

Failure to Mobilize in Reliability-Seeking Organizations: Two Cases from the UK Railway*

J. S. Busby

Lancaster University

ABSTRACT There is a considerable line of research on organizations dealing with large scale, intrinsic hazards. We know a good deal, as a result, about both the causation of catastrophic failure and its avoidance. Past work has explained failure in terms of (for example) structural vulnerabilities and organizational degradation – and reliability in terms of collective mindfulness, rigorous enculturation and high levels of social redundancy. This paper presents a study, based on a qualitative analysis of two disastrous collisions on the UK railway, of organizations that are strongly reliability seeking yet ultimately experience catastrophic failure. It argues that these cases implicated an organizational incapacity to mobilize systemic reform. The possibility of the two failures had been well-known in the organizations before their occurrence, but this knowledge could not be converted into modification. A model is presented to explain how processes of systemic reform co-exist with a set of phenomena that tend to undermine them. It is these that need to be the principal focus of efforts at managing catastrophic failure risks.

INTRODUCTION

A central concern for a range of organizations is how to avoid killing people in the course of operating technologies that present intrinsic, large-scale physical hazards. Such organizations – operating in industries like mass transportation, chemical processing, oil and gas production – simultaneously face a physical context of increasing scale, complexity, distributedness and interconnectedness (Schieb, 2003), and a social context increasingly oriented around risk (Beck, 1992). Managing the risk of catastrophic failure has thus become a pivotal problem. The occasional but highly consequential catastrophes that have occurred have led to a substantial line of research (e.g. Elliott and Smith, 2006; Perrow, 1984; Rasmussen, 1990; Reason, 1990; Turner, 1976) on how catastrophic failure takes place in sociotechnical systems, and how organizational phenomena are implicated in such failures. This has revealed much about organizational vulnerabilities and the processes of gradual degradation that generate them.

Address for reprints: J. S. Busby, School of Management, Lancaster University, Lancaster LA1 4YX, UK (j.s.busby@lancaster.ac.uk).

At the same time a further line of work (e.g. La Porte and Consolini, 1991; Roberts, 1990) has been investigating how certain organizations – such as naval aircraft carriers – nonetheless achieve extraordinarily high levels of reliability and safety, despite the intrinsic physical hazards they deal with. We are left to wonder how far this achievement could be extended beyond those organizations, and how organizations that approach – but seem not to attain – high reliability in fact fail. We might ask how, in particular, catastrophic failure of a kind that has been foreseen in an organization with a very substantial history of experiencing and managing large-scale, widely-distributed risk, still occurs. The study we describe here was an attempt to do this. It is a study of reliability-seeking that ultimately proves inadequate.

It is also a study that stresses shortfalls in the capacity of organizations to bring about a safe system – rather than shortfalls in their action within that system. Thus, whereas much previous work stresses capacities and flaws in that part of an organization engaged in day-to-day operations, our study points to a particular vulnerability in the part that is one stage removed from operations – the part that assesses and changes the systems and norms that structure these operations. Whereas previous work has emphasized such problems as interruption, stress and the absence or loss of requisite variety, our work concludes that reliability can ultimately be limited by inability to bring about a suitably robust system. The failure of the operating organization to make sense of emergent situations seems not to be the core problem in the cases we analyse: it is more a consequence of the maintaining organization failing to mobilize systemic reform.

The study consists of a qualitative, grounded analysis of two catastrophic failures on the UK railway. Both involved train collisions; both killed between 30 and 40 people and injured another several hundred; both involved signalling as the central, problematic technology. One occurred before the privatization and wholesale reorganization of the UK railway, and one after. As a grounded analysis the only prior commitment was to finding out what organizational phenomena had undermined the structures and mechanisms that should have kept the railway safe.

BACKGROUND

Our point of departure is in the literatures of failure and reliability. The failure literature is diverse but can be broadly divided into two main parts: the first concerned with the nature of an organization's vulnerability to disaster at any particular time, the second with the gradual processes of degradation that lead an organization into a state of vulnerability. The epitome of the first is perhaps Perrow's (1984) 'normal accidents' model, in which the systemic properties of interactive complexity and tight coupling – and the contradictory requirements they impose on organizational decision making – make accidents virtually inevitable. Perrow's ideas have been criticized on several grounds (Bierly and Spender, 1995; Hopkins, 1999; Roberts, 1990; Weick, 1990). But the intractability of contradictory requirements on organizational capacities has been a theme in other work (Roberts, 1990; Turner, 1994). And systemic properties like tight coupling are strongly linked to the temporal conditions that produce vulnerability – such as compressed timescales, the need to produce multiple, simultaneous, critical outcomes

(Roberts and Rousseau, 1989), and the need to maintain continuous operation (Klein et al., 1995). The problem is exaggerated in systems that are sensitive to small, discrepant conditions. Roberts et al. (1994), for instance, provide an analysis of how an aircraft carrier operation deals with the problem of an unidentified part being found on the deck. The fact that a single, stray component in such a vast system can be profoundly troubling indicates how vulnerable the system is to a minor discrepancy.

The literature that deals with degradation has arisen with the observation that generally catastrophic failure tends not to take place in pristine systems in which organizational mechanisms are functioning in the way they are designed. It takes time before systemic vulnerabilities escape the capabilities of organizations to deal with them. Turner's (1976, 1978) model of man-made disaster was among the earlier approaches, and laid particular stress on a period of 'incubation' in which there is a gradual progression towards a failure that is not seen as such at the time. Models of latent failure (Reason, 1990) and latent error (Ramanujam and Goodman, 2003) similarly propose processes in which the preconditions of failure accumulate until some precipitating event turns them into identifiable failure sequences. This accumulation tends to go unnoticed, as failures of organizational cognition (Pidgeon and O'Leary, 2000) and political processes (Gephart, 1984) lead to the discounting of signals of an incipient disaster.

This degradation has also been linked to notions of entropy and the absence of maintenance in organizational systems (Grabowski and Roberts, 1996), the erosion of protective forms of slack (Schulman, 1993), and the gradual incorporation into organizational practice of short-cuts and 'patches' that render their operation hard to understand (Weir, 1996). It is associated with reinforcing feedback loops (e.g. Toft and Reynolds, 1994) that move such practices steadily further from their normative forms. Sometimes such drift is linked to local adaptations that ignore the safety of the system as a whole (e.g. Rasmussen, 1990; Snook, 2000). Sometimes it arises from the denial of resources to protective capacities (Marcus and Hichols, 1999), the 'gradients' that arise from competitive pressures (Rasmussen, 1997) and the tendency in organizations to convert gains in protection into gains in production (Reason, 1997).

