Unpublished interim working document

October 2008

A review of safety management literature

social care institute for excellence

scie

A review of safety management literature

Eileen Munro

London School of Economics

January 2008

Unpublished report;
Interim working document
for SCIE Project “Learning Together”

Contents

Chapter One

1. Introduction

Accidents and disasters in any area of life raise urgent questions about why and how they happened. There are usually two key concerns: to find out who, if anyone, is to blame and to learn lessons to avoid a repeat of the incident. Child welfare services have been no exception to this response, especially in the UK, US, Canada, Australia and New Zealand. When children die or there are other adverse outcomes, inquiries have been conducted to examine the services offered to the families and to learn lessons to improve practice. Some recommendations from these inquiries have been highly influential in shaping the development of child welfare practice. One of the earliest carried out in the UK was the Dennis O'Neil inquiry in 1945 (Home Office, 1945) . This revealed the inadequate standard of assessment and supervision of the foster home in which Dennis was placed and subsequently killed, but also reported that such low standards were prevalent. This informed the 1948 Children Act which set out higher standards of the management of foster care that were legally binding on local authorities in England and Wales. The next major public inquiry was into the death of Maria Colwell in 1973 (Department of Health and Social Security, 1974) . This was also a milestone in leading to new policies and structures that transformed the way that the various professional groups in child welfare worked together to detect and respond to child abuse. However, there have since been hundreds of other inquiries that have had a less obvious impact. Many of these are done by local services and are not publicly available but reviews of published inquires reveal a disheartening message that they are producing the same lessons and formulating similar recommendations for improving practice (Department of Health and Social Security, 1982, Department of Health, 1991, Munro, 1999) This raises the question why the lessons are not learned and further incidents avoided.

A similar history in high risk engineering industries of inquiries producing recommendations that failed to lead to the desired improvement led investigators to reformulate their view of how errors can be studied, offering a deeper analysis of the chain of causation and the development of an alternative approach to investigating why and how mistakes occur. The shift from the traditional to the new form of inquiry has been described as that from a person-centred to a system-centred approach (Reason, 1990b). A key feature is that when human error is found, it is taken as the starting point for investigation, not as the conclusion as it so often is in traditional inquiries. This leads to the performance of individual practitioners being seen as part of a more complex chain of causality running through the whole organizational system.

This review of the literature on error investigation aims to provide an account of these developments that have been taking place in safety management in high risk industries. Much of the work has been done in areas of engineering such as aviation or nuclear power so the relevance to child welfare needs to be teased out. A more closely linked field of development is in healthcare where many government agencies and the World Health Organization have been making efforts to adopt the lessons from engineering and modify them to fit a health service. These sets of literature can be the basis for devising ways of adapting the ideas for use in child welfare services. The aim of this review is to provide an overview of the dominant frameworks and their associated models, and to start to draw out their implications for developing a model in child welfare. It will cover the key concepts, the key debates, and the key methods for collecting and analysing data.

This is not a systematic review because (a) the literature is too vast, and (b) it pays attention to conceptual material, not just the results of empirical studies. The pool of relevant material is potentially huge, drawing on numerous different disciplines and ranging across diverse fields including mechanical engineering, aviation, mental and physical health services, management, etc. Literature has been selected primarily using the criteria of citation rates (reviewing the works that are repeatedly mentioned by later authors) and by apparent relevance to child welfare.

The criterion of citation rates reveals some key people in the development of the theoretical foundations and the methodologies of a systems approach and their work is drawn on heavily in this review: Peter Reason, Jan Rasmussen, Charles Perrow, Erik Hollnagel and Robert Helmreich have written extensively on the subject. In relation to health care, the team led by David Woods at Ohio University, US and that led by Charles Vincent at St Mary's Hospital, London have done much of the pioneering and development work in adapting the approach to medicine.

There are some publications in healthcare that are a particularly useful source because their aim is comparable to this review in that they seek to explain the novelty of the approach compared with traditional modes of investigation and to identify how the methodology can be transferred to areas of work where the social dimension of the socio-technical system is more dominant, e.g. "To Err is Human: Building a Safer Health Service" (Institute of Medicine, 1999) is an American publication putting the case for changing to a systems approach at a national level. (Woloshynowych et al., 2005) reviewed the methods of accident investigation in high-risk industries and of critical incidents in healthcare with a view to identifying what could usefully be adopted in healthcare. They identified twelve methods for closer scrutiny, and concluded that there is considerable potential for further development of techniques and a need for evaluation and validation of methods used.

