as to learn from them, is a common safety strategy in a number of industries (Heinrich, 1959; Helmreich and Merritt, 2001; Webster *et al.*, 2001; Webster *et al.*, 2015). Incidents may comprise near-miss or accident events varying in severity, and can be used to develop new procedures and to redesign physical systems to reduce the occurrence of events in the future. The benefit of incident reporting can also be seen diagrammatically in terms of the state-space approach. Incident reports increase the set of known system states at the cost of unknown system states, thus expanding the known-state circle (shown as the new solid line in Figure 4a). This allows better and more inclusive procedures to be developed for previously unexpected off-normal system behaviour. In this way, reasoning from first principles, which is effortful, error-prone and time-consuming, will be required less often. After the TMI-2 accident, an expanded training programme for TMI-1 personnel was developed. Much of the training was carried out using simulation in an US$18 million, full-scale replica of the TMI-1 control room. Extensive revisions of standard procedures also occurred, including the establishment of a more active incident reporting scheme (Schlager, 1994; Chiles, 2001).

Importantly, the expansion of the known-state circle creates increased scope for engineering solutions, or system-based safety mechanisms to guard against a now larger set of known disaster states. This is possible because a larger portion of the set of disaster states is now included in the set of known states (shown as an increased hatched area in Figure 4a – compare with Figure 3). For example, the TMI-2 accident prompted over 100 safety modifications to be made at the twin TMI-1 reactor, costing US$95 million (Schlager, 1994; Chiles, 2001).

In addition, new generation system designs, based on incident data, can also be shown to affect the state-space for the better. Inherently safer designs create the possibility of systems

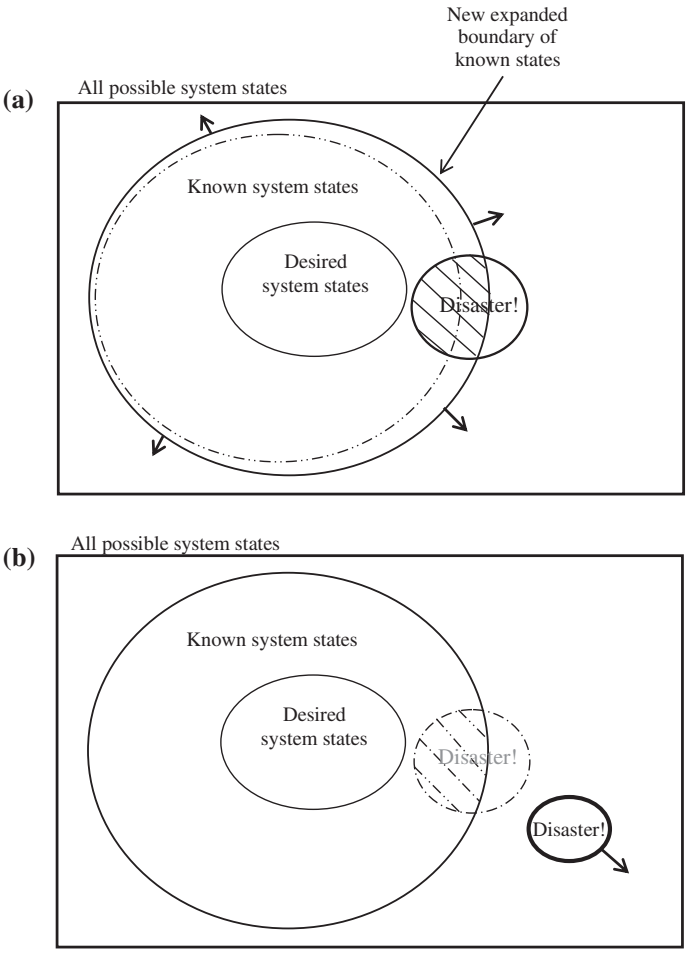

Figure 4. State-space diagram in Figure 3 after it has undergone adjustment to reflect two distinct safety initiatives.