The reliability literature provides a basic critique of this general line of work on failure. Studies of failure generally privilege one state of affairs (failure) over another (the absence of failure), use retrospection as the main tool of analysis, and distract attention from the fact that catastrophic failure is uncommon. 'High reliability organizations', such as those found on aircraft carriers and submarines, give particular evidence of how organizations avert disaster. For example, they incorporate vast redundancy in both organizational and physical systems (Roberts, 1990). They distribute decision making to the lowest possible level of the organization (Roberts, 1990). They seem to be capable of making rapid transitions from rigid, centralized structures to high tempo modes of organization in which hierarchical rank defers to technical expertise, and to emergency response modes that involve pre-programmed, practised behaviours (La Porte and Consolini, 1991; Roberts et al., 1994). They have organizational members who engage in 'heedful inter-relating' (Weick and Roberts, 1993), 'respectful interaction' (Weick, 1993) or 'solicitude' (Pidgeon, 1997). They enjoy powerful processes of enculturation that heavily emphasize reliability, and this allows them to maintain flexibility in local decision making while simultaneously ensuring an homogeneous set of decision premises (Weick, 1987).

They achieve requisite variety without making excessive cognitive demands by employing generalized scripts (Zohar and Luria, 2003). They take strong cues from external scrutineers, such as regulatory staff (La Porte and Thomas, 1995) – not simply internalizing but also extending the norms that regulators impose. And they engage in rigorous learning: processes of continual search for improvement, systematic feedback and review, and the rewarding of discovery of error (La Porte, 1996). They continually reflect on practice (Pidgeon, 1997) and generally appear to conform to Westrum's (1992) idea of 'generative' organizations that actively seek information, welcome new ideas and train messengers. They continually rehearse familiar failure scenarios, strive hard to imagine new ones, and generalize on their experience of failures (Reason, 2000). And they tolerate and even protect key organizational ambiguities in order to learn and cope with the unforeseen (Schulman, 1993).

There is, however, a wide-ranging critique that can be directed at the high reliability organizations work in turn. The supposed protective benefits of redundancy are severely limited by the complexity it introduces (Sagan, 1993), by social shirking (Sagan, 1994) and the prevalence of common-mode failure in human activity (Hollnagel, 1993) – and what protective redundancy there is often reflects a limited, single-loop learning (Bain, 1999). Decentralized decision making, even simply decentralized vetoes, can threaten rather than support safety, which is a global criterion and therefore highly vulnerable to local adaptations (Rasmussen, 1990). The claims made for enculturation are problematic, both for the limited efficacy of culture in overcoming basic structural vulnerabilities (Perrow, 1999), and for the strictly limited extent to which organizations in an individualistic democratic society can achieve the kind of high reliability culture observed in a military organization like Rickover's nuclear-powered submarine service (Bierly and Spender, 1995). La Porte and Thomas's account (1995) of an organization extending regulatory norms suggests, on a closer reading, that organizational members were at least in part being motivated by the possibility that the regulator would catch them out, not by intrinsic questions of reliability or safety. And the heavy emphasis on learning seems to gloss over several difficulties. For example, there is Carroll's (1998) observation of the disparate logics among the groups involved in managing hazardous systems, and the way this disparity can impede learning. Sagan (1994) points out that, in learning from both simulation and exercises, near-misses are easily covered up by powerful actors. And Rijpma (1997) argues that while learning might start off as a cognitive process it rapidly becomes a political one: it is more often about learning to build coalitions than learning to understand events.

The difference in emphasis found in the traditions of studying failure and studying success naturally leads those who study failure to be pessimistic about the perfectibility of operational controls (Sagan, 1993), and those who study reliability to be optimistic. This has meant that the two approaches have sometimes been put in opposition. But the results of comparing their explanatory power have been inconclusive (Rijpma, 2003), and from a logical standpoint it is something of a false opposition (Bain, 1999; Jarman, 2001). A more productive question is perhaps to ask what occurs somewhere near the margins of what we have taken to be high reliability organizations – to ask what occurs in organizations that do experience catastrophic failure against a background of strong reliability-seeking. This ought to tell us much about what ultimately

limits our capacities to manage risk, and what we can reasonably expect of organizations for which risk is an inevitable accompaniment to operation.

THE CASES

The essence of the Ladbroke Grove disaster in 1999 was that a local train collided with an express at a closing speed of about 130 mph, killing 31 people and injuring 416. The local train was heading west from the nearby terminus and crossing the line occupied by the express heading east into the terminus. It had passed a signal set at danger, probably without the driver realizing. Signallers, on seeing this, were slow to react – to set the signal at danger in the face of the express train and to communicate over a secure radio with the driver of the local train. They appeared to expect the local train to stop within the safety ‘overlap’ beyond the signal at danger as that was their normal experience. After the signal in question there was an unusually large overlap before the fouling point. The driver was inexperienced and had passed through a flawed training regime that was in a state of transition. Among the flaws in this regime was the failure to describe to trainees specific risks on the line they were to work on, and to assess them on their particular knowledge of that line. Moreover, the signalling was problematic in several respects – its visibility was poor in particular conditions, it could only be seen for a short interval before it was passed, it was hard to associate with the correct track, and it had an unusual configuration. The inquiry found that there were considerable organizational deficiencies lying behind these immediate events. Problems with the signalling were well-known to both the organizations involved in the disaster and to the regulatory organization that supervised them. Attempts to deal with these problems had been uncoordinated and compromised. The intrinsically hazardous track layout had been conceived with productivity rather than safety in mind. There had been organizational failures to learn from the experience of the same signal having been passed at danger on several previous occasions. In addition, the organizational fragmentation that had followed privatization had left a legacy of lost expertise and instability in organizational structure.

The Clapham Junction disaster in 1988 again involved a collision of two passenger trains, and the subsequent collision of one of these with an empty train. Again the signalling system was implicated, but in a quite different way. A full commuter train, rounding a bend, came upon a signal set at green. When the train was about 30 metres short of this signal its aspect changed to red. The driver cruised to the next signal and stopped there to call the signalling centre. The signal that preceded this (the one that had changed from green to red) was an automatic one, and should have been set to red indefinitely while a train was in the section beyond. Due to a wiring fault, this signal reverted to green as the traffic ahead of the halted train cleared. A second train approached, around the bend, and saw the green aspect. This train collided with the rear of the stationary train at about 35 mph. This displaced the second train, and an oncoming empty train collided with this – killing 35 people and injuring 500. The wiring fault arose from a combination of two errors: one a systematically flawed working practice, the other a lapse occasioned by an interruption. Both errors revealed profound

deficiencies in the signalling organization that meant poor practices could persist and become endemic, and systems of checking and verification could become ineffectual.

