RLIT0000467

Phil. Trans. R. Soc. Lond. B. **327**, 475–484 (1990)

Printed in Great Britain

The contribution of latent human failures to the breakdown of complex systems

By J. Reason

Department of Psychology, University of Manchester, Manchester M13 9PL, U.K.

Several recent accidents in complex high-risk technologies had their primary origins in a variety of delayed-action human failures committed long before an emergency state could be recognized. These disasters were due to the adverse conjunction of a large number of causal factors, each one necessary but singly insufficient to achieve the catastrophic outcome. Although the errors and violations of those at the immediate human–system interface often feature large in the post-accident investigations, it is evident that these 'front-line' operators are rarely the principal instigators of system breakdown. Their part is often to provide just those local triggering conditions necessary to manifest systemic weaknesses created by fallible decisions made earlier in the organizational and managerial spheres.

The challenge facing the human reliability community is to find ways of identifying and neutralizing these latent failures before they combine with local triggering events to breach the system's defences. New methods of risk assessment and risk management are needed if we are to achieve any significant improvements in the safety of complex, well-defended, socio-technical systems. This paper distinguishes between active and latent human failures and proposes a general framework for understanding the dynamics of accident causation. It also suggests ways in which current methods of protection may be enhanced, and concludes by discussing the unusual structural features of 'high-reliability' organizations.

1. Introduction

The past few years have seen a succession of major disasters afflicting a wide range of complex technologies: nuclear power plants, chemical installations, spacecraft, 'roll-on-roll-off' ferries, commercial and military aircraft, off-shore oil platforms and railway networks. If we were to focus only upon the surface details, each of these accidents could be regarded as a singular event, unique in its aetiology and consequences. At a more general level, however, these catastrophes are seen to share a number of important features.

(i) They occurred within complex socio-technical systems, most of which possessed elaborate safety devices. That is, these systems required the precise coordination of a large number of human and mechanical elements, and were defended against the uncontrolled release of mass and energy by the deliberate redundancy and diversity of equipment, by automatic shut-down mechanisms and by physical barriers.

(ii) These accidents arose from the adverse conjunction of several diverse causal sequences, each necessary but none sufficient to breach the system's defences by itself. Moreover, a large number of the root causes were present within the system long before the accident sequence was apparent.

(iii) Human rather than technical failures played the dominant roles in all of these accidents. Even when they involved faulty components, it was subsequently judged that appropriate human action could have avoided or mitigated the tragic outcome.

RLIT0000467

Thanks to the abundance and sophistication of engineered safety measures, many high-risk technologies are now largely proof against single failures, either of humans or components. This represents an enormous engineering achievement. But it carries a penalty. The existence of elaborate 'defences in depth' renders the system opaque to those who control it. The availability of cheap computing power (which provided many of these defences) means that, in several modern technologies, human operators are increasingly remote from the processes that they nominally govern. For much of the time, their task entails little more than monitoring the system to ensure that it functions within acceptable limits.

A point has been reached in the development of technology where the greatest dangers stem not so much from the breakdown of a major component or from isolated operator errors, as from the insidious accumulation of delayed-action human failures occurring primarily within the organizational and managerial sectors. These residual problems do not belong exclusively to either the machine or the human domains. They emerge from a complex and as yet little understood interaction between the technical and social aspects of the system.

Such problems can no longer be solved by the application of still more 'engineering fixes' nor are they amenable to the conventional remedies of human factors specialists. Further improvements in reliability will require more effective methods of risk management. These, in turn, depend upon acquiring a better understanding of the breakdown of complex socio-technical systems, and the development of new techniques of risk assessment. This paper sketches out some of the issues that must be confronted if this ambitious programme is to succeed.

2. Active and latent human failures

Close examination of several recent disasters (especially Bhopal, Challenger, Chernobyl, Zeebrugge and King's Cross) shows the need to distinguish two ways in which human beings contribute to the breakdown of complex systems (see also Rasmussen & Pedersen (1984)).