Note: Incident reporting has allowed the number of known system states to be increased (hence the known states circle is larger in Figure 4a). This has allowed better procedures to be developed. More importantly, however, the expansion of the known state circle creates increased scope for system-based safety mechanisms to guard against a larger set of known disaster states. This is reflected in an increased hatched area in Figure 4a. In the second safety initiative, inherently safer system designs in which no known system state can lead to disaster, reduces the set of disaster states and removes the intersection between disaster states and known states (Figure 4b).

that have a significantly smaller set of states that lead to disaster (hence the smaller set of disaster states in Figure 4b). Based on the analysis of previous accidents, new generation nuclear plants have been proposed in which there are purportedly no known credible events or sequences of events capable of leading to a meltdown. Some of these designs involve reactor cores capable of complete self-stabilisation by passive processes, rather than requiring the active intervention of many complex automatic safety or cooling systems (Morone and Woodhouse, 1986). In theory at least, such systems are safer by virtue of having less complex interactions between sub-systems and being relatively less tightly coupled. A system in which no known credible event can lead to disaster removes the intersection between disaster states and known states (shown as the repositioned set of disaster states

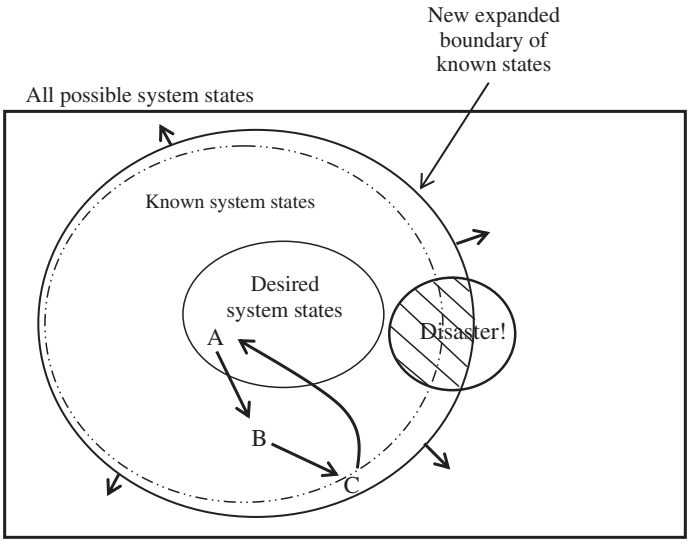

Figure 5. State-space diagram for the successful execution of a recovery pathway in a generic complex system.

Note: The set of known states has been expanded by incident data, allowing the refinement of crisis management strategies such that an accident pathway that previously resulted in disaster has led back to a desired system state primarily with the use of rule-based reasoning. (Figure reproduced with permission from John Wiley and Sons.)

in Figure 4b). Safer, new generation reactors of this sort do not remove the possibility of an unknown state causing disaster, but do place a wider safety margin between disaster states and desired reactor states (Figure 4b).

Accident recovery pathways – incident data in action

After the TMI-2 accident, a number of commentators pointed out that a similar incident in which a PORV had failed to close had occurred at the David-Besse reactor in Ohio in September 1977 (Perrow, 1984; Chiles, 2001). Disaster there had been avoided, partly because the reactor was not operating at full power at the time. In-house procedures were subsequently revised at David-Besse to include the possibility of a stuck PORV, but this information was not shared with other plants. Had the TMI-2 operators been aware of this incident, it is likely they would have diagnosed a loss of coolant due to an incorrectly open PORV before the core began to melt. Further evidence of the value of incident reporting in the aftermath of the TMI-2 accident was demonstrated in 1982 at the Robert E. Ginna nuclear power plant in Ontario. During a pressure release incident, a PORV again failed to close causing transient loss of coolant from the core. With knowledge of previous open PORV incidents, operators quickly diagnosed the problem, and damage to the reactor core was avoided (Schlager, 1994).

