Chapter 16

Conclusions

This book provides a broad survey of incident reporting techniques. This focus is justified because national and international reporting schemes have recently been established to support the aviation industry [310], chemical production [162], marine transportation [389], military acquisition [289] and operations [807], nuclear power production [384], the rail industries [666], healthcare [105]. Chapter 1.3 has presented a wide range of reasons why incident reporting systems are being used to support the operation of safety-critical applications. For example, incident reports help to find out why accidents do not occur. Many incident reporting forms identify the barriers that prevent adverse situations from developing into a major accident. These insights are very important. They help analysts to identify where additional support is required in order to guarantee the future benefits of those safeguards. The higher frequency of incidents permits quantitative analysis. It can be argued that many accidents stem from atypical situations. They, therefore, provide relatively little information about the nature of future failures. In contrast, the higher frequency of incidents provides greater insights into the relative proportions of particular classes of human 'error', systems failure, regulatory weakness etc. Incident reporting systems also provide a reminder of hazards. Incident reports provide a means of monitoring potential problems as they recur during the lifetime of an application. The documentation of these problems increases the likelihood that recurrent failures will be noticed and acted upon. Reporting systems can also provide feedback that keeps staff 'in the loop'. Incident reporting schemes provide a means of encouraging staff participation in safety improvement. In a well-run system, they can see that their concerns are treated seriously and are acted upon by the organisation. Greater insight into national and global safety issues can be gained.

A further argument in favour of incident reporting schemes is that they can encourage the sharing of safety-related lessons and data. Incident reporting systems provide the raw data for comparisons both within and between industries. If common causes of incidents can be observed then, it is argued, common solutions can be found. However, in practice, the lack of national and international standards for incident reporting prevents designers and managers from gaining a clear view of the relative priorities of such safety improvements. Incident reporting schemes are also cheaper than the costs of an accident. These is an argument that the relatively low costs of managing an incident reporting scheme should be offset against the costs of failing to prevent an accident . This is a persuasive argument. However, there is also a concern that punitive damages may be levied if an organisation fails to act upon the causes of an incident that subsequently contribute towards an accident.

Many of these potential benefits have been cited by the politicians and civil servants who are responsible for settig up new schemes. On the 13th June 2000, the UK Health Secretary, Alan Milburn, and the Chief Medical Officer for England, Liam Donaldson, announced the establishment of a centralised reporting facility for adverse incidents across the UK National Health Service (NHS) [105]. The Chief Medical Officer said; "At the moment there is no way of knowing whether the lessons learned from an incident in one part of the NHS are properly shared with the whole health service". The Health Secretary said; "Patients, staff and the public have the right to expect the NHS to learn from its mistakes so we can ensure the alarm bells ring when there are genuine concerns so they can be nipped in the bud". On April 2nd, 2001 the U.S. Department of Transportation announced a reduction in the threshold for reporting hazardous liquid pipeline incidents from 2,100 gallons, or 50 barrels, to just 5 gallons. This extension of the reporting system was intended to 'heighten the quantity, quality, and usefulness of reported accident information' [705]. In 2002, the FAA Administrator Jane Garvey stated that the GAIN international incident reporting initiative "is one of our best hopes for enhancing aviation safety in the next century". David Hinson, the Former FAA Administrator, observed that "the inception of the GAIN concept" was the most "significant accomplishment of my tenure" [658].

Numerous studies have described tools and techniques that are intended to help realise the potential benefits of incident reporting systems [444, 844]. Chapter 14.5 has, however, argues that very few papers analyse the reasons why previous initiatives have failed to yield these safety improvements [411]. This is a significant omission. Some reporting systems only elicit a very small number of contributions. The Royal College of Anaesthetist's recently concluded that self-reporting only retrieves approximately 30% of the incidents that are detected by independent audit [716]. Those submissions that are obtained may only come from particular sections of the workforce [119]. Under-reporting is, however, only one of the problems that complicate the introduction and maintenance of incident reporting systems. For example, finite resources can also be exhausted if schemes elicit too many contributions. The difficult balancing act between encouraging participation and discouraging excess contributions can be illustrated by the FDA's MedWatch program. Clinicians do not have to be convinced that a device has actually caused an incident in order to report; "Causality is not a prerequisite for MedWatch reporting; suspicion that a medical product may be related to a serious event is sufficient reason for a health professional to submit a MedWatch report". Contributors are, however, warned that the system does not encourage 'a report on every adverse event' only serious events [257]. The aspirations for novel reporting systems, cited in previous paragraphs, must be balanced against the experience of many managers who are responsible for maintaining such systems. For example, the following quotation describes the results of an investigation into the processing of safety hazard reports and unsatisfactory condition reports (UCR). This was triggered by a letter that was sent by a Corporal in the Canadian National Defence Forces' Safety Digest publication in which he expressed frustration at the apparent failure of their reporting system:

> "The investigation revealed that, at least four years ago, Cpl Krygsveld's immediate supervisors failed to action or register four legitimate hazard reports that he submitted. Thus, the Unit Ground Safety Officer was not aware and no immediate action was taken. The fact that two supervisors tore up two of Cpl Krygsveld's hazard reports is unacceptable and clearly reflects a disturbing lack of safety consciousness." [138]

Although the investigation report argued that this is an isolated problem within the Canadian Airforce scheme, similar problems have affected many other reporting systems. For instance, Chapter 14.5 has described how the FDA have pioneered the development of Sentinel monitoring systems as a means of addressing the problems of under-reporting and reporting bias in accounts of medical device failure.

This following pages, therefore, provide a critical analysis of recent initiatives to introduce incident reporting systems. In particular, it is argued that the proponents of many reporting schemes over-estimate the impact that they can have upon the operation of complex, safety-critical systems. One reason for this is that many recent proposals show a limited understanding of the tools and techniques that have supported reporting systems in other countries and other industries.

16.1 Human Problems

Many of the problems that affect the introduction and operation of incident reporting systems stem directly from the difficulty of encouraging individuals and teams to participate in the system. This scale of this problem is often under-estimated in the official announcements that are used to launch new reporting schemes. The publicity surrounding the launch of a new reporting system is often intended to help elicit this participation. Unfortunately, such official announcements and press releases are insufficient to sustain most reporting systems.

16.1.1 Reporting Biases

The long-term elicitation of incident reports depend upon a number of factors. In particular, it can depend upon whether or not particular groups in the workforce perceive there to be any potential benefits to them from participating in the scheme. For example, nursing staff contributed about 90% of all of the reports that have been submitted in a local intensive care unit over the last decade. 621 reports were submitted by nurses compared with 77 reports by medical staff [119]. These figures must, however, be interpreted with great caution. It is important to consider the total number of staff who might contribute to such a system. The teams that contribute to this reporting scheme consist of three medical staff, one consultant, and up to eight nurses per shift. The larger number of reports contributed by nursing staff can also be explained in terms of the involvement in, or exposure to, the types of workplace incidents that were solicited under this particular scheme. Nursing staff had the most direct contact with the patients who remain the focus of the reporting system. They then have a proportionately greater opportunity to witness adverse events [119]. Such reporting biases have important consequences. The system may tell us a great deal about the execution of medical procedures. It may, however, tell us relatively little about more complex problems in the planning, coordination and administration of treatment within a department. Such factors are often overlooked by the proponents of reporting systems when they make strong claims about the quality and quantity of information that might be obtained about safety-critical systems.

Automated logging and tracking systems provide means of addressing the problems both of low contribution rates and of biased participation in a reporting system. The proponents of such systems often have an unfortunate way of advocating their introduction; "competent personnel love them, while incompetent personnel loathe them" [223]. Such assessments hide the difficulty and expense that is often involved in interpreting the data provided by automated, incident-monitoring systems. There is also a concern that any data will be used to punish rather than support staff performance through additional training. These concerns have acted as powerful barriers against the introduction of monitoring equipment onto UK trains. Recommendation 9 of the HMRI report into the accident at Watford South Junction advocated the use of these systems to monitor driving technique. In 1999, however, less than 20% of trains carried this equipment [351]. More recently, the action plan to implement the recommendations of the Southall accident report included steps to extend both voluntary incident reporting systems and automated monitoring equipment [319]. This link between voluntary reporting systems and automated monitoring is instructive. The success of reporting systems in aviation is often explained in terms of the pilot's fear that any incident may have been observed and reported by their colleagues or by tracking equipment. Participation in reporting system often provides a limited degree of protection and support in any subsequent investigation. It is less clear whether such a 'no-blame' approach might also be extended to other domains where automated monitoring and confidential reporting systems have been used together. For instance, some initial steps have been taken to use computer-based tools to automatically identify adverse occurrences in healthcare applications [402]. This often leads to ethical problems when failures have a long-term effect on the patient's prognosis.