(i) Active failures: those errors and violations having an immediate adverse effect. These are generally associated with the activities of 'front-line' operators: control room personnel, ships' crews, train drivers, signalmen, pilots, air traffic controllers, etc.

(ii) Latent failures: these are decisions or actions, the damaging consequences of which may lie dormant for a long time, only becoming evident when they combine with local triggering factors (that is, active failures, technical faults, atypical system conditions, etc.) to breach the system's defences. Their defining feature is that they were present within the system well before the onset of a recognizable accident sequence. They are most likely to be spawned by those whose activities are removed in both time and space from the direct human–machine interface: designers, high-level decision makers, regulators, managers and maintenance staff.

Two recent accident investigations, in particular, have dramatically reversed the usual practice of focusing upon the actions of the 'front-line' operators (Sheen, 1987; Fennell 1988). Both the Zeebrugge and King's Cross inquiries concluded that rather than being the main instigators of these disasters, those at the human–machine interface were the inheritors of system defects created by poor design, conflicting goals, defective organization and bad management decisions. Their part, in effect, was simply that of creating the conditions under which these latent failures could reveal themselves.

There is a growing awareness within the human reliability community that attempts to discover and remedy these latent failures will achieve greater safety benefits than will localized

RLIT0000467

efforts to minimize active failures. So far, much of the work of human factors specialists has focused upon improving the immediate human–system interface. Whereas this is undeniably an important enterprise, it only addresses a relatively small part of the total safety problem, being aimed at reducing the active failure tip of the causal iceberg. The remainder of this paper will focus upon latent rather than active failures, beginning with some quantitative evidence from the nuclear power industry.

3. Some data in support of the latent failure argument

The Institute of Nuclear Power Operations (INPO) manages the Significant Event and Information Network for its member utilities both within and outside the United States. In 1985 they issued an analysis of 180 significant event reports received in 1983–84 (INPO 1985). A total of 387 root causes were identified. These were assigned to five main categories: human performance problems, 52%; design deficiences, 33%; manufacturing deficiences, 7%; external causes, 3%; and an 'other unknown' category, 5%.

The human performance problems were further broken down into the following sub-categories: deficient procedures or documentation, 43%; lack of knowledge or training, 18%; failure to follow procedures, 16%; deficient planning or scheduling, 10%; miscommunication, 6%; deficient supervision, 3%; policy problems, 2%; and 'other', 2%.

There are two important conclusions to be drawn from these data. First, at least 92% of all root causes were man-made. Secondly, only a relatively small proportion of the root causes (approximately 8% of the total) were initiated by the operators. The majority had their origins in either maintenance-related activities, or in fallible decisions taken within the organizational and managerial domains.

The major role played by maintenance-related errors in causing nuclear power plant events has also been established by two independent studies (Rasmussen 1980; NUMARC 1985). Of these, simple omissions (the failure to carry out necessary actions) formed the largest single category of identified human problems in nuclear power plant operations.

4. A resident pathogen metaphor

It is suggested that latent failures are analogous to the 'resident pathogens' within the human body, which combine with external factors (stress, toxic agencies, etc.) to bring about disease. Like cancers and cardiovascular disorders, accidents in complex, defended systems do not arise from single causes. They occur through the unforeseen (and often unforeseeable) concatenation of several distinct factors, each one necessary but singly insufficient to cause the catastrophic breakdown. This view leads to a number of general assumptions about accident causation.

(i) The likelihood of an accident is a function of the total number of pathogens (or latent failures) resident within the system. All systems have a certain number. But the more abundant they are, the greater is the probability that a given set of pathogens will meet just those local triggers necessary to complete an accident sequence.

(ii) The more complex, interactive, tightly coupled and opaque the system (Perrow 1984), the greater will be the number of resident pathogens. However, it is likely that simpler systems will require fewer pathogens to bring about an accident as they have fewer defences.

RLIT0000467

(iii) The higher an individual's position within an organization, the greater is his or her opportunity for generating pathogens.

(iv) It is virtually impossible to foresee all the local triggers, though some could and should be anticipated. Resident pathogens, on the other hand, can be assessed, given adequate access and system knowledge.