The term used in this review for the new way of investigating errors is a 'systems approach'. Some readers may be more familiar with the closely overlapping concept of 'root cause analysis' but this term is not used because it is seriously misleading. First, it implies a single root cause when, typically, errors arise from the interaction of a number of factors. Also the adverse outcome may not have a root but an active cause, such as a malicious action by front line worker. A third factor is that the word 'system' in the name draws attention to the key feature that the investigation is an opportunity for studying the whole system. It is a chance to learn not just of flaws in the system but of what is working well. It is important to remember that error and success are two sides of the same issue. Avoiding error implies improving performance.

Indeed, the systems approach to understanding errors offers not only a method for investigating tragedies in child welfare but also for studying normal practice with a view to gaining a deeper picture of how the system is operating to support front line workers, to learning what is working well for children, and identifying weak areas where there is a need for development. The overall goal of a child welfare system is to improve children's safety and welfare; reducing error is a means to that end not the end in itself.

The following chapters outline the key elements of the subject. Chapter Two explains the new view of safety management, and why the traditional view was unsatisfactory. The next chapter looks at the theoretical foundations of the different approaches, the assumptions made about how to study technical and social systems. The difficulties of defining and classifying errors are then addressed. The following three chapters look at error classification and analysis in different dimensions of the system: factors in human performance, in the working relationship between people and the range of artefacts designed to improve their performance, and factors in the overall organizational system in which the individual operates. Chapter Eight presents the ways that data can be collected, detailing both continuous mechanisms to facilitate continual learning and incident-focused methods. Chapter Nine recounts the three main areas of implementation in healthcare: in the acceptance of the approach at the highest organizational levels, the development of reporting systems to facilitate learning from minor mishaps, and the design of methods for investigating specific incidents. The final chapter raises some issues that need to be addressed in transferring the approach to the field of child welfare

> *Since the purpose of the literature review is to inform the adaptation of the systems approach for use in child welfare, there are, at relevant points, indented sections in italics that highlight the questions that section raises for such an adaptation.*

Chapter Two

2. The new approach to improving safety

The literature contains a recurrent theme of two pictures of human error: the traditional and the newer systems picture. This repeated imagery emphasises the fundamental shift in the framework for understanding and reducing error from the traditional 'person-centred' to the new 'system-centred'(Reason, 1990b). While even quite recent publications refer to the systems approach as 'new' (e.g. (Dekker, 2006), the first steps in its development can be seen in the 1940s, for example, American Air Force research (Fitts and Jones, 1947). Since then, it has been further developed, with some significant changes and disputes about theoretical assumptions, and been adopted in more and more areas of safety management.

This chapter begins with an account of the methodology of the traditional inquiry, the type of solutions it produces, and some of the reasons for its appeal. It then examines its limitations and presents the alternative approach of seeing human error in its wider systemic context. A more detailed account of the theoretical foundations of the systems approach is provided in the subsequent chapter.

2.1 The traditional approach

When an accident or tragedy occurs, it has been a standard and understandable response for people to ask why and how it happened. This leads to an inquiry into its causation. These generally have at least two aims: first to learn where in the process the error occurred with a view to learning how to prevent a recurrence and, secondly, to judge whether any individual(s) were responsible and to allocate blame.

In analysing why something happened, we follow a chain of events back into the past. The behaviour of the complex, real world is a continuous dynamic flow which we divide into discrete events in order to identify causes and create explanations. In developing an explanation of an unexpected outcome, we look at previous events and seek to pick out ones which were unusual as the focus for our analysis of why the process did not proceed as expected. 'Unusual' can only be defined against a background knowledge of what is normal in those circumstances. The search back through the history of the adverse outcome has no logical end point. Whatever events are identified, it is always possible to ask further questions about why or how they occurred. To avoid infinite regress, some 'stop-rule' is needed. Rasmussen studied many different types of investigations and identified three main reasons for stopping:

> Stop-rules are not usually formulated explicitly. The search will typically be terminated pragmatically in one of the following ways: (a) an event will be accepted as a cause and the search terminated if the

> causal path can no longer be followed because information is missing; (b) a familiar, abnormal event is found to be a reasonable explanation; or (c) a cure is available. The dependence of the stop-rule upon familiarity and the availability of a cure makes the judgement very dependent upon the role in which a judge finds himself. An operator, a supervisor, a designer, and a legal judge will reach different conclusions.
> To summarise, identification of accident causes is controlled by pragmatic, subjective stop-rules. These rules depend on the aim of the analysis, that is, whether the aim is to explain the course of events, to allocate responsibility and blame, or to identify possible system improvements to avoid future accidents (Rasmussen, 1990).