16.1.2 Blame

The US National Patient Safety Foundation (NPSF) recently argued that the names of individual health workers who file incident reports must not be released to the public or to licensing bodies. They maintained that 'the culture of health care is rife with guilt, blame, and fear, and these are the greatest obstacles to effective reporting systems' [585]. Similarly, the Australian Transportation Safety Bureau (ATSB) exploit a 'no-blame' systemic safety analysis that is intended to encourage a climate 'in which people are prepared to report their errors' [53]. In the UK, the National Occupational Safety and Health Committee has argued that 'too often' there is a tendency to blame the victim rather than search for the causes of adverse events [584]. 'The most important thing' is to establish why they were not prevented rather than focus narrowly on how they happened.

Again, however, many of these statements avoid the moral and practical problems that arise from the maintenance of 'no blame' reporting systems. The Cullen enquiry into the Ladbroke Grove accident acknowledges this complexity when considering the possible extension of no-blame, confidential reporting systems [195]. It was argued that people must be accountable for their actions. A no-blame approach can also, paradoxically, lead to a system in which workers are more ready to accept responsibility for a failure in order to conclude the investigation as rapidly as possible. Instead of no-blame reporting, the aim should be to develop an industry culture 'in which information is communicated without fear of recrimination and blame is attached only where this is justified' [195]. Cullen notes that the UK rail industry is some distance away from this ideal situation. Many large-scale systems, address employees' fear of retribution and blame by developing elaborate legal safeguards to protect potential contributors [59]. Unfortunately, most reporting systems lack the funding and the necessary managerial support to provide this level of assurance. Even when they are present, there can be other workplace factors that dissuade employees from participating in a reporting system. In extreme cases, employees may even neglect medical treatment rather than expose themselves to workplace harassment:

> "FRA has become increasingly aware that many railroad employees fail to disclose their injuries to the railroad or fail to accept reportable treatment from a physician because they wish to avoid potential harassment from management or possible discipline that is sometimes associated with the reporting of such injuries. FRA is also aware that in some instances supervisory personnel and mid-level managers are urged to engage in practices which may undermine or circumvent the reporting of injuries and illnesses." [235]

In the medical domain, a number of high-profile cases have acted as a powerful disincentive to the participation in incident reporting systems. For example, the High Court recently intervened to recommend the reinstatement of a surgeon who had expressing worries about the success rate of a colleague in his hospital. The trust initially refused to comply with the Court of Appeal's finding [109]. This parallels the case of Stephen Bolsin who first uncovered an unusually high death rate among babies undergoing cardiac surgery at the Bristol Royal Infirmary [436]. He subsequently claimed that he was unable to continue working in the NHS as a result of his 'whistle blowing' and was forced to move to a hospital at Geelong, near Melbourne. These causes have resulted in the Public Interest Disclosure Act (1998), which allows whistle blowing staff who feel they have been victimised to take their employers to an industrial tribunal. There is no limit to the compensation that can be awarded and employees simply need an "honest and reasonable" suspicion that malpractice has occurred or is likely to occur. Such protection has, however, proven to be insufficient to persuade employees to contribute to many voluntary reporting systems. For example, the 2001 Royal College of Nursing congress explicitly backed a call for action to protect nurses who 'speak out'. One of the delegates argued that whistle blowing was often seen as 'grassing up' or betraying colleagues. Theoretically, such additional protection should not be necessary under the 1998 Act. Some of these concerns can be explained by the informal pressures to conform to the norms of a particular working group. They can also be explained by the practical problems of preserving anonymity within small teams. Given the limited numbers of staff who perform particular tasks on particular shifts, potential contributors can often be identified through a simple process of elimination.

16.1.3 Analytical Bias

It is important not to underestimate the potential biases that influence the analysis of near misses and adverse occurrences. Over the past three years, we have conducted a series of interviews, surveys and observational studies of incident investigators and safety managers [750]. This work has helped to identify a range of influences that can affect the decision making processes that are intended to distinguish causal factors from the mass of other contextual information that is extracted from an initial report. At its most extreme, incident data can be used in a post hoc way to justify decisions that have already been made and positions that have already been adopted. For example, the proponents of Crew Resource Management training have used data from the

US Aviation Safety Reporting System in this way [412]. Chapter 10.4 briefly introduced these sources of analytical bias. For instance, *author bias* arises when individuals are reluctant to accept the findings of any causal analysis that they have not themselves been involved in. For instance, incidents at US highway-rail crossings can trigger investigations by federal organisations, such as the National Transportation Safety Board (NTSB) and the FRA. They can also result in state level enquiries. In some states, responsibility is divided between public agencies and the railroad operators. Elsewhere, responsibility is assigned to regulatory agencies such as the Public Utility Commission, Public Service Commission, or State Corporation Commission. In other states, investigations involve representatives of state, county, and city jurisdictions. Both state and local law enforcement agencies will also be involved if an incident involves the enforcement of traffic laws. Local government bodies are given responsibility for operational matters related to crossings through their ordinances. Each of these organisations can, and often do, hold different views about the causes of adverse events. The situation is slightly simpler for incident investigations in the UK. However, railway privatisation has created a situation in which conflict can arise between operating companies, Railtrack and HMRI. This is neatly encapsulated in Anthony Scrivener's recent article on Ladbroke Grove entitled 'Pass the signal - pass the blame' [733].

Confirmation bias arises when investigators attempt to ensure that any causal analysis supports hypotheses that exist before an incident occurs. In other words, the analysis is simply conducted to confirm their initial ideas. *Frequency bias* occurs when investigators become familiar with certain causal factors because they are observed most often. Any subsequent incident is, therefore, likely to be classified according to one of these common categories irrespective of whether an incident is actually caused by those factors [396]. There are many examples of these two forms of bias in the handling of SPAD reports prior to the Ladbroke Grove accident. Cullen estimates that approximately 85% of all such incidents were classified as the result of driver 'error' [195]. The frequency of such findings helped to reinforce this analysis as an acceptable outcome for any SPAD investigation; "I am led to conclude that the ready acceptance of blame by drivers, encouraged by the no blame culture, may have contributed to this poor analysis of root causes". The subsequent report argued that operating companies should review their incident investigation practices to ensure that there is no presumption that driver error is the sole or principal cause of SPADs.

Recognition bias arises when investigators have a limited vocabulary of causal factors. They actively attempt to make any incident 'fit' with one of those factors irrespective of the complexity of the circumstances that characterise the incident. These pressures can be illustrated by the response to initial reports of problems in the performance of cardiac surgery at Bristol Infirmary. The Society of Cardiothoracic Surgeons of Great Britain and Ireland discussed the reports of poor outcomes in 1989. Further information emerged during site visits in 1990. The sub-optimal results were attributed to the low volume of work because an increasing number of cases was widely believed to be associated with better outcomes. Adverse reports were, therefore, interpreted in a way that encouraged the generation of more work and that avoided questioning existing practices at lower volumes. The eventual enquiry argued that "the focus on throughput may with hindsight be thought to have distracted attention from further inquiry, as the Bristol results, with the exception of the figures for 1990, showed no real improvement" [436].

Political bias arises when a judgement or hypothesis from a high status member commands influence because others respect that status rather than the value of the judgement itself. This can be paraphrased as 'pressure from above'. *Sponsor bias* occurs when a causal analysis indirectly affects the prosperity or reputation of the organisation that an investigator manages or is responsible for. This can be paraphrased as 'pressure from below'. *Professional bias* arises when an investigator's colleagues favour particular outcomes from a causal analysis. The investigator may find themselves excluded from professional society if the causal analysis does not sustain particular professional practices. This can be paraphrased as 'pressure from beside'. The influence of these workplace issues can be difficult to assess. For example, the FRA Safety Board conducted an analysis of incidents from January 1990 to February 1999. This found that only 18 coded 'operator fell asleep' as a causal or contributing factor. The NTSB found these figures difficult to believe given the prevalence of such incidents in other modes of transportation [610]. Two NTSB investigations that had found fatigue as a causal factor were not coded in the FRA database as fatigue-related but as a failure to comply with signals. A number of influences might explain such different interpretations of the same incidents. For instance, the FRA plays a significant role in the promotion of the rail industry as well as in its regulation. The NTSB focuses more narrowly on the investigation of safety-related incidents. In consequences, the political, sponsor and professional influences that act on those organisations will be quite different.