(v) It therefore follows that the efforts of safety specialists could be directed more profitably towards the proactive identification and neutralization of latent failures, rather than at the prevention of active failures, as they have largely been in the past.

These assumptions raise some further questions: how can we best gauge the 'morbidity' of high-risk systems? Do systems have general indicators, comparable to a white cell count or a blood pressure reading, from which it is possible to gain some snapshot impression of their overall state of health?

5. A GENERAL FRAMEWORK FOR ACCIDENT CAUSATION

The resident pathogen metaphor is far from being a workable theory. Its terms are still unacceptably vague. Moreover, it shares a number of features with the now largely discredited accident proneness theory, although the pathogen view operates at a systemic rather than at an individual level.

Accident proneness theory floundered when it was established that unequal accident liability was, in reality, a 'club' with a rapidly changing membership (see Reason 1974). In addition, attempts to find a clearly defined accident-prone personality proved largely fruitless.

The pathogen metaphor would suffer a similar fate if it turned out that latent failures could only be identified retrospectively in relation to a specific set of accident circumstances in a particular system. For the analogy to have any value, it is necessary to establish a generic set of indicators relating to system 'morbidity', and then to demonstrate clear connections between these indicators and accident liability across a wide range of complex systems and in a variety of accident conditions.

In what follows, an attempt will be made to develop the pathogen metaphor into a theoretical framework for considering the aetiology of accidents in complex technological systems. The challenge is not just to provide an account of how active and latent failures combine to produce accidents, but also to show where and how more effective remedial measures might be applied.

Before considering the pathology of complex systems, we must first identify their essential, 'healthy' components. These are the basic elements of production. All complex technologies are involved in some form of production, whether it be energy, a chemical substance, or the mass transportation of people by land, sea or air. There are five basic elements to any productive system: decision makers, line management, preconditions, productive activities and defences.

(i) Decision makers. These include both the architects and the senior executives of the system. Once in operation, the latter set the production and safety goals for the system as a whole. They also direct, at a strategic level, the means by which these goals should be met. A large part of their function is concerned with the allocation of finite resources. These comprise money, equipment, people and time.

(ii) Line management. These are the departmental specialists who implement the strategies

RLIT0000467

of the decision makers within their particular spheres of operation: operations, training, sales, maintenance, finance, safety, engineering support, personnel, and so on.

(iii) Preconditions. Effective production requires more than just machines and people. The equipment must be reliable and of the right kind. The workforce must be skilled, alert, knowledgeable and motivated.

(iv) Productive activities. These are the actual performances of machines and people: the temporal and spatial coordination of mechanical and human activities needed to deliver the right product at the right time.

(v) Defences. Where the productive activities involve exposure to hazards, both the human and mechanical components of the system need to be provided with safeguards sufficient to prevent forseeable injury, damage or costly outages.

The human contributions to accidents are summarized in figure 1. They are linked there to each of the basic elements of production, portrayed as 'planes' lying one behind the other in an ordered sequence. The question at issue is: how do fallible decisions translate into unsafe acts capable of breaching the system's defences?