The backtracking continues until the investigator reaches a causal factor that is familiar to him/her. Technical malfunction is one such familiar cause and leads to solutions that improve the accuracy or reliability of technology. The second, even more common, is human error; a worker, usually one whose action was close in time to the accident, made an error of omission or commission that played a significant causal part in the subsequent accident. *If only* this worker had taken the correct action *then* the accident would not have occurred.

In the traditional inquiry into accidents and mishaps, human error has readily been accepted as an adequate explanation. It has been identified as the major cause of adverse outcomes in a range of industries and activities. There is a remarkably consistent finding of 70-80% of inquiries across a range of industries and professions attributing tragedies to human error: in anaesthesia (Cooper et al., 1984, Wright et al., 1991), in aviation (Boeing Product Safety Organization, 1993). These inquiries conclude that the key source of error was in human performance: surgeons, pilots, or other members of the workforce made mistakes. Among the myriad causal factors in the complex sequence of events that led to the final accident, the actions of one or more humans are picked out as of crucial causal significance.

> *My own study of child abuse inquiry reports, carried out before learning of this statistic on human error, collected data on the judgments reached by the inquiry teams on whether or not human error had been a significant factor. In 75% of cases, it was cited (Munro, 1999).*

The causal explanation that centres on the human being who took the significant action proximate to the accident offers a familiar and plausible story: the engineer failed to notice the temperature was rising; the surgeon misread the notes. However, attributing the adverse outcome to the human operators who were nearest to it in time and space:

> ultimately depends on the judgement by someone that the processes in which the operator engaged were faulty and that these faulty processes led to the bad outcome. Deciding which of the many factors surrounding an incident are important and what level or grain of analysis to apply to those factors is the product of *human* processes

> (social and psychological processes) of causal attribution. What we identify as *the cause of the incident* depends on with whom we communicate, on the assumed contrast cases or causal background for that exchange, and on the purposes of the inquiry (Woods et al., 1994).

When the traditional inquiry is satisfied with human error as the explanation then it produces solutions based on that conclusion. If adverse outcomes are seen as due to human error, it follows that solutions are sought that target human performance in order to improve it and reduce the risk of error. It looks as if a basically safe system has been corrupted by poor human performance so that improving safety requires reducing or controlling that human element. Three broad approaches to improving performance are identified in the literature:

Psychological strategies: using punishments or rewards to shape performance and encourage people to operate at a higher level. Naming, shaming, and blaming those deemed responsible gives a powerful message to others about the need to improve the standard of work. Management, too, can introduce strategies that monitor and reward a greater attention to complying with accepted standards of good practice.

Reducing the autonomous role of humans as much as possible. In engineering, increased automation, replacing human operators with machines, has been a major solution. Even where individuals cannot be removed from the process, there are ways to reduce their scope for independent decision making by introducing increasingly detailed protocols to provide a step-by-step account of how the operation should be carried out.

Increasing monitoring of the workforce to ensure they are complying with rules and guidance. As protocols take a more significant part in practice, there is a corresponding need to check that they are being followed and to intervene and punish if deviations are observed.

> *In child welfare, all of these solutions are visible. The many public inquiries that have aroused strong media attention have typically named and shamed key personnel. The key social worker in the care given to Maria Colwell before her death in 1973 had her photograph in all the main newspapers at the time, accompanied by critical, personal comments. Her equivalent in the more recent Victoria Climbie inquiry suffered a similar fate plus losing her job, being struck off as a social worker and being placed on the list of those deemed unsafe to work with children in any capacity. The level of public criticism is seen by many as a contributory factor to the current recruitment and retention problems being experienced in many child welfare systems.*
>
> *Reductions in the role of the individual are most prominently evidenced in the increased amount of guidance and prescription of the minutiae of practice. A more recent phenomenon has been the introduction in some agencies of actuarial tools for making risk assessments or*

decisions about interventions, reducing the worker's role to inputting the data but giving the tool the task of computing the data and producing a conclusion.