This section has reviewed a broad range of 'human factors' problems that complicate the development and maintenance of incident reporting system. Some groups may choose to contribute to a scheme while others do not. Such participation patterns can be caused by a fear of retribution even in confidential no-blame systems. Analysts must, therefore, develop techniques to address the problems of under-reporting by eliciting contributions from potential participants. Alternatively, they can be forced to develop extrapolation techniques that can be used to make inferences about the nature of any potential incidents that might otherwise be reported by these groups. In either case, the problems of ensuring consistent participation across many different working groups can be exacerbated by the pressures that lead to analytical bias. For example, political bias can be exerted to ensure that a lack of reports from some groups is interpreted as positive evidence for a good safety record. Alternatively, high participation rates from certain groups can lead to forms of recognition bias that make analysts more likely to reach similar conclusions for incidents reported by different groups of workers.

16.2 Technical Problems

The interaction between particular participation patterns and the problems of analytical bias can frustrate attempts to obtain many of the potential benefits from incident reporting that were introduced in the opening sections of this Chapter. These problems would not be so serious if investigators and safety managers were equipped with an appropriate armoury of well-developed techniques. The methods and tools might then be applied to address the problems of under-participation. Individuals could then claim that any residual under-reporting had persisted 'in spite' of the most stringent efforts to elicit incident reports. Similarly, safety managers might claim that appropriate measures had been taken to combat various forms of analytical bias. Chapter 9.3 and 10.4 have described some of these techniques. Unfortunately, these techniques are not widely exploited. In contrast, most systems rely upon a range of ad hoc and 'in house' techniques to support both the elicitation and subsequent analysis of incident reports. This proprietary nature of these approaches can create barriers to information sharing. The use of different 'in house' methods also prevents comparisons being made between similar schemes in different countries or industries.

16.2.1 Poor Investigatory and Analytical Procedures

Previous chapters have identified a range of theoretical and practical issues that are often ignored during the development of small scale 'in house' analytical techniques. These issues, typically, have little importance during the initial stages of a reporting system. They can, however, become increasingly significant as the scope of any system expands to cover more potential contributors or as external regulatory intervention imposes increasing demands on those who are responsible for maintaining the reporting system.

We have identified two key theoretical ideas that must be considered when developing appropriate techniques for the analysis of adverse events: Mackie's Causal Fields and Hausman's Causal Asymmetries. Mackie argues that events result in effects that together form a 'causal field' [508]. For complex events, individual only observe a subset of these effects. Hausman's view of causal asymmetry builds on this argument [315]. If we know the cause we can predict the likely consequences, however, if we only know the consequences then it is far harder to unambiguously identify a single cause. An individual's interpretation of the cause of an incident, therefore, depends upon their observations of the effects and the relationship between those effects and a range of alternative possible causes. Additional complexity stems from the way in which most failures stem from several different factors that together form what Mackie terms a 'causal complex'. These theoretical ideas are reflected in the UK Health and Safety Executive's guidance on the incident and accident analysis that support railway safety cases:

> "There is much evidence that major accidents are seldom caused by the single direct action (or failure to act) by an individual. There may be many contributing factors that may not be geographically or managerially close to the accident or incident. There might also be environmental factors arising from or giving rise to physical or work-induced pressures. There is often evidence during an investigation that some of the contributory factors have been observed before in events that have been less serious. Accident and incident investigation procedures need to be sufficiently thorough and comprehensive to ensure that the deep-rooted underlying causes are clearly identified and that actions to rectify problems are carried through effectively." [352]

Unfortunately, many incident and accident techniques significantly under-estimate the complexity of causal analysis. Several existing approaches attempt to identify a single 'root cause'. Other techniques, fail to consider the range of alternative causes that can account for the same observed effects. This creates problems during subsequent enquiries and litigation when it can be shown that investigators failed to consider other plausible accounts. Such caveats can be levied at some of more advanced analytical tools, including WBA and Tripod.

There are a number of reasons why reporting systems fail to adopt existing investigation and analysis techniques. This book is, in part, intended to address the lack of reference material in this area. There are other problems. For instance, many small scale systems lack the resources that are necessary to hire or train existing staff in some of the more complex techniques. This creates particular problems when safety managers and investigators consider human factors issues. There is a tendency to blame incidents on inadequate attention or on poor staff performance. Such findings obscure or neglect the 'performance shaping factors' that contribute to human failure. This can be illustrated by Busse and Wright's analysis of incidents reported in an intensive care unit. The clinicians and nursing staff who were responsible for the system argued that a number of incidents stemmed from inattention and 'thoughtlessness'. This often led to recommendations that focussed on reminders, including numerous posters that describe recommended procedures [121]. The same events were then analysed by a human factors expert who argued that such reminders could only provide short-term protection against certain classes of adverse events. Their effectiveness declines rapidly over time. In contrast, the application of incident investigation techniques derived from the Tripod method, introduced in Chapter 10.4, revealed that many of the incidents of 'thoughtlessness' could also be interpreted as the result of 'work arounds' to support poorly designed or faulty equipment.

Larger-scale reporting systems can avoid some of these problems by ensuring that their staff are trained in appropriate analytical techniques. Unfortunately, there is little agreement about which approaches might support the causal analysis of incidents in many industries [453, 195]. It is instructive to note that even the GAIN initiative, which many regard as the most advanced attempt to create industry-wider reporting standards in aviation, has still to agree on a core set of analytical techniques. This lack of consensus has important consequences. It can undermine confidence in the findings of any investigation, especially when there are misgivings about the intent or purpose of any enquiry.

16.2.2 Inadequate Risk Assessments

It can be argued that safety managers and investigators are justified in their decision to reject many existing analytical tools in favour on 'in house' solutions. Previous chapters have argued that very few of the existing techniques can be integrated directly into the design and development of future systems. In particular, they are very poorly integrated with risk assessment. This lack of integration can have unfortunate consequences. For instance, I recently witnessed a design team deriving rough reliability estimates for the same components that their colleagues had already been studying using automated monitoring systems [423]. Incident reporting systems can provide evidence about the consequences of a potential failure and approximations for the likelihood of particular hazards. Such information can help to increase the accuracy of risk assessments which can be notoriously inaccurate [251].

The lack of integration between risk assessment and incident reporting not only affects the pro-active use of failure information to support future development. It can also prevent the effective allocation of resources within a reporting system. If investigators do not assess the risks associated with the recurrence of a previous incidents then it can be difficult to justify why one failure deserves closer investigation than another [420].

One of the main conclusions from this book is, therefore, that more support must be provided to support the two-way flow of information between risk assessment and incident investigation. The products of risk analysis can be used to gude the allocation of investigatory resources. The products of incident reporting can be used to inform estimates about the consequence and likelihood of future failures. A number of problems complicate the use of incident reporting data to guide the application of risk assessment techniques. Most risk analysis centres on the frequency and consequence of an event. It is, however, often unwise to assume that any recurrence of a near-miss incident will have the same outcome. Many reporting systems therefore assume that any recurrence will have the 'worst plausible outcome' [423]. This creates problems because different investigators can have very different opinions about what is, and what is not, a plausible outcome from any future failure. This is most apparent in the differences that can arise between the risk assessments that are produced by the safety managers in operating companies and those of regulatory organisations. The US Department of Energy will issue a Preliminary Notice of Violation as a way of warning management that they have under-estimated the risks associated with any recurrence of a particular incident. For example, a series of 'unplanned worker contaminations' in a national laboratory during 1999 led to exposures that were well within specified limits. However, the Department concluded that 'the lack of adherence to radiological work controls and the amount of radioactive material potentially available for uptake in the body' created the potential for more serious incidents in the future. The subsequent investigation argued that, in contrast, 'laboratory management was reluctant to acknowledge the serious nature of the concerns and treated them as a series of individual personnel errors' [207].