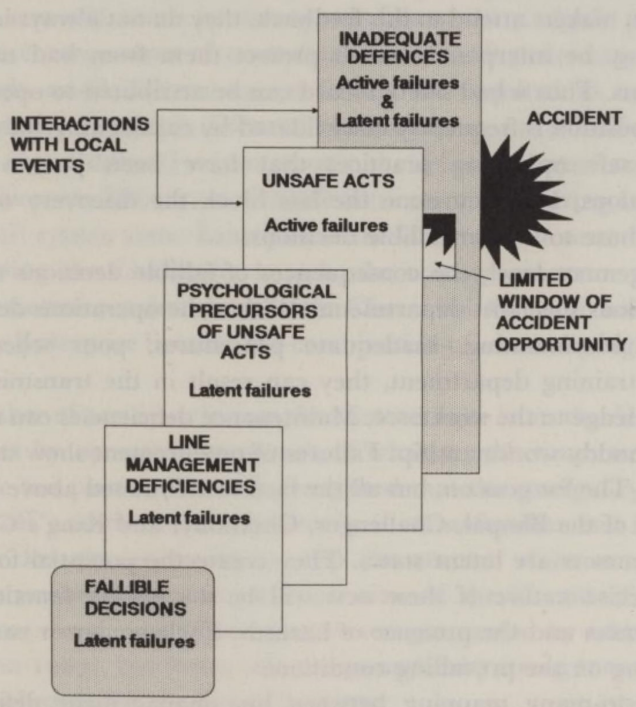

FIGURE 1. Showing the relationship between the various human contributions to accidents and the basic elements of production. Latent failures have their primary systemic origins in the fallible decisions made by senior executives. They are subsequently translated into different forms as the effects of these decisions pass through the system during the production process.

It is assumed that latent failures (resident pathogens) have their primary systemic origin in the errors of high-level decision makers. But they are also introduced into all levels of the system by the human condition. Error proneness and the capacities for being stressed, failing to perceive hazards, being ignorant of the system, and having less than ideal motivation are brought by each individual into the workplace.

RLIT0000467

Even in the best run organizations, a significant number of influential decisions will subsequently prove to be mistaken. Fallible decisions are an inevitable part of the design and management process. The issue is not so much how to prevent them, but how to ensure that their adverse consequences are detected and recovered.

All organizations must allocate limited resources to two distinct goals: production and safety. In the long term, these are clearly compatible. But short-term conflicts of interest will arise in which the resources given to production could diminish safety, and conversely. There are a number of reasons why these dilemmas will tend to be resolved in favour of production rather than safety goals.

First, resources directed at improving productivity have relatively certain outcomes; those aimed at enhancing safety do not, at least in the short term (Brehmer, 1988). This is due to the large part played by stochastic factors in accident causation.

Secondly, the feedback generated by the pursuit of production goals is generally unambiguous, rapid, compelling and (when the news is good) highly reinforcing. In sharp contrast, that derived from the pursuit of safety goals is largely negative, intermittent, often deceptive and perhaps only compelling after a major accident or a string of incidents.

Even when decision makers attend to this feedback, they do not always interpret it correctly. Defensive 'filters' may be interposed, which protect them from bad news and encourage extrapunitive reactions. Thus a bad safety record can be attributed to operator carelessness or incompetence. This position is frequently consolidated by cataloguing the various engineering safety devices and safe operating practices that have been put in place. These are understandable reactions, but they none the less block the discovery of effective counter-measured and contribute to further fallible decisions.

At the line management level, the consequences of fallible decisions manifest themselves differently in the various specialist departments. Within the operations department, they can take the form of undermanning, inadequate procedures, poor scheduling and unsafe assignments. In the training department, they can result in the transmission of inadequate skills, rules and knowledge to the workforce. Maintenance deficiencies can reveal themselves in poor planning and shoddy workmanship. Failures of procurement show up as dangerous and defective equipment. The list goes on, but all the factors mentioned above played a significant part in the aetiology of the Bhopal, Challenger, Chernobyl and King's Cross disasters.

Psychological precursors are latent states. They create the potential for a wide variety of unsafe acts. The precise nature of these acts will be a complex function of the task, the environmental influences and the presence of hazards. Each precursor can give rise to many unsafe acts, depending on the prevailing conditions.

There is a many-to-many mapping between line management deficiencies and these psychological precursors. Failures in the training department, for example, can translate into a variety of precursors: high workload, undue time pressure, inappropriate preception of hazards, ignorance of the system and motivational difficulties. Likewise, any one precondition (for example, undue time pressure) could be the product of many line management deficiencies: poor sheduling, inadequate procedures, inappropriate training and maintenance failures.

A useful way of thinking about these transformations is as types converting into tokens. Deficient training is a general failure type that can reveal itself, at the precursor level, as a variety of pathogenic tokens. Such a view has important remedial implications. Rectifying a

RLIT0000467

failure type could, in principle, eliminate a large class of tokens. The type–token distinction is a hierarchical one. Precondition tokens at the precursor level become types for the creation of tokens at unsafe act level.