Methodology

The approach was to undertake an inductive, grounded analysis of the public inquiry proceedings following the two disasters, with the general objective of deepening our understanding of how organizations seek yet fail to attain reliability. Our data consisted of two texts – the reports of the inquiries made by the inspectors (Cullen, 2000; Hidden, 1989) – an approach that has similarities to, *inter alia*, that of Turner (1976), Gephardt (1993), and Woo and Vicente (2000). This process assumes a basic positivism in which the inquiries have a relationship to a certain set of events that are independent of the observers, even if the view of those events is necessarily a mediated one. Its reliance on the recall and candidness of witnesses to the inquiries is a considerable limitation, but is offset by the large numbers of witnesses and the use of large quantities of documentary evidence. It is difficult to avoid the effects of hindsight – and Pidgeon (1997) argues that very few accounts of failure have made a serious attempt to present the uncertainties experienced at the time by those involved in the failure. But again this is mitigated by the considerable efforts that both inquiries made to understand how situations and systems appeared to those involved in them, at the time. It is also important to bear in mind the advantages of this form of data. Typically it is impossible to observe rare but catastrophic events at first hand, and first-hand observation is only ever partial. Moreover, formal investigations, especially those for public enquiries, are rigorous, well-funded, highly-resourced and mostly exhaustive. The array of experts who participate in them could not be replicated in a research study. At the smaller of the two inquiries (Clapham Junction), evidence was heard from 122 witnesses over a period of 56 days, and over 13,000 pages of documents were considered. Both inquiries had been preceded by immediate investigations carried out by the regulator and industry, but the public inquiries did not draw on these prior investigations and there appear to have been no allegations of bias or partiality. Moreover, the Cullen inquiry was permitted by the government's legal officer to give an undertaking that evidence submitted to the inquiry would not be used in any legal action against the person providing it. Details of the inquiries' methods, witnesses and materials (such as the computerized recordings made at the signalling centre) are laid out in the reports.

The analysis consisted of three main steps. The first was to assemble a causal tree for each case – collecting together all the attributions and diagnoses made in the inquiry according to a theoretically neutral organizing principle (the apparent pattern of causation). There were 232 unique elements in the Ladbroke Grove causal tree, and 102 in Clapham Junction. Figure 1 shows a simplified fragment of the causal tree for Ladbroke Grove, in which the convention is that immediate causes of an element are shown below and indented to its right. The second step was a pre-theoretical selection of significant fragments of these trees that referred to reliability seeking. For example, a fragment of the Ladbroke Grove tree demonstrated how the organization tended to suffer a particular degradation of its systems shortly ahead of large-scale investments. These investment programmes tended to draw together the replacements of multiple, obsolescent tech-

- ...
- Delay in signaller transmitting emergency stop to driver on cab secure radio
 - Signaller was reading worksheets not looking at VDU when first alarm sounded
 - With automated signalling signaller does not have immediate knowledge of state system is in
 - Identification of which alarm had sounded in control panel takes finite time
 - Control panel gives common sounding for many conditions
 - No training given in emergency use of cabin secure radio
 - Use of cabin secure radio not entered common signalling organization practice
 - Failure of signallers to comprehend importance of immediate driver communication
 - Signallers generally diagnosed SPADs [Signals Passed At Danger] as driver error
 - There were no formal debriefings following SPADs
 - There was no training in management of SPADs
 - SPADs occurring in this area were frequent
 - SPAD investigators lacked expertise in human factors and root cause analysis
 - Signallers fail to see themselves as part of the system of preventing SPADs
 - Absence of ATP [Automatic Train Protection] on train 2
 - Cost benefit analysis concluded costs of ATP outweighed benefits
 - Historical ATP report concluded cost exceeded normal safety investment criteria
 - Train operator's analysis in 1996 concluded costs outweighed benefits
 - Consultants' study in 1998 concluded costs outweighed benefits
 - Operating experience of ATP reliability was poor
 - TPWS [Train Protection and Warning System] likely to proceed as alternative to ATP
 - Secretary of State had accepted case for TPWS in 1995
 - Regulator had supported TPWS in report of 1997/1998
 - Expectation that TPWS would provide benefit quicker
 - Provision made in train operator's 2001 budget for TPWS
- ...

Figure 1. Fragment of the causal analysis of Ladbroke Grove

nologies in single, simultaneous undertakings. This 'massing' phenomenon meant that there were lengthy periods in which known deficiencies went uncorrected and simply accumulated. The third step in the analysis involved a further selection process – this time of fragments that in some way provided a refinement or modification to existing theory outlined in the Background – and a grouping of these into explicable categories. For example, the 'massing' effect was selected because it implied that degradation and corrective processes within organizations could be highly discontinuous. This was then grouped with a number of other effects to do with ways of organizing that somehow impeded the direct link between learning and acting.

The whole process is strongly idiographic. It does not provide a comprehensive test of existing theory, but simply attempts to find out what it does not account for. It also covers a great deal of data so it is only the outcome of the final stage of reduction that we have space to reproduce in the Results section that follows.

RESULTS

In this section the results are presented in two parts. The first surveys the categories that emerged from the selection and induction just outlined, and the second expands briefly on each category by offering descriptions and examples. This reflects the significance both of the general structure that emerges in the selected phenomena, and of the specific phenomena in their own right.

A Survey of the Categories

Table I lists the categories of significant phenomena arising from the analysis, together with brief descriptions and examples. These are the residue of the selection and abstraction process just described, and are collected under two more general categories into which they naturally appeared to fall. The first is a category of problematic efforts after reliability. This includes, for example, ‘mobilization’ processes in which the organizations are attempting to convert knowledge about systemic vulnerabilities into modifications, and ‘engendering’ processes in which one organization is attempting to induce reliability enhancements in another. The second general category is of effects that confound reliability seeking. This includes, for example, ‘disarticulation’ processes that interrupt plans for systemic reform, and outbreaks of ‘propriety’ where norms of respectfulness hamper conclusive action.

In the remainder of this section the categories are presented in a little more detail, while their theoretical significance is taken up in the Discussion.

Mobilization

The ‘mobilization’ category is of phenomena that influenced the organizations’ ability to turn knowledge into action – to embody or embed learning in the system. We have already referred to the ‘massing’ of systemic renewal, where necessary modifications are delayed once actors know there will be some kind of general change to the system. There was also a ‘proliferation’ effect, especially at Ladbroke Grove. As new problems and complaints emerged, all around the general problem of signal sighting and signals passed at danger, new working groups were established. The greater the number of groups that had been established, the easier it was to set up another – and the harder it became to synthesize effort across the groups. Thus proliferation became increasingly more likely than synthesis. There was an ‘intervention risk’ effect, such that intervening in certain systems was riskier than not intervening, at least up to some point where the risk of the status quo reached the risk of intervention. Prior to the Clapham Junction accident the signalling system over a large area had become so antiquated that large scale renewal was required – but it had taken many years of deterioration before the balance of risk led to renewal. And there was a ‘constrained redundancy’ effect, in that the kind of protective slack that the organizations needed was simply not available due to over-constraint. Thus, for example, it was not feasible to provide a ‘repeater’ signal that would (redundantly) alert the driver to the state of a signal because of various physical limitations. In all these instances there is a sense of there being ways of organizing that – somewhat unexpectedly – confound attempts to make systemic changes.