Increased surveillance is apparent in the growing requirement to document actions and to work towards targets and performance indicators set by government. The increased surveillance arises not just from the recommendations of abuse inquiries but also from the wider movement in government towards an audit society (Power, 1997) in which, for a mixture of social, economic and political reasons, there is a stronger demand on public sector services to be transparent and accountable.

Consequently, one would expect that the repercussions of this traditional approach discussed below, both good and bad, should be visible within child welfare services.

2.2 Why the traditional approach seems so plausible

The bias caused by hindsight

A major reason for considering that the traditional investigation produces a satisfactory explanation in blaming human error is the bias inherent in examining past events with the benefit of hindsight.

(Woods et al., 1994), list four reasons why people are ascribed blame for bad outcomes:

> First, operators are available to blame. Large and intrinsically dangerous systems have a few, well-identified humans at the sharp end. Those humans are closely identified with the system function so that it is unlikely that a bad outcome will occur without having them present. … The second reason that human error is often the verdict after accidents is that it is so difficult to trace backward through the causal chain of multiple contributors that are involved in system failure. It is particularly difficult to construct a sequence that goes past the people working at the sharp end of the system. … The third reason that human error is often the verdict is paradoxical: human error is attributed to be the cause of large system accidents because human performance in these complex systems is so good. Failures of these systems are, by almost any measure, rare and unusual events. … Hindsight is the fourth major reason that human error is so often the verdict after accidents. Studies have consistently shown that people have a tendency to judge the quality of a process by its outcome: information about outcomes biases their evaluation of the process that was followed.

Once we know the outcome of a process, we have a tendency to over-estimate what could have been anticipated with foresight (Fischhoff, 1975). The significance of new information, such as an observed change in the conditions, looks so clear to those who know how significant it turned out to be. But to judge the operator as faulty for not spotting its significance at the time requires an investigation that examines many factors in the evolving sequence such as the background of normal events, routine practice, attentional demands, and strategic dilemmas.

When discussing error, one should make a distinction between *outcome* failures and mistakes in the *process*. Process errors are defined in relation to some standards as to how the work should be carried out. Since decisions in practice have to be made in conditions of uncertainty, there can be no guarantee that they lead to the desired outcome. The distinction between process and outcome is important because a good decision process can lead to a poor outcome and a poor decision process can be followed by a good outcome.

People have a tendency to judge a process by its outcome. A large body of research demonstrates the power of hindsight bias. In a typical study, two groups are asked to evaluate human performance in cases with the same descriptive details but with the outcomes randomly assigned to be either bad or neutral. The consistent finding is that those who have been told the outcome was bad judge the same decision or action more harshly than those who were told the outcome was neutral. For example, Caplan and his colleagues carried out such a study with anaesthetists asked to judge the quality of care the patient received. Those told the outcome was bad consistently rated the care as substandard while the other group considered it up to standard (Caplan et al., 1991).

Other research shows that once people know the outcome, they tend to view it as having been more probable than other possible outcomes. They also tend to be unaware of the influence their knowledge of the outcome has on their judgment about its probability. The research conducted by Fischhoff (Fischhoff, 1975) demonstrates these tendencies working in conjunction. Participants were told about some event and asked to judge the likelihood of its happening compared with other possible outcomes. Some of the participants were also told what actually happened but told to ignore this information in judging the probability or, in some cases, told to respond as they thought others without the outcome information would respond. Those who knew the actual outcome judged it more likely than those who did not.

Hindsight bias has significance in understanding how we investigate adverse outcomes:

> It appears that when we receive outcome knowledge, we immediately make sense out of it by integrating it into what we already know about the subject. Having made this reinterpretation, the reported outcome now seems a more or less inevitable outgrowth of the reinterpreted situation. … In trying to reconstruct our foresightful state of mind, we

> will remain anchored in our hindsightful perspective, leaving the reported outcome too likely looking (Fischhoff, 1982: 343).