Figure 5 describes the averted disaster at the Robert E. Ginna nuclear plant in terms of the state-space approach. The dotted boundary of the known-state circle indicates the extent of known system states without knowledge of the possibility of an incorrectly open PORV (that is, the set of known states available to TMI-2 personnel). The inclusion of information about open PORV incidents increases the set of known states to the

solid boundary in Figure 5. The accident pathway from A to C is identical to that which occurred during the TMI-2 accident, but now point C is included in the set of known states for which a standard, rule-based procedure is available. Timely implementation of the new standard procedure causes the state of the reactor to migrate from C to A, and disaster and reactor damage are avoided.

The 'garden' in the machine and more advanced incident reporting

The nuclear power plant is simultaneously one of the most complex and potentially dangerous man-made systems in existence. Early engineers discovered that this complexity made them behave very differently from previous steam boiler technology, despite many of the same sub-systems being used. Safety specialists of the 1960s and 1970s discovered that, despite huge investment in safety systems, the question of just how much safety was safe enough was often unanswerable.

Anthropologist Constance Perin, after studying nuclear power facilities throughout the US, directly compares nuclear power plants with living things because of the unpredictability of their behaviour (Perin, 1998). She concludes that their behaviour is more akin to the phenomenon of the natural world than the technological:

> Is there a garden in this machine? Once it is operating, this technology can be no less puzzling than nature. Mishaps are evidence of as yet unanalysed arrangements between parts and whole, the unintended consequences of the dynamically unstable and spontaneous activity of this whole: phenomena more familiar to the natural and social sciences than to technoscience. (Perin, 1998, p.118)

Perin suggests that the most effective approach to dealing with this inherent complexity and unpredictability is to treat reactor operations as experiments. Experiments are intended to supply information to refine current knowledge and help predict future behaviour. As such, more in-depth, on-going data should be collected on the continuous variability of the reactor, rather than simply reporting accidents. Perin gives a telling example of a sequence of events that started with a switch in the wrong cabinet of switches being flipped and ended with the shutting down of the reactor. Because on this occasion a reactor shut-down occurred, the sequence was reported as an accident. But as a plant operator admitted, 'We've made the wrong cabinet mistake a hundred times, but it never before shut down the plant' (Perin, 1998, p.108). No data were collected on previous occasions as to the circumstances of the wrong cabinet mistakes and so no proper analysis could be performed to consider ways in which to correct a clear ergonomic problem in the plant's design. In addition, the lack of data from previous similar mistakes means little progress can be made to understand why, on this particular occasion, the error led to a reactor shut down when on numerous previous occasions it had not. The collection of incident data on mistakes of this sort, or any factor perceived as predisposing a more serious error or accident (precursor events), lowers the threshold for incident reporting and generates richer data on why things go wrong (Webster et al., 2001; Webster, 2012; Webster et al., 2015). The threshold for the reporting of accident precursors in the US nuclear power industry has recently been defined as any event with an estimated probability of as little as 1×10^{-6} (or 1 in 1 million) of causing damage to the reactor core (Belles et al., 1996).

Table 1. Safety milestones in socio-technological systems derived from safety strategies in the nuclear power industry.*

1	*A priori* design of purportedly fail-safe facilities based on an understanding of the constituent parts of the system rather than their interaction or emergent properties
2	Accident reporting leading to relatively minor equipment and procedure refinement
3	Simulator training to improve operator and teamwork performance, often in response to rare and emergent system behaviour
4	Major system-based redesign based on in-depth analysis of previous failures (i.e. the design of inherently safer, new generation equipment and facilities)
5	The lowering of the threshold for incident reporting to include pre-cursor events, other than accidents, that may adversely affect the efficient operation of the system (Perin's operating-as-experimenting approach – see text)

Note: *Adapted and expanded from Webster (2005).