Continuity

The next category was of ‘continuity’ phenomena. In both cases, and especially at Ladbroke Grove, the organizations experienced difficulty in sustaining a coherent line or thread of development. In one case a thread (the improvement of driver training) came to a premature end when a manager moved to another post. In another case a thread

Table I. The categories obtained in the analysis

| <i>Category</i> | <i>Sub-category</i> | <i>Explanation</i> | <i>Example</i> |
|-------------------------|---------------------|---|--|
| Reliability seeking | Mobilization | Conversion of intelligence about reliability in the system to action towards it | Change to the confusing track system in the area was resisted in the expectation that automatic train control technology would ultimately be introduced network-wide |
| | Continuity | Maintenance of sustained pattern of reliability-oriented action | A requirement to implement recommendations on route learning, and concerns of the regulator on drivers at risk, were not passed on to an incoming operations director |
| | Instantiation | Generation of specific reliability actions from general stipulations | The implicated signal was configured in an unusual 'reverse L' aspect. This contravened a railway group standard but was allowed by a derogation. The issuer of the derogation had insufficient authority to do so but did not realize this |
| | Engendering | Influence of externally-derived reliability-seeking goals and norms | The regulator had undertaken a national 'signal passed at danger' audit which showed the implicated signal to be among the worst. It took no enforcement action because of the supposed risk of demotivating an organization that was generally trying to make improvements |
| | Disarticulation | Supplanting of organizational plans by events | The commissioning engineer failed to carry out a wire count. His post was temporary and followed displacement from a previous post. He had received no local induction training or handover from previous incumbent because it was known he was in place temporarily |
| Reliability confounding | Lag | Delay between action and its systemic efficacy | There was a failure to act on an action agreed in a working group because <i>inter alia</i> the feasibility study was subsequently disputed, the action in question was thought to have problematic side-effects, and the accident intervened before a proposal could be reviewed by operating companies |
| | Propriety | Norms of respectful behaviour displacing necessary action | The regulator had believed its involvement was inappropriate because the report of an earlier inquiry into a similar set of events was still pending, and this report had been delayed by questions of law |
| | Speciousness | Availability of wrong but plausible justifications | It was unclear whether the actions recommended by an inquiry in the same area was a matter of zonal or national action so they were neglected by both |

ended in the absence of any verification or checking that it had achieved its intended effect. In a further case a thread involving the consistent application of a particular risk assessment model over time was interrupted by a general change in organizational structure. Processes like risk assessment, which typically involve the use of historical data for future reasoning, fail around discontinuities of organization.

The 'continuity' phenomena thus refer to situations where there is some requirement for activity to be constituted in lines of connected action that need to be carried through to a state of completion, or over a prolonged period, in order to be effective. It is ironic that these lines of action, often aimed at achieving some reform, are undermined by some other act of reform.

Instantiation

The next category was of 'instantiation' phenomena: that is, phenomena to do with converting the general into the particular. They concerned the organization's dealing with specific situations, its translation from general rules to specific actions, and its treatment of the anomaly and idiosyncrasy typically found at the level of specific instances. Perhaps the starkest and simplest of these was the local empiricism that was especially evident in the Clapham Junction case. The technician had learned an unsafe practice, probably through a combination of experiential and social observation – which can only ever be partial, is sometimes idiosyncratic, and can be lacking in causal understanding. This empiricism displaced canonical, standardized practice. Moreover, the unsafe practice had become widespread, so that the protective redundancy of one person checking the work of another also became ineffectual.

Another kind of instantiation effect involved the 'averaging' that certain ways of organizing imply. At Ladbroke Grove, problems became aggregated as they ascended the management hierarchy in such a way that senior managers judged the mass of problems in terms of their average significance. At Clapham Junction, problems categorized as 'wrong side failures' received less attention than those categorized as 'signals passed at danger' because *on average* they were less consequential. But both processes overlooked outliers. Safety and reliability in a system of this kind cannot be managed as an averaged construct, since it only takes one outlier to cause a catastrophe. There is thus a basic problem for organizational cognition in preserving the outlier status of particular instances and avoiding natural managerial processes of averaging.

A third problem with instantiation involved the competing needs for standardization and adaptability. A distinctive aspect of railway operations is that there has traditionally been a strong culture of engineering standardization that should ensure uniformly high performance of safety-critical activity throughout the organization. But it had been necessary to allow 'derogations', whereby informed local decision-making could suspend the blunt action of a standard that is mostly right but occasionally wrong. The use of a problematic signal configuration at Ladbroke Grove arose from one such derogation.

Engendering

The category of 'engendering' phenomena concerned one organization or group seeking to engender reliability seeking in another. This particularly involved a regulatory orga-

nization, and the problem that – like most regulators – it can only sample the operations it regulates. Regulation is a central aspect of organization in industries that deal with large-scale hazards but the necessity of sampling means that regulatory actions extend beyond scrutiny and enforcement to persuasion. Regulatory action can be much more efficient if the regulated organizations internalize regulatory objectives.

One of the problems revealed by the railway cases was that discontinuity in the regulated organization can cause such internalized objectives to be undone. At the Ladbroke Grove inquiry the regulator was said to have placed excessive trust in the operating organization, having inappropriately extrapolated its experience of the organization as it had been before privatization to the organization that existed after. A second problem at Ladbroke Grove involved the regulator finding – before the accident – that the signal implicated in the accident had been among the worst in the regulator's national audit of 'signals passed at danger'. But it had not taken the consequent enforcement action because there was a supposed risk of de-motivating people in the operating organization who were 'doing good work'. The claim is thus that maximal use of external power impedes, rather than promotes, engendering – yet, if there is a failure and maximal power has not been used it is implicated in the failure. This causes a considerable bind for regulatory organizations.

Disarticulation

The 'disarticulation' category collected together phenomena that arose from the apparent undermining of organizational structure by emergent events. Becoming hostage to events as they emerge, being distracted from following plans and strategies, and having to relinquish standard routines in order to make *ad hoc* adjustments, all create a disarticulation of activity. At Ladbroke Grove the restructuring that had continued after privatization, and the regular appearance of new initiatives, led the managers who were meant to be dealing with signals passed at danger (SPADs) to suspect that whatever conclusions they drew could easily be overtaken by new possibilities and new thinking, and undermined their determination to carry their findings through. Prior to Clapham Junction, a pivotal individual had been one of several that were displaced from one post to another. The destination post was itself temporary, and led to the organization concluding that it would not be worthwhile inducting and training him for this post. This lack of training was implicated in the failure.