The evidence about bias in our reasoning gives ground for caution about our reasoning as we investigate back through time to identify the cause of an outcome. Woods et al sum up the two key lessons it provides:

- Decisions and actions having a negative outcome will be judged more harshly than if the same process had resulted in a neutral or positive outcome. We can expect this result even when judges are warned about the phenomenon and have been advised to guard against it.
- Judges will tend to believe that people involved in some incident knew more about their situation than they actually did. Judges will tend to think that people should have seen how their actions would lead up to the outcome failure. Typical questions a person exhibiting the hindsight bias might ask are these: 'Why didn't they see what was going to happen?' It was so obvious. Or, 'How could they have done X? It was clear it would lead to Y' (Woods et al., 1994: 180).

Efforts have been made to find a way to combat hindsight bias since it can produce a significant distortion in error investigations. Telling people about it and warning them to be on their guard does not seem to be effective (Fischhoff, 1977, Wood, 1978). The method that seems to have had the most success is to ask people to consider alternative outcomes to the actual one, for example, asking them to explain how things might have gone otherwise (Hoch and Lowenstein, 1989). Another strategy with some success is to ask people to list reasons both for and against each of the possible outcomes (von Winterfelldt and Edwards, 1986, Fraser et al., 1992).

The fundamental attribution error

Another bias in our reasoning that is said to make the person-centred explanation so plausible is the fundamental attribution error (Reason and Hobbs, 2003: 15, Dekker, 2002). We tend to explain other people's behaviour differently from our own. When looking at our own actions, we are very aware of the context in which they were performed and so we tend to offer explanations in terms of that context. When explaining other people, we are most aware of their behaviour itself and so focus on our explanations of that rather than the context. Therefore, we attribute a greater causal role to the individual's actions than to the local circumstances in which it took place.

> When we see or hear of someone making an error, we attribute this to the person's character or ability – to his or her personal qualities. We say that he or she was careless, silly, incompetent, reckless or thoughtless. But if you were to ask the person in question why the error happened, they would almost certainly tell you how the local

> circumstances forced them to act in that way. The truth of course lies somewhere in between(Reason and Hobbs, 2003: 15).

The risk of distorted thinking due to hindsight bias and the attribution error exists in a systems investigation too, since the investigators are normal human beings, and so they need to be remembered and conscious efforts taken to minimise their impact.

Responsibility and blame

A final factor that makes the person-centred explanation appealing is that it fits with widespread cultural assumptions about human psychology and morality that focus on individual responsibility for one's actions. The train driver who went through a red light is seen as responsible for taking the wrong action and, as such, is culpable.

Many authors writing on the purpose of pubic inquiries have commented on how they often meet the public's need for someone to blame. Identifying the person to blame is psychologically satisfying. It is also reassuring because it suggests that one bad apple has caused the error and the system is basically sound.

There is a serious issue here for systems investigators that is further discussed in Chapter Eight when discussing learning organizations and reporting systems. Although they aspire to a no-blame approach and aim to investigate why the person took the action that is now seen as wrong, there are significant legal and moral obstacles to a full blown dismissal of the concept of personal responsibility. Moreover, practitioners themselves would be loath to be re-framed as cogs in a machine with no individual control and hence no responsibility for their contribution.

In other disciplines, there are some criteria for differentiating between imperfect and negligent actions. In medicine, for example, the Bolam case (1957) established the standard by which clinical practice should be judged, i.e. the standard of a responsible body of practitioners operating in that particular type of medical situation. The law also sets a standard of 'material contribution' in relation to causation. If the clinician can demonstrate that his or her actions were limited by resource constraints, the causal responsibility is diminished or eliminated (Eastman, 1996: 149).   However, in child welfare, there will be difficulties in developing comparable criteria because of the lack of a well tested and shared knowledge base in many aspects of practice. One option that seems frequently used in inquiries is to judge whether the individual's actions complied with procedures and to blame him or her if there are significant deviations. The problems with this are (a) procedures only at best provide outline advice on what to do and so, in many cases, one would want to conclude that procedures had been followed but the practice was still faulty; (b) the procedures are generated from the wisdom of experienced workers but there is no empirical evidence to show that they are 'right' (leaving aside for the *moment all the problems with saying any element of practice is right or wrong).*

2.3 The limitations of the traditional approach

These solutions all, at first glance, look very sensible and, indeed, it is possible to find numerous examples where they have contributed to substantial advances in safety management. Psychological pressure on the workforce prioritises the importance of safety. Automation has, in many cases, replaced fallible humans with highly reliable machines. Automated cash tills in shops have a far higher accuracy than human operators and, when mistakes occur, they are most often due to errors in the human component of the task e.g. inputting the data. Procedures and protocols try to capture the wisdom of the most expert and make it available to all operators, reducing the chances of error occurring due to ignorance. Surveillance improves the organization's knowledge of what is going on and so increases the possibility of spotting weak points in the process that need further attention.