Critical safety milestones in socio-technological systems

Safety strategies developed in the nuclear power industry since its conception in the 1940s can be broadly summarised into five steps or milestones (Table 1). The steps occur in historical order, each adding more sophisticated and robust safety strategies. Each step may also be viewed as indicating changes in the underlying presumptions about the nature of the particular technological system involved. In terms of nuclear power plants, a presumption underlying Step 1 was that these were inherently predictable, controllable and therefore safe devices. This confidence on the part of the engineers and designers of nuclear plants came from their understanding of the plant's constituent parts, many of which had been reliably employed in other industrial settings for many decades. Engineers and designers did not anticipate that the new complexity and tight coupling of nuclear plants would lead to unanticipated component interactions and unplanned system behaviours, with significant implications for controllability and safety.

Data collected in Step 2 did not necessarily contradict the presumptions underlying Step 1, but indicated only that refinements on the Step-1 designs were required. Presumptions underlying Step 3 were that safety in existing facilities could be maintained at appropriate levels through simulation training and adapting the operators to the peculiarities of the system. However, by Step 4, the *a priori* designs from Step 1 were viewed as inherently unsafe, prompting wholesale modification of existing facilities and the design of a new generation of inherently safer reactors. By Step 5, reactors were considered inherently unpredictable because of their complexity, making them more akin to living things (the antithesis of Step 1). This means that even the experts' understanding of the system is incomplete. Therefore, long-term safety improvements at Step 5 can be achieved only through treating operation of the reactor as an on-going experiment in which the collection of data allows better understanding and prediction of the system's behaviour in future.

Use of the framework

The framework presented here comprises two components – a state-space approach for better understanding system dynamics, and a set of milestones describing the development of safety in complex socio-technological systems. The state-space approach allows better conceptualisation of system failures, incident reporting and remedial initiatives – demonstrated here using examples from the nuclear power industry, and elsewhere with an example of anaesthesia in the operating room (Webster, 2005). The milestones can be

Table 2. Implications for healthcare at each framework milestone.

1	Complexity in healthcare is increasing rapidly, meaning that specialist knowledge of the isolated elements and episodes of care are no longer sufficient to guarantee acceptable levels of patient safety
2	Accident reporting has led to refinements in certain healthcare systems, but more substantial system redesign is required
3	Simulator training to improve operator and teamwork performance is becoming more common in healthcare, but in order to achieve safety targets set by the US Institute of Medicine (see text), safety initiatives must address more than the individual and team behaviour of clinicians
4	Evidence-based system redesign has occurred in certain practice areas in healthcare, but these benefits are unevenly distributed
5	Thresholds for the reporting of incidents to include error and accident precursors has occurred in certain high-risk and safety-conscious areas of healthcare, but reporting elsewhere remains patchy

used as a yardstick to assess the development of safety in socio-technological industries. For example, Table 2 demonstrates the use of the framework milestones to assess the relative safety progress of healthcare. Healthcare in first world countries has recently achieved Step 3 (simulation training); however, much work remains to be done at Step 4 and Step 5 (respectively, major evidence-based system redesign and the lowering of the threshold for the reporting of incidents). Taken together, this work formalises and strengthens the analogy of systems safety between the nuclear power and healthcare industries, and demonstrates a relatively mature safety culture in the former and a still developing safety culture in the latter (Webster, 2005; Cook and Rasmussen, 2005; Pronovost and Hudson, 2012). Although significant piecemeal improvements in safety have occurred in healthcare in recent years, these fall considerably short of the goal of a 50% reduction in medical error across the board set by the United States Institute of Medicine in 2000 (Institute of Medicine, 2000; Webster, 2012). Greater investment in safety strategies at Steps 3–5 would yield safety benefits capable of achieving the Institute of Medicine's goal (Table 2).