The disarticulation of an organization – the undermining of those structures that are needed to attain reliability – is not simply experiencing external events, but also about treating the central construct of organization as the flow of events in which it is engaged rather than the structures and regularities it must preserve over time.

Lag

Various phenomena were categorized as being concerned with 'lag': the lag that occurs between an organization having a concept of some way of organizing and that concept being enacted. The general effect is that modifying a system takes a finite time, and

because modification of some kind will always be needed to reflect changing circumstances this discrepancy between intentionality and actuality will be permanent.

These modifications are not necessarily technological ones: two instances were to do with training regimes. At both Clapham Junction and Ladbroke Grove, the training and development of staff had been subject to progressive, and sometimes discontinuous, improvement over time. This meant that more recently recruited people tended to have a better formal knowledge, and that there was a persistent legacy of staff who had joined the organization and had become experienced during poorer training regimes. The temptation is to believe that this is temporary, and that it is eventually put right as newer staff displace older staff. But as long as training regimes change regularly this problem will persist indefinitely.

Propriety

The category of 'propriety' effects involved notions of the proper way of doing things that turn out to be problematic. One instance (in the Ladbroke Grove case) involved the regulator declining to participate in a process because the report into an earlier accident was still pending, having been delayed for legal reasons. In one respect this is quite normal. An organization of this kind should not offer opinion or make judgements when the implications of a recent disaster remain unclear. Yet in another respect this kind of propriety is bizarre – leaving the system in question in a condition that a previous disaster has already proved to be inadequate. Another instance (in the Clapham Junction case) involved managers being reluctant to veto a plan that had become outdated. Installation work on the signalling system had been planned a long time in advance, and changes in circumstance had made the plan inapplicable. Yet none of the managers involved seemed to think it appropriate for them to challenge the plan. The issue is perhaps related to notions of 'model 1' theories-in-use (Argyris, 1992) and the over-emphasis on avoiding embarrassment. But it suggests how, in a complex setting, behaviour is almost bound to be regulated by general norms of conduct. Of course if one knew in advance that this conduct would contribute to a failure on a particular occasion one would suspend the conduct, but since this generally will not be known in advance one would have to suspend it indefinitely.

Speciousness

The 'speciousness' category collected together phenomena related to the use of specious rationalizations in support of actions that were unfounded in any appropriate logic. This is a contentious category from a methodological standpoint because the analysis is based on the judgment in hindsight that an argument was specious at the time it was made. This means what we have to say is tentative, but it emerged from the analysis as an important issue.

Some of the speciousness effects appeared simply to involve self-serving mistruths, but others were more subtle, such as self-serving attributions of causality. For example, signallers attributed complete responsibility for SPADs (signals passed at danger) to drivers, and this let the signallers avoid taking any concern in SPADs – even when

SPADs are plainly near-misses that might presage future disaster. Such attributions naturally present a strong barrier to Weick and Roberts' (1993) heedful inter-relating. There was a telling fragment of testimony quoted in the inquiry report (Cullen, 2000, Section 6.34):

Chair: Was that the overall feeling? If a SPAD happens, that is essentially a driver problem?

Witness: Well how can it not be? If a person at a traffic light goes through a red light, is it not his fault?

Given that, at the time, there were several working groups trying to get to grips with the systematic problem of signal visibility, it is hard not to regard this as a specious argument.

Another 'speciousness' effect involved blame in a different way. The regulator had, prior to Ladbroke Grove, encouraged the development of a 'no-blame culture' in which people were not discouraged from revealing near-misses because they were afraid of being blamed for their role in them. A side-effect of this blamelessness was that when SPADs occurred, and causation was attributed to drivers, the drivers supposedly accepted this attribution since it was not accompanied by blame. Because the attribution was not contested there was apparently no attempt to reveal the genuine, systemic causes of SPADs. There is a sense in which – in a 'no blame culture' – people might collude to accept simplistic attributions because they can then avoid the need to change the system. These attributions are specious, and perhaps most people know it, but in the absence of blame there is apparently little at stake.

DISCUSSION

The results presented in the last section were organized around two broad categories: of what the organization was attempting to do, which became problematic; and of what was making this problematic. In Figure 2 we have suggested there are therefore two co-existing elements at work – 'requisite processes' and 'limiting processes' – and we have used the lower level categories to add detail to these. Thus, within 'requisite processes', engendering is needed to distribute the values and capacities required for safety; instantiation is needed to generate specific plans that deal with the specific threats in question; mobilization is needed to initiate and provide resources to corrective activity; and continuity is needed to sustain this activity over time as other contingencies arise in the organization. On the other hand, within the 'limiting processes', are processes of disarticulation that cut across sustained corrective activity, of lag that prevent such activity being rapidly efficacious, processes of propriety that inhibit organizational members from following through into conclusive outcomes, and of speciousness – allowing actors to justify the ultimate failure to correct the system.

It also seems to be the case that processes of reform rather than processes of operation are primarily implicated – that is, processes acting on the system that fails catastrophically, rather than acting within it. The central problem at Ladbroke Grove was the inability to reform a signal and track layout that was known to be highly