As mentioned, few analytical techniques provide explicit support for the use of risk assessments to drive the allocation of finite investigatory and development resources. Even if risk assessments are integrated into other investigatory techniques, there is no guarantee that investigators will respond in an appropriate manner:

> "During the almost five years preceding the Ladbroke Grove accident, there had been at least three occasions when some form of risk assessment analysis on the signaling in the Ladbroke Grove area has been suggested or proposed. The requests were: the Head of Technical Division's letter of 11 November 1996 which requested a layout risk assessment of the re-signaling (paragraph 43); the Field Inspector's letter of 16 March 1998 to Railtrack (paragraph 64); and the Railtrack Formal Inquiry of 1 July 1998 (paragraph 66). In addition there was an earlier request for details of measures taken to reduce the level of SPADs in the area around SN109 recorded in the Head of Technical Division's letter of 1st March 1995 (paragraph 39). None of these requests appears to have been pursued effectively by HMRI." [353]

Such comments illustrate another of the numerous paradoxes that arise in incident reporting. It is easier to identify situations in which risk assessment has failed to prevent a recurrence than situations in which it has successful mitigated the risks of future failure. For instance, safety managers might use a risk assessment to justify intervention to mitigate the consequences and likelihood of a particular failure. If their intervention has been successful then the number of similar incidents may fall and the outcomes of these events will be less 'severe'. However, such situations are indistinguishable from those in which the manager introduces unnecessary measures to address a risk that was lower than they had anticipated.

16.2.3 Causation and the Problems of Counter-Factual Reasoning

Investigators must adopt a consistent approach to the causal analysis of adverse events. Confidence can be compromised if different causes are identified for apparently similar events. Unfortunately, the proponents of many reporting systems underestimate the difficulty of causal analysis. The previous section outlined some of the theoretical problems that complicate this task. Counterfactual reasoning provides the main analytical technique for improving the consistency of causal analysis in incident and accident investigations [250, 469]. Chapters 9.3 and 10.4 have described how this technique has been integrated into a wide range of methods including Events and Causal Factor analysis as well as WBA. Counterfactual reasoning takes the general form that 'if a causal factor had not occurred then the incident also would not have taken place' [491]. If an incident would still have taken place whether or not a event had occurred then it cannot be thought of as causal factor. This style of argument is illustrated by an NTSB marine incident; "...had the main switchboard been subjected to thorough and timely inspections as part of an effective preventive maintenance program, any faulty connections or conductive objects would have likely been identified and corrected, and the fire might have been avoided." [620]. The same style of reasoning can be used beyond the immediate 'causes' of an incident to look at the actions that might have mitigated the consequences of the failure. For example, the report argued that "a firefighting team that was trained in the techniques of combating an electrical fire should have led the response to the fire in the control room...such a team probably would have extinguished the fire more quickly and with minimum risk". The author is using a counterfactual argument because a trained firefighting team was not available to combat the initial incident. The NTSB investigator also deploys counterfactual argumentation to eliminate potential causal factors. The report argues that 'even without a fuse, a transient voltage spike of sufficient magnitude to create an arc that could jump the gap probably could not have been created'. This is a counterfactual argument because there was evidence to suggest that a form of fuse had been present in the system before the incident. Hence with this additional safeguard, we can discount the cause of the fire being a transient voltage spike.

Counterfactual reasoning is both complex and error prone. For example, how sure can we be that an incident would not have occurred if a causal factor had not been present? Causal asymmetries suggest that many different causal complexes will have the same outcome. For instance, there are no guarantees in the previous incident that the inspections would have found a particular faulty connection. Previous incidents have shown that inattention and fatigue often compromise such safeguards. Chapter 10.4 has argued that the strengths and weaknesses of counterfactual reasoning remain an area for future research. Byrne and her colleagues have, however, conducted a number of preliminary studies [123, 124]. This work argues that deductions from counterfactual conditionals differ systematically from factual conditionals and that, by extension, deductions from counterfactual disjunctions differ systematically from factual disjunctions. This is best explained by an example. If we argue that "...had the main switchboard been subjected to thorough and timely inspections as part of an effective preventive maintenance program, any faulty connections or conductive objects would have likely been identified and corrected, and the fire might have been avoided" then readers will infer that the inspections had not taken place. This counterfactual style of argument can have such a persuasive effect that readers overlook contradictory evidence elsewhere in a report [426]. There are more complex examples of the inferences that readers draw from counterfactual arguments. The statement that *the fire was caused by a faulty connection within the main switchboard that initiated an arc fault or by a conductive object falling onto the switchboard bus bars* is a factual disjunction. Byrne argues that such sentences encourage the reader to think about these possible events and decide which is the most likely. There is an implication that at least one of them took place. The statement that *had the switchboard been covered by an effective preventive maintenance program or a thorough inspection by the Alaska Marine Highway System then the presence of faulty connections would have been identified* is a counterfactual disjunction. Chapter 9.3 has shown that this use of the subjunctive mood communicates a presumption that neither of these events actually occurred.

This theoretical work has pragmatic implications for incident investigation. If factual disjunctions are used then care must be taken to ensure that one of the disjuncts has occurred. If counterfactual disjunctions are used then readers may assume that neither disjunct has occurred. The distinction between counterfactual and factual disjunctions forms part of a wider concern to ensure that analytical biases are not hidden through the inappropriate use of language in incident reports. For example, rhetorical devices known as tropes can be used to increase the impact and effectiveness of everyday prose. Chapter 12.4 has briefly introduced the way in which tropes can be used to achieve particular effects on the readers of an incident report. For instance, *anaphora* uses repetition at the beginning of successive phrases, clauses or sentences. It can create an impression of climax in which the repetition leads to a particularly important insight or conclusion.

> "Both patients had implanted pacemakers, and both had experienced unintended maximum pacing rates up to 120 beats per minute. Medical intervention was needed to turn off the minute ventilation sensor in each pacemaker. When the sensors were turned off, the patients' heart rates returned to normal." [275]

This example illustrates the successive use of 'both' to emphasise the link between events happening to the patients. The investigator uses this repetition to draw the reader's attention to relationships between the consequences of a single cause for both patients. In this case, a clinical device such as a cardiac monitor or mechanical ventilators, was assumed to have generated a weak electrical signal that was sufficient to interfere with the ventilation sensors on the patient's devices. This, in turn, resulted in the incorrect measurement of thoracic impedance and ultimately in pacemaker rate increases. It is important to emphasise that such techniques are not of themselves either 'good' or 'bad'. Rhetorical devices can be used to convince us of well-justified conclusions or to support half-baked theories. It is important, however, to be sensitive to the effects that such techniques might have on the readers of an incident report. For instance, the previous citation can be interpreted to provide readers with a clear summary of the evidence that supports the investigators' conclusions. It can also be interpreted in a more negative light. For example, further investigation might establish independent causes for the effects observed in both patients. The rhetorical device creating a link between each individual might dissuade investigators from conducting such additional investigations.

Antithesis uses juxtaposition to contrasts two ideas or concepts. This technique is often used to contrast some form of normative or correct behaviour with the events that are presumed to have caused an accident. This can be illustrated from the following analysis of an incident reported to the Canadian Defence Forces:

> "...instead of braking gently, the driver's foot accidentally hit the accelerator. The vehicle jumped forward out of control, veered to the right, sped over the ditch, and crashed into the front wall of a 7-unit multi-family dwelling..." [147]

This technique is important because readers may make a number of additional inferences based upon such constructions. In this context, it is tempting to infer that the resulting collision would not have happened if the accelerator had not been pressed. The rhetorical construct diverts attention away from alternative hypotheses. The car may have been travelling too fast for any braking manuever to have prevented the eventual incident.

Most investigators and safety managers are unaware that they exploit such rhetorical devices. They draft prose to support their arguments. They may inadvertently stress conclusions that are not well supported by the available evidence. They may also cast doubt on other findings that contradict their version of events. Unfortunately, this inadvertent use of rhetorical devices is often exposed at litigation. In particular, it is often possible to show that particular linguistic constructs reflect the unsupported assumptions of investigators. In the previous example, it would be necessary to demonstrate that the accidental use of the accelerator was the cause of the incident and not the failure to brake well before the accelerator was applied. The key point here is that the problems of bias and interpretation not only affect the causal analysis of incident reports, they also complicate the way in which adverse events are documented and presented by investigators. If these influences are not considered then there is a danger that alternative explanations will be prematurely discounted and potential lessons lost.