Concluding remarks

Despite significant safety challenges and the potential for devastating failure, the US nuclear power industry has made early and systematic use of the science of safety. Nuclear power plants are tightly coupled and highly complex socio-technological systems, and this combination of characteristics renders the nuclear power industry a kind of worse case test-bed for the development of safety strategies. Based on key events in the development of safety in the US nuclear power industry, I have described and formalised a framework for the conceptualisation of safety goals in socio-technological systems. Safety analogies can be important and powerful ways to conceptualise safety goals and to give direction to new safety initiatives. In this regard, much has been made of the safety culture of the aviation industry as a model for safety in healthcare. However, the healthcare industry comprises a tightly coupled and highly complex socio-technological system with characteristics more similar to those of the nuclear power industry than those of aviation, suggesting that the nuclear power industry is a more appropriate comparator for healthcare safety analogies. Applying the framework to the analysis of safety in healthcare indicates that healthcare has a safety culture that is still developing, and that substantial safety progress remains to be made through major evidence-based system redesign, and the lowering of the threshold for the reporting of incidents to include precursors to errors and accidents.

Notes

1. The Manhattan Project was a research programme of the US government, active between 1942 and 1945, which produced the first atomic bombs.
2. This simplified version of events is intended to demonstrate some of the extraordinary forces, costs and unpredictability involved in this kind of complex, intensified technology. US dollar amounts have not been adjusted for inflation
3. US statistics on preventable deaths during healthcare are used for comparison simply because they are so well known. Evidence suggests that the problem of treatment-related harm during healthcare is an international one, and is no worse in the US.

Disclosure statement

No potential conflict of interest was reported by the author.

ORCID

Craig S. Webster http://orcid.org/0000-0002-6997-4263

References

Anonymous (nda) 'Early nuclear power plants', *Wikipedia*, available from https://en.wikipedia.org/wiki/BORAX_experiments [accessed April 2016].
Anonymous (ndb) 'Nuclear reactor accidents in the United States', *Wikipedia*, available from https://en.wikipedia.org/wiki/Nuclear_reactor_accidents_in_the_United_States [accessed April 2016].
Anonymous (ndc) 'World-wide nuclear and radiation accidents by death toll', *Wikipedia*, available from http://en.wikipedia.org/wiki/List_of_nuclear_and_radiation_accidents_by_death_toll [accessed April 2016].
Barach, P. and Johnson, J. (2006) 'Understanding the complexity of redesigning care around the clinical microsystem', *Quality and Safety in Health Care*, 15, pp.i10–i16.
Barach, P. and Small, S. (2000) 'Reporting and preventing medical mishaps – lessons from non-medical near miss reporting systems', *British Medical Journal*, 320, pp.759–63.
Belles, R., Cletcher, J., Copinger, D., Dolan, B., Minarick, J. and O'Reilly, P. (1996) '1994 accident sequence precursor program results', *Nuclear Safety*, 37, pp.73–83.
Casey, S. (1998) *Set Phasers on Stun and Other True Tales of Design, Technology, and Human Error*, Aegean, Santa Barbara.
Chiles, J. (2001) *Inviting Disaster – Lessons from the Edge of Technology*, Harper Collins, New York.
Clarfield, G. and Wiecek, W. (1984) *Nuclear America – Military and Civilian Nuclear Power in the United States 1940–1980*, Harper and Row, New York.
Cook, R. and Rasmussen, J. (2005) '"Going solid": a model of system dynamics and consequences for patient safety', *Quality and Safety in Health Care*, 14, pp.130–34.
Department of Health (2000) *An Organisation with a Memory – Report of an Expert Group on Learning from Adverse Events in the NHS*, Stationery Office, London.
Gander, P., Millar, M., Webster, C. and Merry, A. (2008) 'Sleep loss and performance of anaesthesia trainees and specialists', *Chronobiology International*, 25, pp.1077–91.
Gawande, A. (2009) *The Checklist Manifesto – How to Get Things Right*, Metropolitan Books, New York.
Haynes, A., Weiser, T., Berry, W., Lipsitz, S., Breizat, A., Dellinger, E., Herbosa, T., Joseph, S., Kibatala, P., Lapitan, M., Merry, A., Moorthy, K., Reznick, R., Taylor, B. and Gawande, A. (2009) 'A surgical safety checklist to reduce morbidity and mortality in a global population', *New England Journal of Medicine*, 360, pp.491–99.
Healy, J. (2011) *Improving Health Care Safety and Quality – Reluctant Regulators*, Ashgate, Aldershot.
Heinrich, H. (1959) *Industrial Accident Prevention –A Scientific Approach*, McGraw-Hill, New York.
Helmreich, R. and Merritt, A. (2001) *Culture at Work in Aviation and Medicine*, Ashgate, Aldershot.