A psychological precursor, either alone or in combination with others, can play a major role in provoking and shaping an almost infinitely large set of unsafe acts. The stochastic character of this onward mapping reveals the futility of 'tokenism': the concentration of remedial efforts upon preventing the recurrence of specific unsafe acts. Although certain of these acts may fall into a recognizable category (for example, failing to wear personal safety equipment) and so be amenable to targeted safety programmes, the vast majority of them are unforeseeable and occasionally quite bizarre.

This view of accident causation suggests that unsafe acts are best reduced by eliminating their psychological precursors rather than the acts themselves. However, it must be accepted that whatever measures are taken, some unsafe acts will still occur. It is therefore necessary to provide a variety of defences to intervene between the act and its adverse consequences. Such defences can be both physical and psychological. The latter are as yet relatively unexploited, and involve procedures designed to improve error detection and recovery.

Very few unsafe acts will result in damage or injury. In a highly protected system, the probability that the consequences of an isolated action will penetrate the various layers of defence is vanishingly small. Several causal factors are required to create a 'trajectory of opportunity' through these multiple defences. Many of the causal contributions will come from latent failures in the organizational structure, or in the defences themselves. Others will be local triggering factors. These could be a set of unsafe acts committed during some atypical (but not necessarily abnormal) system state. Examples of the latter are the unusually low temperature on the night preceding the Challenger launch, the voltage–generator tests carried out just before the annual maintenance shut-down in the Chernobyl-4 reactor, and the nose-down trim of the Herald of Free Enterprise because of a combination of unusually high tide and unsuitable docking facilities.

A significant number of accidents in complex systems arise from the deliberate or unwitting disabling of defences by operators in pursuit of what, at the time, seem to be sensible or necessary goals. The test plan at Chernobyl required that the emergency core cooling system should be switched off, and the need to improvise in an unfamiliar and increasingly unstable power regime later led the operators to strip the reactor of its remaining defences. At Zeebrugge, the overworked and undermanned crew of the Herald of Free Enterprise left harbour with the bow doors open. This was an oversight caused by a bizarre combination of active failures (Sheen 1987), but it was also compounded by strong management pressures to meet the stringent schedule for the Dover docking.

6. Managing safer operations

An effective safety information system has been found to rank second only to top management concern with safety in discriminating between safe and unsafe companies, matched on other variables (Kjellen 1983). The feedback loops and indicators that could go to make up such a system are shown in figure 2.

Loop 1 (reporting accidents, lost time injuries, etc.) represents the minimum requirement.

RLIT0000467

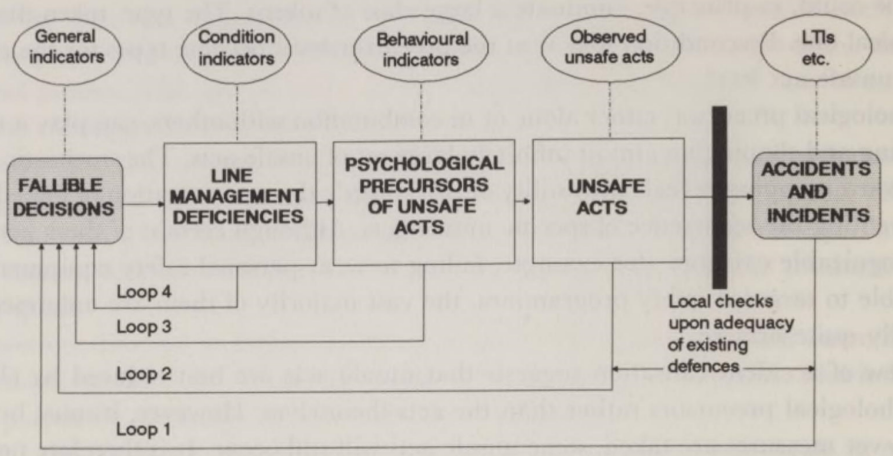

FIGURE 2. The actual and potential feedback loops and indicators associated with each of the basic elements of production. Loop 1 already exists in most systems, and communicates information about accidents, lost time injuries, etc. Loop 2 is potentially available through Unsafe Act Auditing. Loops 3 and 4 could provide information regarding latent failures; though, in practice they are rarely seen in place.