The traditional approach has clearly made a significant contribution to improving safety and the quality of practice. It remained dominant during the nineteenth and early twentieth century but began to be questioned primarily because of its empirical limitations: accidents were still occurring; its solutions while eradicating some problems were not sufficient to reduce the risk of accident to an acceptable level. The Institute of Medicine in the US, in arguing for the need to change to a systems approach, offers statistics on the rate of error in medicine in the US healthcare system (Institute of Medicine, 1999: 1). Between 44,000 and 98,000 people die in hospitals each year as a result of medical errors that could have been prevented. Even the lower estimate makes death from preventable medical errors more common than deaths from car collisions, breast cancer, or AIDS. Beyond the toll in deaths, there are many other significant adverse outcomes: the cost of extra care, the loss of income, long term disability, loss of trust, and loss of satisfaction. Reducing the rate of error is clearly a major avenue to improving patient care.

In child welfare, it is difficult to find comparable estimates. Child deaths from abuse or neglect are relatively rare (estimated about a 100 per year in the UK) but many, if not most, of these were not known to child welfare services. Even in the known cases, human error by a practitioner may not have contributed. However, the adverse outcomes for children and parents from less destructive errors are harder to estimate but are likely to be significant.

Besides a concern that the traditional approach was not producing good enough solutions, there was increasing concern that the solutions it produced were themselves contributing to new forms of error.

The pressure to find a more effective approach was experienced most strongly in high risk industries, where mistakes caused the loss of life not just the loss of industrial output. Foremost among the researchers for improving safety was the American military forces, in particular the Air Force. This research started to reveal how features of the work environment made human error more or less likely.

The following is an early example of how investigators began to extend their focus from the individual operator to features of the work environment. In 1947, Fitts and Jones demonstrated how features of World War Two airplane cockpits systematically influenced the way in which pilots made errors (Fitts and Jones, 1947). The direct cause of the accident was often that pilots confused the flap and landing-gear handles because these often looked and felt the same and were located next to one another. In the typical accident, a pilot would raise the landing gear instead of the flaps after landing, damaging the propellers, engines and air-frame. Such errors were shown to be not random individual failings but systematically connected to features of people's tools and tasks. The mistakes became more understandable when researchers looked at the features of the world in which the pilots were operating, and analysed the situation surrounding the pilot. The potential to operate the wrong control was built into the design and error was particularly likely when there were a lot of tasks demanding the pilot's attention, as there were when coming in to land. This underlying cause was called the 'root cause' or 'latent error' (Reason, 1990b). During the war, pilots had developed a short-term solution of fixing a rubber wheel to the landing-gear control and a small wedge-shaped end to the flap control. This basically solved the problem by making it easier for pilots to select the right handle and therefore reducing the rate of human error. In this example, human error is seen as relative to the context, not as an absolute, and this is a key feature of the systems approach.

Rasmussen challenged the existing orthodoxy of seeing human error, in itself, as a satisfactory explanation of an adverse outcome:

> it …[is] important to realize that the scientific basis for human reliability considerations will not be the study of human error as a separate topic, but the study of normal human behaviour in real work situations and the mechanisms involved in adaptation and learning (Rasmussen, 1985: 1194).

Woods et al also stress the importance of studying people in context not in isolation:

> To study error in real-world situations necessitates studying groups of individuals embedded in a larger system that provides resources and constraints, rather than simply studying private, individual cognition (Woods et al., 1994: 20).

The arguments against the traditional, person-centred approach combine two strands: highlighting the defects in the solutions it has produced and emphasising how a systems approach offers a more detailed, accurate, and constructive theoretical framework within which to investigate and explain error and safety.