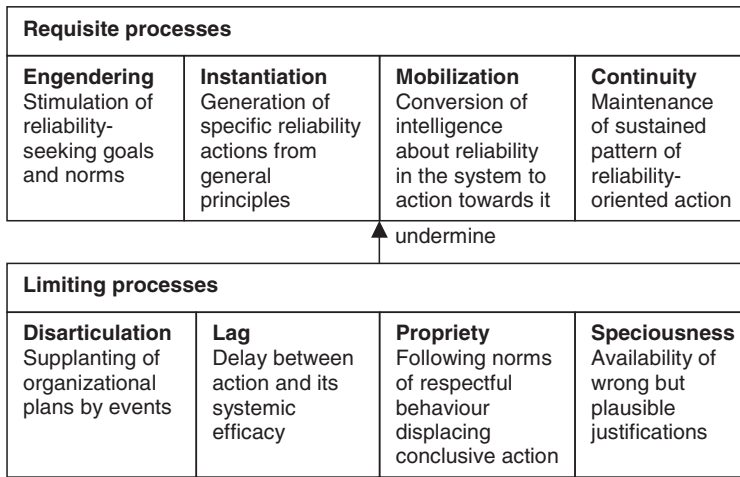


Figure 2. Co-existing requisite and limiting processes

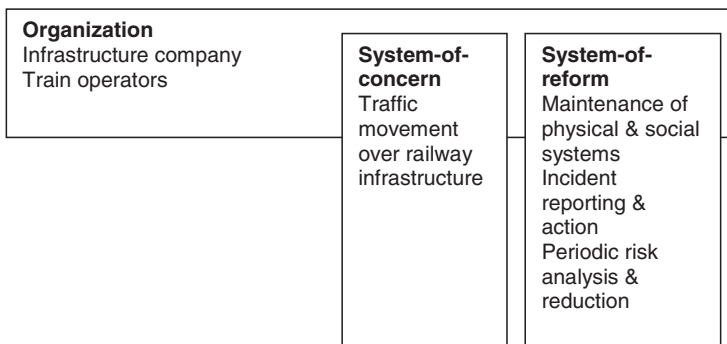


Figure 3. Organizations in and around systems

vulnerable. The central problem at Clapham Junction was the inability to reform a set of work and verification practices known to be deficient. And the specific phenomena collected under our categories reflect this emphasis on reforming rather than operating a system – from the proliferation of working groups under ‘mobilization’ to the neglect of systemic explanations of error under ‘speciousness’. Figure 3 is intended to show the broad distinction between the system that is the object of reform – the system that fails catastrophically – and the system in which processes of reform reside. The suggestion is that organizations exist in and around two main systems of this kind, and our results indicate the second is at least as important as the first as a unit of analysis.

The importance and problematic quality of maintaining and reforming a system – and its central place in organizing around risk – are perhaps self-evident in a number of respects. First, organizational systems cannot be unvarying, even when perfectly reliable, since they must respond to continual pressures on production, new technological opportunities, shifting expectations and so on. There is a danger, in reasoning

about the capacity to manage risk, that this is seen as an asymptotic progression towards some static optimum. Second, reform can be effortful, uncomfortable, disrupting of routines, and lacking in routine itself, yet organizations must simultaneously maintain reliability while doing it. A railway cannot stop working during processes of reform. Third, reform involves a more-or-less public admission that a previous set of organizing assumptions had been inadequate and possibly imperilled reliability in the past. This must sometimes be uncomfortable, and the prospect of blame for a prior state of affairs creates an incentive to minimize the evident discrepancies between the systems before and after reform. Fourth, the effectiveness of reform is typically problematic: once effective it stops and the people engaged in it have to find other things to do; once effective there is often no particular evidence that it is effective (especially if it is geared towards safety – the absence of a notable outcome); and its effectiveness might seem to be only the smallest of advances when the next systemic deficiency requiring correction comes to the surface.

Finally, it seems evident that the limiting processes are themselves products of organization – not phenomena that are in some way independent of it. Disarticulation is the occurrence of unexpected events that only matter because we construct organizations to deal with expected events; lag accompanies efforts that are large in scale because organization itself is large in scale; propriety is a behaviour that seems to be needed to make organization work; and speciousness is a natural by-product of situations in which rationalization is necessary and different rationalizations are plausible. Similarly, we suggested in the Results that effects like the ‘massing’ of investments are chosen ways of organizing that can impede mobilization; that instantiation problems arise because ways of organizing that involve averaging and aggregation; that continuity problems arise because we organize in a way that requires consistency over some temporal or spatial unit; and that engendering becomes a problem when we choose to organize in a way that includes regulation. Figure 4 is a simple model that is intended to reflect all this, proposing that reliability seeking organization tends to induce the processes that limit what can be achieved by reform.

Thus our general thesis is that organizing around risks of catastrophic failure tends to be limited by the conditions that organization itself produces, in ways that we are familiar with from past work; but also that it critically involves processes of systemic reform, which are themselves limited by the conditions produced by organization.

As well as changing the emphasis from processes of operation to those of reform, our analysis indicates other differences in focus from the literature on organizational reliability. Thus, the high reliability organizations theme of migrating authority structures (for example La Porte and Consolini, 1991) seems to be one of transitions between forms that are otherwise stable. Presumably it is vitiated in organizations that suffer disarticulation – that exist in a state of flux – when (for example) they seem unable to keep individual actors in defined roles for any length of time. Similarly, the high reliability organizations principle of strong enculturation (Weick, 1987) is not straightforward because cultural understandings have to be particularized or transformed to be relevant to specific situations, and our study suggests how this process of instantiation can be problematic. At Ladbroke Grove the possibility of derogations to engineering standards had led to a flawed signal configuration in a way that is redolent of Vaughan’s (1996,