16.2.4 Classification Problems

It can be difficult to detect patterns of failure amongst the natural language accounts of adverse events that are produced by many reporting systems. The volume of prose produced in national and international systems can make it difficult for any individual to keep track of common causes or consequences across many incidents. In consequence, many reporting systems use keyword-based summaries. Analysts represent the causes of an incident or near-miss by an enumeration of terms drawn from an agreed glossary. This approach can also provide a concise representation of a range of other contextual information, including mitigating factors and the potential consequence of an adverse event. The use of keyword summaries helps to reduce the interpretation problems that stem from tropes in natural language accounts. This approach strips out the rhetorical techniques that emphasise particular interpretations through the use of anaphora, antithesis etc. There are further benefits. For instance, the use of an agreed taxonomy can help to ensure that different organisations all consider a consistent set of terms when describing adverse events. Chapter 13.5 has also shown how classification schemes can be based upon the data models that support relational databases. Not only do these terms provide a vocabulary for describing individual incidents, they also provide the keywords that can be used to form the queries used to extract information about previous events. This use of classification schemes also supports the compilation of statistical data. Safety managers and regulators can provide information about the frequency of incidents that are attributed to each of the causes included in the taxonomy.

A number of practical problems complicate the use of such taxonomies to support the indexing and retrieval of incident reports. In particular, it can be difficult to establish reliable procedures for the codification of each adverse event. Each incident can be codified locally, within the group or organisation in which an incident occurred, or by a central unit who are responsible for codifying a large number of events sent in by different participants. If the codification of incidents is performed centrally then it is important that staff understand enough about the context in which an incident occurs for them to ensure that the correct codes are assigned. Alternatively, if incidents are to be codified at a local level then it can be difficult to ensure that safety managers assign the same codes to similar incidents in different locations [419]. The FDA illustrate some of t he problems in incident classification through a 'real' case study in which a violent patient in a wheelchair was suffocated through the use of a vest restraint that was too small. The risk manager, JC, proceeded as follows:

> "She finds the list of event terms, which was detached from the rest of the coding manual... She muses: 'Mr. Dunbar had OBS which isn't listed in these codes; he had an amputation which is listed; he had diabetes which isn't listed; and he had hypertension which is listed'. JC promptly enters 1702 (amputation) and 1908 (hypertension) in the patient codes. She then finds the list for Device-Related Terms... She reviews the terms, decides there was nothing wrong with the wheelchair or the vest restraint, and leaves the device code area blank." [277]

The success of any classification system, therefore, depends upon the procedures that are used to identify appropriate terms. The resulting classification of 1702 (amputation) and 1908 (hypertension) provides few insights into the nature of the incident. A range of techniques can be used to identify potential mis-classifications. Many of these rely upon comparing the results from any classification with those obtained from more reliable sources. For example, the frequency of particular terms in an incident classification can be compared with those from Sentinel schemes. The additional resources and training provided to Sentinel systems should ensure that their classifications provide a more accurate reflection of adverse events than those submitted by other reporting systems. Unfortunately, it can be difficult to judge whether any differences are due to misclassifications or to underlying differences in the nature of events that are reported to different units within the same industry. Alternatively, analysts can compare the results of an independent reclassification of previous incidents with those that were originally returned from a reporting system. Such comparisons again rely upon an appropriate sampling technique to ensure that this approach detects 'genuine' differences in any subsequent reclassification. If the sample focuses on particular classes of adverse event then this approach may fail to uncover wider problems of misclassification in the incidents that wre not selected as part of the sample.

A range of further problems complicate the application of taxonomies. Many incidents involve 'wicked' failures that cannot easily be described by a number of discrete terms [468]. For instance, computer-related failures often stem from a combination of requirements and design errors. System components fail because some necessary tasks are not identified and because others are identified but had not fully developed. For complex systems it can also be difficult to distinguish between a requirements failure and a design failure. This is especially difficult if detailed design and requirements documents are distributed across the sub-contractors that are responsible for implementing component functionality.

Further problems arise from the difficulty of classifying human behaviour. For instance, a recent HMRI report identified the difficulty of distinguishing between 'misjudgement' and 'disregard' [351]. The allocation of these different terms has a profound impact on the consequences of any investigation. 'Misjudgement' implies that the operator may have behaved in a reasonable manner even if they ultimately failed to safeguard the system. 'Disregard' suggests a more willful neglect of necessary precautions. The HMRI report cites the example of a train driver who appeared to make every effort to brake at a signal in poor weather conditions, yet the incident was categorised as 'disregard' rather than 'misjudgement'. The allocation of such terms involves a level of analysis and discretion that goes beyond the FDA's taxonomy, which focuses on observable features such as the role of hypertension in an incident.

In many industries, it can be difficult to ensure agreement over the components of incident taxonomies. In particular, there is a trade-off between the coverage of a taxonomy and the reliability of any analysis. In general, analysts are less likely to achieve a consistent classification if more terms are introduced into a scheme. The development of an appropriate taxonomy is further complicated by the need to respond to changes in the types of incident that are reported to a scheme. This creates particular problems if analysts are forced to go back and manually reclassify hundreds or thousands of records to reflect new distinctions and definitions of the components of any taxonomy. Some reporting schemes now hold more than 500,000 reports [59]. If previous records are not updated to reflect the new classification system then safety managers may fail to discern that recent incidents form part of a wider pattern, which is obscured by weaknesses in the previous classification scheme. This problem is particularly acute when taxonomies are extended to describe human behaviour. The field of human factors research has changed rapidly over the last decade with an increasing focus on group interaction. However, few of these changes have been reflected in incident reporting systems because of the costs associated with manually analysing and re-classifying existing records.

As mentioned, many classification systems are derived from or inform the development of databases. These systems support the retrieval of individual incident reports using queries that are constructed in terms of the components of incident taxonomies. For instance, analysts can use the FDA's MAUDE system to retrieve information about all incidents that were tagged with the device codes, mentioned in previous paragraphs. Unfortunately, the theoretical under-pinnings of these systems are often poorly understood by the people who use them. Safety managers, therefore, often rely upon queries that are pre-programmed by system administrators. Unfortunately, it can be difficult to ensure that safety managers clearly communicate their information requirements to technical support staff [416]. In consequence, safety managers often do not receive the information that they think they have requested when they issue a query. It can also be difficult for safety managers to formulate more ad hoc, exploratory queries because they lack necessary technical knowledge about relational database technology [416].

Chapter 13.5 introduced alternative technologies that avoid some of the limitations associated with the use of relational databases in incident reporting schemes. Web-based techniques can also be used to automate the indexing of incident reports in response to changes in a classification scheme. This avoids the overheads associated with the manual reclassification of many thousands of previous records. Similarly, probabilistic information retrieval systems enable users to search for information without the need to form complex queries [416]. Information requests can be expressed in the vernacular. The retrieval system analyses compares attributes of the query, such as the frequency of key terms, to automatically identify potential matches within a collection of incident reports. For example, it is relatively rare to find the term 'explosion' in medical incident reports. The use of this term in a query can help retrieval systems to identify a relatively small number of potential matches.

In contrast, less attention might be paid to the use of the term 'patient', which is likely to appear in many of the records within the system. Probabilistic information retrieval systems do not, however, provide a panacea. For example, it can be difficult to ensure that particular queries yield appropriate levels of precision and recall. A system exhibits poor precision if it returns many incidents that the user does not believe are related to their query. The user must then manually filter the large number of incidents that the system considers to be a match. Conversely, poor recall occurs when a system fails to return an incident that the user believes is related to their query. Systems that provide good recall are often imprecise. Conversely, systems that offer high degrees of precision will often exclude incidents that ought to have been returned as a potential match. A number of existing research projects have, therefore, begun to look for alternative computational techniques including conversational case based reasoning [416]. The intention is to avoid the static data models that limit the application of relational databases but also to help users interactively address the problems created by poor precision and recall. Conversational case based reasoning techniques enable users to filter the presentation of incident information through the iterative refinement of natural language queries.

16.3 Managerial Problems

Previous sections have focussed on problems that can prevent reporting systems from yielding the benefits that are often claimed for them. These range from the difficulty of eliciting sufficient reports through the problems of ensuring consistent analysis through to the computational challenges that complicate the storage and retrieval of large-scale incident collections. In contrast, this section focuses on the problems of managing incident reporting systems. It is important not to underestimate these problems because they are often far more significant than the technical issues that have been summarised in previous paragraphs. If insufficient resources, including time, finance and expertise, are allocated to a reporting system then there is little prospect that it will yield significant safety improvements. If management structures are not established to enable the dissemination, implementation and monitoring of safety recommendations then any insights are unlikely to be acted upon. If higher levels of management, regulators and political interests have unrealistic expectations about the likely benefits of a reporting system then it is unlikely to satisfy any subsequent validation actions.