In most cases, however, the information supplied is too little and too late for effective proactive control. The events that safety management seeks to prevent have already occurred.

Loop 2 is potentially available through unsafe act auditing procedures. In practice, however, this information is usually only disseminated to the lower, supervisory levels of the organization.

The main thrust of the present view of accident causation is towards the establishment of loops 3 and 4. It argues that the most effective way of managing safety is by acting upon types rather than tokens; that is, by influencing system and individual states occurring early on in the history of a possible accident. To identify these indicators, and to find ways of neutralizing the general failure types so revealed constitute the major challenges facing contemporary accident researchers. For the moment, however, we will consider only the most global of these diagnostic signs: the general indicators associated with top-level decision making.

The general indicators shown in figure 2 cover two broad aspects of a system's safety management. The first relates to the variety and sensitivity of its feedback loops. The second deals with the senior executives' responses to safety-related data. No amount of feedback will enhance system safety if the information supplied is not acted upon in a timely and effective manner.

Westrum (1988) has provided a useful classification of the ways in which organizations differ in their responses to safety-related information. These reactions fall into three groups: denial, repair and reform actions.

(i) Denial actions. These may take one or both of the following forms: suppression, 'whistleblowers' are punished or dismissed and their observations removed from the record; and encapsulation, the observers are retained, but the validity of their observation is disputed or denied.

(ii) Repair actions. Externally, these can take the form of a public relations exercise in which the observations are allowed to emerge, but in a reassuring and sugar-coated manner. Internally, the problem is admitted, but it is only addressed at a local level. 'Offending' operators are disciplined or relocated. Dangerous items of equipment are modified to prevent

RLIT0000467

the recurrence of a specific kind of observed failure. The wider implications of the problem are denied.

(iii) Reform actions. These take two forms: dissemination, the problem is admitted to be global, and global action is taken upon it; reorganization, action on the problem leads to a fundamental reappraisal and reform of the system as a whole.

The more effective the organization, the more likely it is to respond to safety data with reform actions. Less adequate organizations will confine themselves to either denial or repair actions. This leads to a tripartite classification of organizations.

(i) Pathological organizations possess inadequate safety measures, even under normal circumstances. They habitually sacrifice safety for greater productivity, often under severe economic pressure, and they actively circumvent safety regulations.

(ii) Calculative organizations try to do the best job they can by using 'by-the-book' methods. These are usually adequate under normal operating circumstances, but often fail to thwart the development of the multiple-cause accidents, discussed earlier.

(iii) Generative organizations set safety targets for themselves beyond ordinary expectations, and fulfill them because they are willing to do unusual things in unconventional ways. They emphasise results more than methods, and value substance more than form. Notable exemplars of this last category have been investigated by La Porte and his colleagues at the University of California, Berkeley.

7. High-reliability organizations

The La Porte group (see La Porte & Consolini 1989) has made an intensive study of three high-reliability organizations: the Federal Aviation Authority's air traffic control system, Pacific Gas and Electric's power generating system and two U.S. Navy nuclear aircraft carriers. These organizations share at least two goals: to avoid altogether major failures that could cripple or even destroy, the system; and to cope safely with periods of very high peak demand and production whenever they arise. All of these organizations perform complex and hazardous tasks under considerable time pressure, and they do so with a very low error rate and an almost total absence of catastrophic failure. What are the ingredients? Can we 'bottle' them?

Perhaps the most significant feature of these organizations is their complex yet highly adaptive structural reactions to changing levels of hazard. Each organization has three distinct authority structures: routine, high-tempo and emergency. Each structure has its own characteristic practices, communication pathways and leadership patterns.