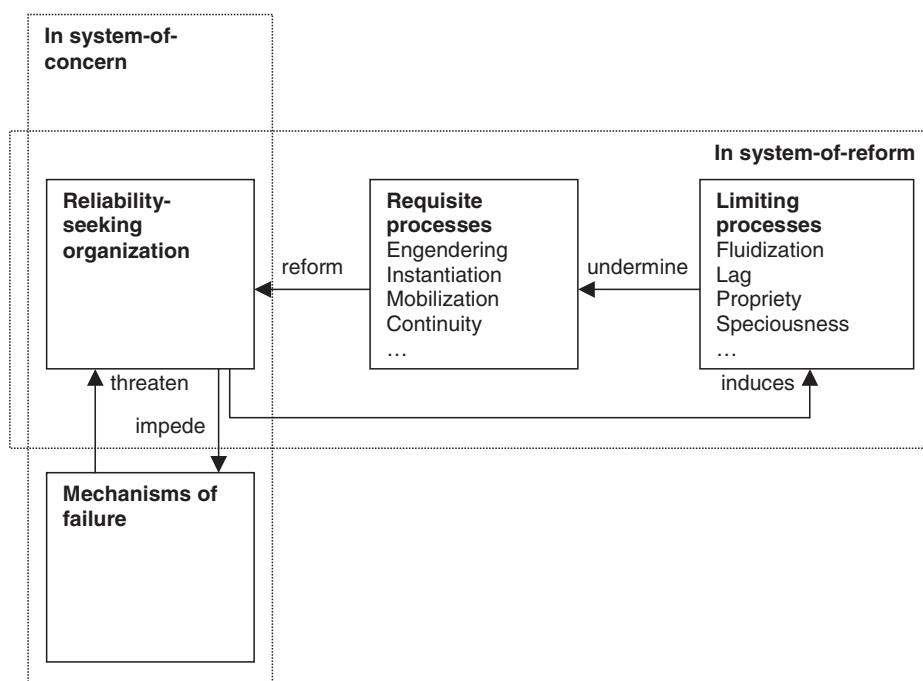


Figure 4. A general model of limited reliability seeking

p. 57) 'normalized deviance'. The problems of engendering – and the transmission of reliability goals from one organization to another – did not seem to occur in La Porte and Thomas's (1995) study of a nuclear power plant as high reliability organization. It appears from their account as though the operator had internalized regulatory demands very effectively. But it also looks as though the operator had been strongly motivated by the prospect of shame, and in particular the shame of the regulatory staff finding some problem before the operator did. Thus the regulatory demand seemed not to be genuinely internalized: it is simply that there was a translation from the standard sanctions of the regulator (such as fines) into more subtle, social sanctions (the shame of being found out). Our findings to do with propriety and speciousness do seem to reflect Weick and Roberts' (1993) advocacy of heedful inter-relating, and Weick's (1993) respectful interaction, and they follow well established themes of rationalization (Staw, 1980). But heedfulness sounds like a counsel of perfection in these settings. It is important to remember that the organizations in our two cases were large bureaucracies, and that people dealing with each other often did not know each other very well, nor interact intensively or frequently. It is implausible that in every case they had sufficient encounters and mutual knowledge to develop the capacity for genuine heedfulness. Individuals were regularly moved from one area to another and one role to another, and often did not collaborate in a tightly-bounded environment, face-to-face, over extended periods of time. In such cases, heedfulness could well fall back to a lesser behaviour in which general norms of propriety become the guide, and in which justifiability is a necessary proxy for effectiveness.

CONCLUSION

This study has been an attempt to understand reliability seeking. Whereas prior work on organizational reliability has looked at the accomplishment of 'high reliability' organizations, this study analysed what undermined the reliability-seeking efforts of organizations that could not, ultimately, achieve sufficient reliability. Our conclusion from two case studies is that organizing around risks of catastrophic failure critically involves processes of systemic reform whose efficacy is limited by conditions that organizing itself tends to produce. Thus – for example – continuity over time is necessary for reform but is subject to various kinds of disarticulation. Mobilization is another necessary element for reform but is undermined by the kind of propriety that is necessary in complex organizations whose members often cannot develop extensive experiential knowledge about one another. There is an implication, perhaps, that future work in this area could profit from further studies of organizations whose reliability is marginal – rather than definitively good enough or clearly inadequate – since the subtleties of what limits reliability seeking are likely to be most evident in such settings.

We also hope to have contributed to knowledge about how organizations do and should manage risk at a more detailed level. An important theme was the way in which specious justifications arose in the choices various actors made. We argued that this can be a benign phenomenon in settings where averaged outcomes are sufficient achievements, but that it is hazardous in settings where a single error can trigger a large scale catastrophic failure. This suggests that organization around risk as a central construct should in part be about finding ways of detecting specious justifications and avoiding organizational members colluding in them in order to avoid difficult reflections on the systems being operated. One of the instructive examples in the cases analysed here was how – in an organization that had laudably moved towards having a 'no-blame' culture – a collective acceptance arose of a partial and inadequate explanation of incidents because little seemed to be at stake and no-one felt the need to mount a vigorous defence of themselves, precisely because they were not blamed.

A case study of this kind is naturally limited in what claims can be made for its generality. Rather than claiming risk management in organizations is *all* about systemic reform and the management of speciousness, then, we are claiming an existence proof: that problematic reform and phenomena like speciousness exist in *some* organizations, *some* of the time. And they are sufficiently important to be implicated in two of the worst events recently seen on the UK railway.

NOTE

*Many thanks are due to the reviewers of an earlier version of this paper for their observations and suggestions.

REFERENCES

- Argyris, C. (1992). *On Organizational Learning*. Cambridge, MA: Blackwell.
Bain, W. A. (1999). 'Application of theory of action to safety management: recasting the NAT/HRT debate'. *Journal of Contingencies and Crisis Management*, 7, 129–40.

- Beck, U. (1992). *Risk Society: Towards a New Modernity*. London: Sage.
- Bierly, P. E. and Spender, J. C. (1995). 'Culture and high reliability organizations: the case of the nuclear submarine'. *Journal of Management*, **21**, 639–56.
- Boden, D. (1994). *The Business of Talk: Organizations in Action*. Cambridge: Polity Press.
- Carroll, J. S. (1998). 'Organizational learning activities in high-hazard industries: the logics underlying self-analysis'. *Journal of Management Studies*, **35**, 699–717.
- Cullen, The Rt Hon. Lord (2000). *The Ladbroke Grove Rail Inquiry*. London: Health and Safety Executive.
- Douglas, M. and Wildavsky, A. (1982). *Risk and Culture. An Essay on the Selection of Technological and Environmental Dangers*. Berkeley, CA: University of California Press.
- Elliott, D. and Smith, D. (2006). 'Cultural Readjustment After Crisis: Regulation and learning from crisis within the UK soccer industry'. *Journal of Management Studies*, **43**, 2, 289–318.
- Freudenberg, W. R. (2003). 'Institutional failure and the organizational amplification of risk: the need for a closer look'. In Pidgeon, N., Kasperson, R. E. and Slovic, P. (Eds), *The Social Amplification of Risk*. Cambridge: Cambridge University Press, 102–20.
- Gephart, R. P. (1984). 'Making sense of organizationally based environmental disasters'. *Journal of Management*, **10**, 205–25.
- Gephart, R. P. (1993). 'The textual approach: risk and blame in disaster sensemaking'. *Academy of Management Journal*, **36**, 1465–514.
- Grabowski, M. and Roberts, K. H. (1996). 'Human and organizational error in large scale systems'. *IEEE Transactions on Systems, Man, and Cybernetics, Part A*, **26**, 2–16.
- Hidden, A. (1989). *Investigation Into the Clapham Junction Railway Accident*. London: The Stationery Office.
- Hollnagel, E. (1993). *Human Reliability Analysis: Context and Control*. London: Academic Press.
- Hopkins, A. (1999). 'The limits of normal accident theory'. *Safety Science*, **32**, 93–102.
- Jarman, A. (2001). '“Reliability” reconsidered: a critique of the HRO-NAT debate'. *Journal of Contingencies and Crisis Management*, **9**, 98–106.
- Klein, R. L., Bigley, G. A. and Roberts, K. H. (1995). 'Organizational culture in high reliability organizations: an extension'. *Human Relations*, **48**, 771–93.
- La Porte, T. R. (1996). 'High reliability organizations: unlikely, demanding and at risk'. *Journal of Contingencies and Crisis Management*, **4**, 60–71.
- La Porte, T. R. and Consolini, P. M. (1991). 'Working in practice but not in theory: theoretical challenges of high-reliability organizations'. *Journal of Public Administration Research and Theory*, **1**, 19–47.
- La Porte, T. R. and Thomas, C. W. (1995). 'Regulatory compliance and the ethos of quality enhancement: surprises in nuclear power plant operations'. *Journal of Public Administration Research and Theory*, **5**, 109–37.
- Marcus, A. A. and Nichols, M. L. (1999). 'On the edge: heeding the warnings of unusual events'. *Organization Science*, **10**, 482–99.
- Meszaros, J. R. (1999). 'Preventive choices: organizations' heuristics, decision processes and catastrophic risks'. *Journal of Management Studies*, **36**, 977–97.
- Perrow, C. (1984). *Normal Accidents*. New York: Basic Books.
- Perrow, C. (1999). *Normal Accidents*, reprint. Princeton, NJ: Princeton University Press.
- Pidgeon, N. (1997). 'The limits to safety? Culture, politics, learning and man-made disasters?'. *Journal of Contingencies and Crisis Management*, **5**, 1–14.
- Pidgeon, N. and O'Leary, M. (2000). 'Man-made disasters: why technology and organizations (sometimes) fail'. *Safety Science*, **34**, 15–30.
- Ramanujam, R. and Goodman, P. S. (2003). 'Latent errors and adverse organizational consequences: a conceptualization'. *Journal of Organizational Behavior*, **24**, 815–36.
- Rasmussen, J. (1986). *Information Processing and Human-Machine Interaction*. Amsterdam: North Holland, 153.
- Rasmussen, J. (1990). 'The role of error in organizing behaviour'. *Ergonomics*, **33**, 1185–200.
- Rasmussen, J. (1997). 'Risk management in a dynamic society: a modelling problem'. *Safety Science*, **27**, 183–213.
- Reason, J. (1990). 'The contribution of latent human failures to the breakdown of complex systems'. *Philosophical Transactions of the Royal Society, Part B*, **327**, 475–84.
- Reason, J. (1997). *Managing the Risks of Organizational Accidents*. Aldershot: Ashgate.
- Reason, J. (2000). 'Human error: models and management'. *British Medical Journal*, **320**, 768–70.
- Rijpmma, J. A. (1997). 'Complexity, tight-coupling and reliability: connecting normal accidents theory and high reliability theory'. *Journal of Contingencies and Crisis Management*, **5**, 15–23.
- Rijpmma, J. A. (2003). 'From deadlock to dead end: the normal accidents-high reliability debate revisited'. *Journal of Contingencies and Crisis Management*, **11**, 37–45.