16.3.1 Unrealistic Expectations

Chapter has presented statistics that indicate the high frequency of adverse events and near-miss incidents in many safety-critical industries. It has been estimated that between 4% and 17% of patients in acute hospitals suffer from iatrogenic injury [850]. Observational studies have found that 45% of patients experienced some medical mismanagement and 17% suffered events that led to a longer hospital stay [28]. The US Aviation Safety Reporting System averages approximately 600 reports and UK railway's Safety Management Information System receives over 1,700 reports each week. It can, however, be difficult to interpret these statistics. The 4-17% figure for iatrogenic injury is based on extrapolations that may not be confirmed by new regional reporting procedures in the UK and the US. Similarly, the number of submissions to the Safety Management Information System is partly due by the way it helps record reports that meet regulatory requirements under the UK's RIDDOR regulations. In spite of these caveats, many of the proponents of incident reporting point to the success of schemes such as the ASRS to justify the development of new systems. They often underemphasise the potential problems that can arise when data is collected unnecessarily or that can occur when finite resources are swamped by a flood of 'low criticality' reports. For instance, since the Ladbroke Grove accident, there has been a requirement to conduct a formal inspection after every report of a SPAD. This has resulted in over 200 formal investigations by Her Majesty's Railways Inspectorate (HMRI). It is likely that there have also been a far larger number of near-misses. Drivers frequently avoid passing the signal by rectifying a potential problem 'at the last minute'. It has been argued that confidential, voluntary reporting systems might be used to elicit information about these events that would otherwise go unrecorded [195]. In particular, they might provide insights both about those measures that helped the driver to detect the potential danger and about those factors that might have turned a near miss into a more serious incident.

Some individuals and organisations have looked beyond the particular safety lessons that might be learned from a reporting system. In addition to identifying successful defences and potential vulnerabilities, they have argued that such schemes will also reduce costs by avoiding the negative consequences of previous failures. For instance, the NHS have promoted the development of voluntary reporting systems as one of several measures that are intended to achieve a number of ambitious objectives:

> "...the Department of Health should establish groups to work urgently to achieve four specific aims: by 2001, reduce to zero the number of patients dying or being paralysed by maladministered spinal injections (at least 13 such cases have occurred in the last 15 years); by 2005, reduce by 25% the number of instances of negligent harm in the field of obstetrics and gynaecology which result in litigation (currently these account for over 50% of the annual NHS litigation bill); by 2005, reduce by 40% the number of serious errors in the use of prescribed drugs (currently these account for 20% of all clinical negligence litigation); by 2005, reduce to zero the number of suicides by mental health inpatients as a result of hanging from non-collapsible bed or shower curtain rails on wards (currently hanging from these structures is the commonest method of suicide on mental health inpatient wards)." [635]

Such high expectations can contrasted with the prosaic problems that limit the utility of incident reporting systems. There are further problems. Chapter 14.5 has described the difficulty of monoring whether or not a reporting system is meeting the ambitious objectives that often set for them. Increases in the reporting frequency can be due to increased participation in a scheme or from a sudden rise in adverse events. Independent audits can be used to distinguish between these two interpretations, however, limited resources prevent their use as a persistent monitoring device. There are further problems; a reduction in the accident frequency cannot always be used as an indicator. In many industries, high consequence adverse events are so rare that any reduction from the beneficial effects of incident reporting cannot be distinguished from random changes even over relatively long periods of time. A number of more sophisticated statistical models have recently been proposed to address these concerns. For instance, UK RailwaySafety's Precursor Indicator Model measures short-term changes in the frequency of events that have led to previous accidents. The precursors that form the focus of this monitoring technique are updated over a longer time period to reflect any changes in the causes of those accidents that do occur. Such approaches recognise the technical complexities that arise in attempting to extract data to support the management of incident reporting systems. One of the aims behind this book is to undermine some of the more simplistic and extreme claims that are made about the benefits of incident reporting. A further intention has been to provide safety managers with information about tools, such as the Precursor Indicator Model described in Chapter 14.5, that can be used to avoid some of the pitfalls that have affected previous reporting systems.

16.3.2 Reliance on Reminders and Quick Fixes

Risk monitoring tools, such as RailwaySafety's PIM techniques, are not an end in themselves. They provide information that should be used to inform future decision making. Information about previous failures must guide the interventions that reduce the likelihood or mitigate the consequences of future incidents. Previous chapters have cited many success stories where incident reports have triggered a prompt and effective response to adverse events. These have ranged from new forms of glazing in the viewports of military bunkers described in Chapter 11.5 [144] through to changes in the ventilation systems in the laundry facilities on cruise ships mentioned in Chapter 12.4 [608]. It is important to recognise, however, that the development of a reporting system does not always guarantee that recommendations will be acted upon. For example, Chapter 8.3 recounted the thirty year campaign by the NTSB to encourage the wider use of Excess Flow Valves in consumer gas lines [589]. Similarly, the concerns voiced by the ASRS about the consequences of tight turnaround times in US airports have had some limited effects [412]. It is unrealistic to expect that every recommendation proposed by a reporting system should be implemented. Economic and commercial pressures have been cited as reasons for the delay in implementing the two recommendations mentioned above.

There are further generic reasons why reporting systems can fail to yield the safety benefits that some have predicted. For instance, these schemes often yield few surprises. They help to reinforce existing safety concerns that many managers will already be very familiar with. For instance, Sexton, Thomas and Helmreich's study of medical incident reporting found that the most common recommendation for improving patient safety in intensive care was to acquire more staff to handle the workload. The most common recommendation from incident reporting in operating theatres was to improve communication [736]. The lack of surprise should be unsurprising. Organisations that establish reporting systems, typically, already have a good idea of the safety issues that affect their working practices. Van Vuuren's research, cited in Chapter 10.4, provides further examples. One study focussed on 19 incidents that were reported over approximately one month to an Accident and Emergency Department [845]. His analysis yielded a total of 93 potential causes of which 45% related to organisational issues while 41% were classified as 'direct' human causes. The organisational causes included the need to secure external services. In particular, incidents were often triggered or exacerbated by the need to secure beds for the patients in the Department. They also included a lack of senior staff during peak periods. Direct human causes included problems that new Senior House Officers experienced in interpreting X rays. They also stemmed from a culture of learning from mistakes and a reluctance to contact senior staff. Many NHS safety managers are already very aware of these issues. Such concerns do not, however, secure the resources and organisational support that is necessary to implement specific improvements.

It is difficult to underestimate the impact that resource issues can have upon the benefits of any reporting system. Some schemes have proven to be very successful in triggering large scale investments in safety measures. Insurance and safety classification companies, including Lloyds Register and DNV, have played an important role in motivating upper levels of management to invest in the recommendations that are derived from reporting systems. In contrast, other more local schemes have often been set the target of becoming 'self funding'. In other words, any safety investments must be paid for by corresponding process improvements that are elicited as 'lessons learned' through the same reporting mechanisms [419]. These constraints can often lead to a form of resource starvation. Managers are forced to make recommendations that they know will never be funded or to focus their attention on 'low cost' and 'no cost' solutions. This partly helps to explain the dominance of the *perfectability approach* to risk management, mentioned in Chapter 11.5. Rather than address the organisational, technical and environmental causes of adverse events, staff are urged to 'try better', 'be more aware' or 'follow established procedures'. Busse and Wright's study of reporting in intensive care found 82 'Remind Staff?' statements in a total of 111 recommendations over a 15 month period [119]. 29 other recommendations focussed on revised procedures and protocols (e.g. 'produce guidelines for care of arterial lines - particularly for femoral artery lines post coiling'), or were equipment related (e.g. 'Obtain spare helium cylinder for aortic pump to be kept in ICU'). A reliance on reminders can also be seen in larger scale, voluntary reporting systems. For example, Chapter 1.3 described how the aviation industry's CHIRP system often relies upon the perfectability approach. Pilots have been urged to check that they have entered the 'correct' data into navigation systems 'then, and only then, should the Execute function button be pressed' [177]. Similarly, they have been urged to complete pre-flight visual inspections of all flight surfaces or in cases where this is not possible, such as a high wing high tail configuration, to ensure 'a sound knowledge' of local de-icing processes in the prevailing weather conditions, 'If there is any doubt as to whether the aircraft is clean, a take-off should not be attempted' [180]. Similarly, the ASRS has encouraged pilots to mitigate the problems of noise cancelling headphones by having them 'half-on' and 'half-off' during take-off and landing [62]. These reminder statements can be interpreted in two different ways. At one level, they provide an important source of practical information for operators. This is reinforced by a recent contributor to CHIRP:

> "It is timely that we remind ourselves of the health and safety hazards that may exist on the aircraft. It is also timely that we remind ourselves that we are individually responsible for our own health and welfare in situations that we know are hazardous.