The routine mode reveals the familiar hierarchical pattern of rank-dependent authority. This is the face of the organization most evident to the casual observer. It functions with the use of extensive 'standard operating procedures'.

Just beneath the surface of this bureaucratic structure is the high-tempo mode, practised by the same individuals, but in quite a different manner. Authority is no longer based upon rank, but upon functional skills. Formal status defers to expertise. Communications switch from largely vertical channels to being richly horizontal among task-related groups.

Within these high-tempo groups, the La Porte team noted an extraordinary sensitivity to the incipient overloading of any one of its members. For example, when an air traffic controller has an unusually large number of aircraft on his screen, supervisors and other controllers will

RLIT0000467

gather around silently and watch for danger points. When found, they are shown by pointing at the screen. Few words are spoken. When the load has eased, the impromptu support group fades away as quietly as it arrived.

The emergency mode is triggered by unequivocal signs that a well-defined danger is imminent. Authority patterns in this mode are based upon a preprogrammed and well-rehearsed allocation of duties. Individuals regroup themselves into different functional units on the basis of a predetermined plan, tailored to the particular nature of the emergency. La Porte and Consolini (1989) comment upon these co-existing structures as follows: 'Contemporary organization theory literature does little to alert one to the likelihood of these multi-layered, nested authority systems. We are familiar with different types of organization that parallel each of these modes. There are bureaucratic organizations, professional ones, and disaster response ones. We have not thought that all three might be usable by the same organizational membership.'

Can we build these adaptive structural ingredients into high-risk organizations at their outset, or must they evolve painfully and serendipitously over many years of hazardous operating experience? It is probably too early to tell. But it is clear that a close study of high-reliability organizations should feature prominently on the research agendas of those concerned with understanding and preventing the kinds of disaster discusssed in this paper. Just as in medicine, it is probably easier to characterize sick systems rather than healthy ones. Yet we need to pursue both of these goals concurrently if we are to understand and then create the organizational bases of system reliability.

Many people contributed to the ideas expressed in this paper. Two, in particular, deserve special mention. Patrick Hudson of the University of Leiden was the first to apply the type–token distinction in the context of accident causation. The production-based framework for accident causation emerged from discussions with John Wreathall of Science Applications International Corporation (Columbus, Ohio).

References

Brehmer, B. 1988 Changing decisions about safety in organizations. World Bank Workshop on Safety Control and Risk Management, 18–20 October, Washington, D.C.

Fennell, D. 1988 *Investigation into the King's Cross underground fire*. Department of Transport, London: HMSO.

INPO 1985 *An analysis of root causes in 1983 and 1984 significant event reports*. Atlanta, Georgia: Institute of Nuclear Power Operations.

Kjellen, U. 193 *The Deviation Concept in Occupational Accident Control, TRITA/AOG-0019, Arbetsolycksfalls Gruppen*. Stockholm: Royal Institute of Technology.

La Porte, T. R. & Consolini, P. M. 1988 Working in practice but not in theory: theoretical challenges of high reliability organizations. Annual Meeting of the American Political Science Association, 1–4 September, Washington, D.C.

Numarc 1985 *A maintenance analysis of safety significant events*. Nuclear Utility Management and Human Resources Committee. Atlanta, GA: Institute of Nuclear Power Operations.

Perrow, C. 1984 *Normal accidents: living with high risk technologies*. New York: Basic Books.

Rasmussen, J. & Pedersen, O. M. 1983 Human factors in probabilistic risk analysis and risk management. In *Operational safety of nuclear power plants*, vol. 1. Vienna: International Atomic Energy Agency.

Rasmussen, J. 1980 What can be learned from human error reports. In *Changes in working life* (ed. K. Duncan, M. Gruneberg & D. Wallis). London: Wiley.

Reason, J. T. 1974 *Man in motion*. London: Weidenfeld.

Sheen, Mr Justice. 1987 *Herald of Free Enterprise*. Report of Court no. 8074. Department of Transport. London: HMSO.

Westrum, R. Organizational and inter-organizational thought. World Bank Workshop on Safety Control and Risk Management, 18–20 October, Washington, D.C.