- Roberts, K. H. (1990). 'Some characteristics of one type of high reliability organization'. *Organization Science*, **1**, 160–76.
- Roberts, K. H. and Rousseau, D. M. (1989). 'Research in nearly failure-free, high-reliability organizations: having the bubble'. *IEEE Transactions on Engineering Management*, **36**, 132–9.
- Roberts, K. H., Stout, S. K. and Halpern, J. J. (1994). 'Decision dynamics in two high reliability organizations'. *Management Science*, **40**, 614–24.
- Rudolph, J. W. and Repenning, N. P. (2002). 'Disaster dynamics: understanding the role of quantity in organizational collapse'. *Administrative Science Quarterly*, **47**, 1–30.
- Sagan, S. D. (1993). *The Limits of Safety: Organizations, Accidents and Nuclear Weapons*. Princeton, NJ: Princeton University Press.
- Sagan, S. D. (1994). 'Toward a political theory of organizational reliability'. *Journal of Contingencies and Crisis Management*, **2**, 228–40.
- Schieb, P. A. (2003). 'Emerging systemic risks: a new frontier?'. *Risikogesellschaft Wohin?*. Centre for Global Dialogue Rueschlikon, 28 March.
- Schulman, P. R. (1993). 'The negotiated order of organizational reliability'. *Administration and Society*, **25**, 353–72.
- Shrivastava, P., Mitroff, I. I., Miller, D. and Mighani, A. (1988). 'Understanding industrial crises'. *Journal of Management Studies*, **25**, 285–303.
- Snook, S. A. (2000). *Friendly Fire: The Accidental Shootdown of US Black Hawks Over Northern Iraq*. Princeton, NJ: Princeton University Press.
- Staw, B. M. (1980). 'Rationality and justification in organizational life'. *Research in Organizational Behavior*, **2**, 45–80.
- Taylor, D. (1987). 'The hermeneutics of accidents and safety'. In Rasmussen, J., Duncan, K. and Leplat, J. (Eds), *New Technology and Human Error*. Chichester: Wiley, 31–41.
- Toft, B. and Reynolds, S. (1994). *Learning from Disasters*. Oxford: Butterworth-Heinemann.
- Turner, B. A. (1976). 'The organizational and interorganizational development of disasters'. *Administrative Science Quarterly*, **21**, 378–97.
- Turner, B. A. (1978). *Man Made Disasters*. London: Wykeham.
- Turner, B. A. (1994). 'The future of risk research'. *Journal of Contingencies and Crisis Management*, **2**, 146–56.
- Vaughan, D. (1996). *The Challenger Launch Decision*. Chicago, IL: University of Chicago Press.
- Weick, K. E. (1987). 'Organizational culture as a source of high reliability'. *California Management Review*, **29**, 112–27.
- Weick, K. E. (1988). 'Enacted sensemaking in crisis situations'. *Journal of Management Studies*, **25**, 305–17.
- Weick, K. E. (1990). 'The vulnerable system: an analysis of the Tenerife air disaster'. *Journal of Management*, **16**, 571–93.
- Weick, K. E. (1993). 'The collapse of sensemaking in organizations: the Mann Gulch disaster'. *Administrative Science Quarterly*, **38**, 628–52.
- Weick, K. E. and Roberts, K. H. (1993). 'Collective mind in organizations: heedful interrelating on flight decks'. *Administrative Science Quarterly*, **38**, 357–81.
- Weir, D. (1996). 'Risk and disaster: the role of communications breakdown in plane crashes and business failures'. In Hood, C. and Jones, D. K. C. (Eds), *Accident and Design: Contemporary Debates in Risk Management*. London: University College London Press, 114–26.
- Westrum, R. (1992). 'Cultures with requisite imagination'. In Wise, J. A., Hopkin, V. D. and Stager, P. (Eds), *Verification and Validation of Complex Systems: Human Factors Issues*. Berlin: Springer-Verlag, 401–15.
- Wildavsky, A. (1983). 'Information as an organizational problem'. *Journal of Management Studies*, **20**, 29–40.
- Wildavsky, A. (1988). *Searching for Safety*. New Brunswick: Transaction Publishers.
- Woo, D. M. and Vicente, K. J. (2003). 'Sociotechnical systems, risk management, and public health: comparing the North Battleford and Walbarton outbreaks'. *Reliability Engineering and System Safety*, **80**, 253–69.
- Zohar, D. and Luria, G. (2003). 'Organizational meta-scripts as a source of high reliability: the case of an army armoured brigade'. *Journal of Organizational Behavior*, **24**, 837–59.