[179]

These reminders can, however, be interpreted in a different way. It is possible to argue that far from improving safety, they illustrate potential weaknesses in the defences that protect safety-critical systems. Human factors research points to the dangers of any reliance on reminders. Unless people are continually reminded then they are likely to forget the importance of safety precautions over time [369]. Any reminder should be seen as a prompt for more concerted action to address underlying technical and organisational issues. In this view, reminders are short-term fixes to deeper safety problems ranging from the design of human-machine interfaces to navigation systems through through to the intergration of noise cancelling headphones into cockpit auditory warnings. There is evidence that the managers of both the CHIRP and ASRS systems have responded in this way, they have initiated and contributed to more sustained safety initiatives following particular incident reports. It is also instructive, however, to examine those incidents that typically yield reminders. As we have seen, in the aviation domain these often concern cockpit design but can also relate to commercial pressures to meet particular ATC slots. In the medical domain, reminders are often used to cope with the lack of beds or key staff or with specific medical devices acquired by other units in the hospital. These reminders are intended to address problems that are, typically, perceived to lie outside the scope of the reporting system. Staff must perform better to cope with issues that cannot be addressed by more direct means, including changes in acquisition policy and commercial practices. The presence of such reminders, therefore, provides insights into the perceived limitations of an incident reporting system.

16.3.3 Flaws in the Systemic View of Failure

The reliance on reminders by 'perfectionist' approaches to incident reporting often stems from an undue emphasis being placed upon the direct human causes of adverse events. Operators are blamed as the main cause of an adverse event, hence they should be exhorted to 'try better, next time'. Correspondingly less attention is placed on the organisational, technical and environmental circumstances the contributed to an incident. The managers of many schemes have become sensitive to these criticisms. For instance, Chapter 12.4 quotes the Swedish Maritime Agency's statement that 'it is to be underlined that it is not the purpose of the investigation to establish or apportion blame or liability' [769]. Rather than focusing on individual human errors, the purpose of their reporting system is to provide 'a complete picture' of adverse events. Similarly, a recent report into SPADs on UK railways stressed that 'no driver sets out to have a SPAD, but all humans are prone to unintentional error on occasions'. It went on to argue that in order to reduce SPADs it was necessary to understand the factors that contribute to these events and that '... there is a growing body of evidence that features in the design and configuration of the signalling system can significantly increase the risk of driver error' [358]. This view has also been embodied in HMRI guidance on the preparation of safety cases, this rejects the identification of operator error as a root cause of incidents on UK railways:

> "In these criteria the term 'root causes' includes consideration of management' s real and perceived messages to workers, environmental and human factors, as we ll as plant failures and inadequate procedures. Human errors arising from poor operating conditions, procedures, management expectations or plant design are not root causes; the predisposing factors are." [352]

In the domain of medicine, the UK NHS' 'Organisation with a Memory' argued that although human error can precipitate an incident, there are usually deeper 'systemic factors' that created the context in which an adverse event was likely to occur [635]. In contrast to the perfective approach, this systemic approach:

> "... takes a holistic stance on the issues of failure. It recognises that many of the problems facing organisations are complex, ill-defined and result from the interaction of a number of factors. This approach starts from the premise that humans are fallible and that errors are inevitable, even in the best run organisations. Errors are seen as being

> shaped and provoked by upstream systemic factors, which include the organisation's strategy, its culture and the approach of management towards risk and uncertainty. The associated counter-measures are based on the assumption that while we cannot change the human condition we can change the conditions under which people work so as to make them less error-provoking. When an adverse event occurs, the important issue is not who made the error but how and why did the defences fail and what factors helped to create the conditions in which the errors occurred. The system approach recognises the importance of resilience within organisations and also recognises the process of learning as enhancing such resilience." [635]

The claim that accidents occur in the 'best run organisations' echoes Perrow's argument that companies will still have accidents no matter how hard they try [677]. His work on 'normal accidents' provides theoretical under-pinning for the systemic view of failure [677] Accidents occur in unexpected ways because operators, managers and regulators cannot hope to control the many different hazards that are created by the complexity and coupling of high-technology systems. It is easy to misinterpret many of Perrow's arguments especially about the nature of politics and control in high-reliability organisations. It seems clear, however, that he views adverse events as part of the price to be paid for technological innovation. Accidents and incidents also provide 'warning signals' about the trade-offs between safety and production that characterise many industries. The previous chapters of this book provide numerous examples to confirm Perrow's view. Railtrack have made considerable efforts to introduce innovative incident reporting systems. They have also been involved in a succession of major accidents. Perrow argues that if we cannot prevent potential catastrophic failures in high-technology systems then they should be 'abandoned, drastically scaled back or drastically redesigned'. However, many of these rail accidents occurred while design changes are being introduced across the network. The introduction of the Train Protection and Warning System, mentioned in Chapter 2.3, is scheduled for completion in January 2004. Unfortunately, it is difficult to abandon or drastically scale back a national rale system in the meantime. The reduction in network traffic following the Hatfield crash played a major role in the UK governments decision to force the infrastructure company into receivership.

At the heart of normal accident theory is the argument that previous technologies supported linear and loosely coupled systems. It is possible to anticipate and counter the hazards that are raised by dams and canals. Accidents are 'foreseeable and avoidable' because if one component fails then there is time to react and mitigate the consequences of adverse events. In contrast, modern production and transportation systems exhibit non-linear behaviours. They exploit techniques such as 'just in time' production that provide efficiency savings but that also create potential vulnerability. Individual failures propagate throughout and beyond system components in ways that would not have been possible in loosely coupled systems. Again, previous chapters have supported these arguments. For instance, Chapters2.3 and 6.4 describe how a decision to increase the traffic on a Brunswick Line rail system led to the placement of a signal before a station so that drivers had sufficient time to break before a hazard on the other side of the station. This contributed to the subsequent accident because investigators argued that drivers were less likely to recollect the previous signal after they had performed a station stop [598]. This incident illustrates many of Perrow's arguments. The decision to increase line capacity implied higher operating speeds and longer stopping distances. This consequent operating changes introduced greater complexity for drivers and signallers. The increased operating speeds reduced the opportunity to correct any driver failure to recollect the signal. High-line capacity also made it more likely that any such failure would have catastrophic consequences; it was more likely that the train would come into conflict with other traffic on the line.

Perrow's arguments for the systemic causes of failure in high-technology systems are persuasive. They do not, however, explain some of the observations in this book. It is difficult to maintain a distinction between linear and non-linear systems. In particular, such distinctions cannot easily be used to identify situations in which accidents are 'foreseeable and avoidable' and those in which they cannot be predicted or mitigated. Even relatively simple 'systems' can fail in complex and unexpected ways. Chapter 7.3 reviewed the catastrophic effects of excavation activities that failed to account for soil characteristics and the knock-on effects of digging holes in particular locations.

This analysis is confirmed by Petroski work on the role of failure in civil engineering [680]. Perrow counters these objections by arguing that linear processes often form part of more complex non-linear systems and hence may lead to unexpected forms of failure. This analysis does not, however, explain the lack of surprise that engineers and managers often express over the failures that affect even the most complex and tightly-coupled technological systems. This is true, for instance, of many of the adverse medical events reviewed in Chapter 13.5. Most iatrogenic incidents stem from recognised problems that managers lack the resources or the institutional support to address. Pressures to attain other organisational or commercial objectives stifle the concerns that are expressed in the aftermath of previous incidents. This analysis shares much in common with Sagan's view that factional interests can oppose safety measures that threaten their position within high-reliability organisations [719].

The lack of surprise that accompanys the publication of many incident reports can be explained in several ways. Theories about the causes of an incident are often formed in the aftermath of an adverse event while primary and secondary investigations are being conducted. This can lead to speculation that informs and is informed by the broadcast media. This line of analysis is unsatisfactory. The same speculation occurs both after accidents and incidents. Hence it cannot explain the differences between our observations and the surprising, catastrophic failures studied by Perrow. These differences can, however, be explained by the relatively high frequency of the adverse events that we have studied. Managers and operators are less surprised by the causes of low consequence incidents because they occur more often than the 'catastrophic' failures studied by Perrow. Further caveats can, however, be raised against this argument. In the medical domain it is possible to identify a range of high-consequence incidents whose causes are very unsurprising, including junior doctors' lack of experience and the high workloads imposed upon some staff [453, 635]. Even if a clinician never witnesses an accident, they will still be familiar with many of the causes of these events.