Hunt, P. (1988) 'Safety in aviation', *Perfusion*, 3, pp.83–96.
Institute of Medicine (2000) *To Err is Human – Building a Safer Health System*, National Academy Press, Washington DC.
Kapur, N., Parand, A., Soukup, T., Reader, T. and Sevdalis, N. (2016) 'Aviation and healthcare: a comparative review with implications for patient safety', *Journal of the Royal Society of Medicine Open*, 7, 1, pp.1–10.
Leveson, N. (2011) *Engineering a Safer World – Systems Thinking Applied to Safety*, MIT Press, Cambridge MA.
McKeown, W. (2003) *Idaho Falls –The Untold Story of America's First Nuclear Accident*, ECW Press, Toronto.
Merry, A. and McCall Smith, A. (2001) *Errors, Medicine and the Law*, Cambridge University Press, Cambridge.
Morone, J. and Woodhouse, E. (1986) *Averting Catastrophe – Strategies for Regulating Risky Technologies*, University of California Press, Los Angeles.
Perin, C. (1998) 'Operating as experimenting – synthesizing engineering and scientific values in nuclear power production', *Science, Technology and Human Values*, 23, pp.98–128.
Perin, C. (2005) *Shouldering Risks – The Culture of Control in the Nuclear Power Industry*, Princeton University Press, Princeton NJ.
Perrow, C. (1984) *Normal Accidents – Living with High Risk Technologies*, Basic Books, New York.
Plsek, P. and Greenhalgh, T. (2001) 'The challenge of complexity in health care', *British Medical Journal*, 323, pp.625–28.
Pronovost, P. and Hudson, D. (2012) 'Improving healthcare quality through organisational peer-to-peer assessment: lessons from the nuclear power industry', *BMJ Quality and Safety*, 21, pp.872–75.
Reason, J. (1990) *Human Error*, Cambridge University Press, New York.
Sagan, S. (1993) *The Limits of Safety – Organizations, Accidents, and Nuclear Weapons*, Princeton University Press, Princeton NJ.
Schlager, N. (ed.) (1994) *When Technology Fails – Significant Technological Disasters, Accidents, and Failures of the Twentieth Century*, Gale Research, Detroit.
Vincent, C. and Amalberti, R. (2016) *Safer Healthcare – Strategies for the Real World*, Springer Open, London.
Webster, C. (2002) 'Why anaesthetising a patient is more prone to failure than flying a plane', *Anaesthesia*, 57, pp.819–20.
Webster, C. (2005) 'The nuclear power industry as an alternative analogy for safety in anaesthesia and a novel approach for the conceptualisation of safety goals', *Anaesthesia*, 60, pp.1115–22.
Webster, C. (2012) 'Overcoming complexity and improving the safety of medical systems', *Prometheus*, 30, pp.320–26.
Webster, C., Anderson, B., Stabile, M. and Merry, A. (2015) 'Improving the safety of pediatric sedation: human error, technology and clinical Microsystems' in Mason, K. (ed.) *Pediatric Sedation Outside of the Operating Room: A Multispecialty International Collaboration*, Springer Science, New York, pp.587–612.
Webster, C. and Grieve, D. (2005) 'Attitudes to error and patient safety', *Prometheus*, 23, pp.253–63.
Webster, C., Merry, A., Larsson, L., McGrath, K. and Weller, J. (2001) 'The frequency and nature of drug administration error during anaesthesia', *Anaesthesia and Intensive Care*, 29, pp.494–500.
Weick, K. (1976) 'Educational organizations as loosely coupled systems', *Administrative Science Quarterly*, 21, pp.1–19.