There is, therefore, an irony in Perrow's use of the term 'Normal Accidents'. The types of catastrophic, unexpected failures that he studies are both atypical and very rare. Coupling and complexity are, of course, major concerns. However, most accidents in high-technology systems occur through combinations of well-known problems. These unexpected combinations lend the element of surprise. In Mackay's terms, we are often familiar with the singular causes of adverse events but cannot anticipate the many causal complexes that lead to incidents or accidents [508]. I would not, however, go as far as Leveson who argues that it 'is often a matter of luck' whether the causal conditions for an incident exist [486].I would argue that the formation of causal complexes is a matter of time rather than luck. The longer organisations continue to neglect well-known singular, general causes of failure then the more likely it is that they will contribute to adverse events or near-miss incidents. Whether they develop into catastrophic accidents depends on the nature of the causal complex and the barriers that can be used to mitigate the consequences of an initial failure.

16.4 Summary

There has been a rapid growth in the number and scale of incident reporting systems. In the UK, the Ladbroke Grove rail accident stimulated a range of initiatives including the expansion of the CIRAS voluntary reporting system [195]. The Bristol Infirmary enquiry had a similar impact on UK healthcare [436]. In the United States, the Institute of Medicine report 'To Err is Human' prompted many states, as diverse as Arizona, New York and Washington, to draft bills that establish additional reporting voluntary and mandatory systems for healthcare professionals [453]. In Japan, the Maritime Labour Research Institute in Tokyo is one of several organisations that have begun to investigate alternative forms and procedures for incident analysis [554]. This builds on work in the nuclear and aviation industries. The Japanese Ministry of Public Welfare has also instructed hospitals to develop reporting systems to help reduce iatrogenic incidents. Strong claims have been made about the potential benefits of these systems. Incident reporting applications are argued to offer valuable insights into the near-miss incidents that have the potential to threaten future safety. They can also be used to elicit information about 'lessons learned' and act as an exchange for best practice [845].

This chapter has, however, argued that significant barriers must be addressed before incident reporting systems can be successfully applied within many industries. These can be summarised as follows:

1. *reporting biases.*
   There are few guarantees that all staff will contribute to a reporting system. Variations in participation rates have been observed both within working groups at the same location, as in hospital systems, and between geographical regions, for example across rail networks. Automated systems are increasingly being introduced to trigger investigations into near-miss incidents. However, some tasks cannot easily be instrumented. Many of the more specialised monitoring systems are unreliable and often provide 'false positives' that consume finite analytical resources. In consequence, it seems likely that reporting rates of less than 20-30% will be typical of many healthcare applications. These problems do not affect some reporting systems. SPAD reports provide a relatively accurate impression of the frequency of these events. However, the monitoring systems that help to detect these incidents tell us very little about events that *almost* resulted in an incident but that were narrowly averted by operator intervention.

2. *blame.*
   Some local systems enjoy good levels of participation while trusted individuals administer the scheme. Staff learn to trust the integrity of those individuals. Participation rates often fall dramatically when they are replaced [119]. This effect is clearly linked to potential contributors' concerns that they will be viewed as 'whistle blowers' either by their colleagues or by those who administer the system.

3. *analytical bias.*
   There are numerous forms of bias that can affect the analysis of incidents once they have been reported. These include author bias, judgement and hindsight bias, confirmation and frequency bias, recognition bias, political, sponsor and professional bias. This is not an exhaustive list but it illustrates the difficulty of ensuring that any investigation is not hindered by 'undue' influences. These issues are particularly important in incident reporting when many stages of an initial investigation and analysis will be performed not by an external authority but by the organisation that was directly involved in the occurrence.

4. *poor investigatory and analytical procedures.*
   Once an adverse occurrence or near miss has been reported, it can be difficult to determine what factors should be included within an investigation. This is important for theoretical reasons because it can be difficult to identify salient factors within what Mackie terms the 'causal field' [508]. Hausman also points to the problems created by 'causal asymmetry' [315]. If we know the cause then we can determine the effects. However, if we observe the effects then it can be difficult to reach firm conclusions about the multiple possible causes of those observations. These theoretical problems are exacerbated by the resource constraints that affect incident reporting. Many organisations lack both the funding and the expertise to investigate more than a single causal hypothesis. This clearly limits the value of any insights that might be obtained from the analysis of near miss incidents.

5. *inadequate risk assessments.*
   The design of safety-critical applications is typically guided by some form of risk assessment. Risk can be thought of as the product of the consequence and the likelihood of a particular failure. Incident reporting systems have been proposed as powerful means of informing risk assessments. They can provide quantitative data about the relative frequency of previous failures [453]. As we have seen, however, analytical and reporting biases undermine such

statements. Similarly, the nature of 'near miss' incidents makes it very difficult to identify the 'worst plausible outcome' that might inform any decision about the consequences of a future recurrence.

6. *causation and the problems of counterfactual reasoning.*
   Many organisations have responded to the problems of analytical bias by recommending techniques that draw upon counterfactual reasoning. This style of argument takes the form; 'X is a causal factor if the incident would not have occurred if X also had not occurred'. Counterfactual reasoning is both difficult and unreliable. For example, there is often an implicit and unwarranted assumption that X occurred [315]. Similarly, Mackie's work on causal complexes suggests that investigators are likely to find many different X's for any single adverse event. The problems of counterfactual reasoning are compounded by Hausman's observations about causal asymmetries. Not only are there likely to be many X's within a causal complex, there are also likely to be many alternative causal complexes that might explain the observed effects of an incident . Unfortunately, all existing analysis techniques rely upon the subjective judgement of individual investigators to determine which of these X's are 'plausible' causes. Even the more formal, mathematically based techniques rely upon weightings or partial orders that ultimately reflect subjective assessments.

7. *classification problems.*
   Many organisations have responded to the problems of counterfactual reasoning by adopting causal taxonomies. These initiatives form part of a wider attempt to classify incidents according to a restricted range of criteria. This offers numerous benefits. In particular, the elements of the classification be used as indexing terms in relational databases. Unfortunately, field studies have shown that few safety managers know how to use these tools to accurately extract information about previous incidents. Problems also arise when the items in a database have to be manually reclassified to reflect changes in a causal taxonomy. This can be particularly onerous for national systems that hold many hundreds of thousands of records. Several prototype systems have been developed to address these problems. For instance, we are using information retrieval techniques that were originally developed for mass-market web-based applications. These approaches are the subject of on-going research and currently suffer from poor precision and recall.

8. *unrealistic expectations.*
   Many people who initiate reporting systems expect reductions in the frequency and consequence of adverse events that are unreasonable given previous experience in running these schemes. These expectations are particularly problematic given that many types of incident will not be reported to confidential systems. There can be strong organisation and cultural barriers that prevent employees from disclosing information about their friends and colleagues;

9. *reliance on reminders and quick fixes.*
   Many reporting systems lack the financial resources that are necessary to address underlying system failures. These systems are, typically, seen as a form of cost reduction rather than as a form of income generation. This separation of reporting systems from sources of investment can result in recommendations that focus narrowly on 'quick fixes'. Studies of previous systems have seen a tendency to adopt a perfective approach in which operators are urged to try harder to avoid future incidents. Such reminder statements provide dubious protection given that they must be continually reinforced if they are not to be forgotten.

10. *flawed systemic views of failure.*
    Rather than focusing on the role of individual operator error in the causes of an adverse event,

attention has shifted to the 'systemic factors' that make failure more likely. For instance, Perrow has argued that the coupling and complexity of high-technology systems make it difficult to predict and prevent potential catastrophes. We would argue, however, that many operators and managers are already very familiar with the singular, general causes of adverse events. The difficulty lies in predicting how these individual factors will combine to form causal complexes. Reporting systems can help to address this problem. Information about previous incidents can be analysed to identify common patterns of failure. Unfortunately, the utility of this approach is compromised by the problems of analytical and reporting bias.

This is a partial list. The previous chapters in this book have mapped out a number of additional problems that complicate the development of incident reporting systems. These include the difficulty of determining whether or not a particular scheme has had any impact on safety at all. Previous chapters have also summarised the 'state of the art' in terms of the techniques that safety managers and regulators might use to address these problems. The presentation of these techniques has been driven by the use of case studies and by practical experience in applying them to previous incidents. Many of these techniques provide only partial solutions, there are no panaceas. Instead, I hope that by bringing together a wide range of material from many different industries it will be possible to learn from the successes and failures that others have experienced in the development of their reporting systems.