Solutions as sources of error

A common criticism of the traditional approach is that the solutions introduced as a response to the traditional accident inquiry can unintentionally contribute to the causation of future accidents. Increased psychological pressure, automation, proceduralisation, and surveillance all alter the context in which people operate and, in some cases, alter it in undesirable ways.

'Alarm overload' is one example of how solutions can unintentionally create new problems. Traditional inquiries have frequently led to the introduction of alarm mechanisms to ensure that operators are alerted if a mechanical failure occurs. Each alarm has the sensible function of alerting operators to the existence of a specific problem in the equipment or process. When a single red light comes on, operators can notice it and know they need to look in a particular part of the equipment to find the cause. The problem arises when several light up at once, as happens in a crisis. Then operators quickly become overwhelmed and confused by the alarm system so that they are unable to interpret what is going on and deal effectively with the crisis.

This was a major factor in the Three Mile Island nuclear power plant disaster where the system almost went into a catastrophic meltdown. Not only did the operators fail to take the right steps, their interpretation of the alarms led them to take actions that made the crisis worse. In the control room, the search for safety had created a control panel with more than 600 alarm lights. In this major crisis, so many lights came on that the operators were unable to interpret them accurately (Wildavsky, 1988). In their efforts to make the system safer, engineers had inadvertently changed the nature of the tasks required of the operators so that they were a challenge to human cognitive abilities and so harder for the operators to carry out well.

A similar example of alarm overload contributed to the Apollo 12 space shuttle accident. One space controller in mission control made the following comment after the Apollo 12 spacecraft was struck by lightening:

> The whole place just lit up. I mean, all the lights came on. So instead of being able to tell you what went wrong, the lights were absolutely no help at all (cited in (Woods and Hollnagel, 2006: 88).
>
> *In child welfare, the closest example to this seems to be the social/political shift since the 1960s to asking (or, in mandatory reporting systems, requiring) people to raise the alarm if they have even a minor concern that a child may be suffering from abuse or neglect. The lowering of the referral threshold and the consequent sharp rise in reports alleging child abuse to the child protection services, the majority of which do not lead to professional services being offered and so are, in some sense, false alarms. The massive*

impact of the rise in referrals has left many agencies struggling to identify the few serious cases in a timely manner.

Whittingham presents the British railway system as an interesting case example to illustrate the cumulative effect of numerous traditional inquiries on the overall system (Whittingham, 2004: 167). On the railways, there has always been a strong link between individual human error and a major accident. Each inquiry into an accident sought solutions to reduce the potential impact of the individual operator's mistake by increasing automation and introducing safety systems that, for example, prevent two trains being on the same piece of track. Whittingham criticises the impact of this approach which has had two major effects:

1. The railways evolved into a pseudo-military organization, where safety depended upon orders, rule, and regulations being obeyed largely without question.
2. Safety improvement measures had to wait for an accident and were directed towards preventing the same or similar accident happening again…. This approach tended to address the direct cause rather than the root cause of accidents. Since the root cause had not been eliminated, it was almost guaranteed that a similar accident would soon occur with a different direct cause (Whittingham, 2004: 168).

Safety measures to prevent trains colliding have taken the form of improving the signals available to drivers to inform them whether the track ahead is clear or not. The signals indicate danger (red) proceed with caution (amber) or clear (green). Within the cab, there are additional means to alert the driver to the state of the signals. An Automatic Warning System (AWS) has been in operation since the 1960s. This alerts the driver to signals and he is required to cancel them and decide what appropriate action, if any, to take. The AWS helps driver vigilance; it does not automatically control the movement of a train. The driver needs to notice and react to the AWS warning.

In the Southall rail accident in 1997, the train had been fitted with both AWS and the newer Automatic Train Protection (ATP) that prevents a train going through a red signal. However, on the day of the accident, neither system was operating. No rules had yet been established for the new ATP system and the rules that had evolved for the ATS system permitted a train to run without it if it was defective, with the proviso that it should being repaired as soon as practicable. This was to minimise disruption to the train service. The direct cause of the accident was the driver's actions but the root cause or latent error was an operational rule that allowed the train to be driven with both train protection systems disabled. Another root cause was the management decision that the ATP system that prevents driver error should not become the norm because of cost.

The Paddington train crash in 1999 illustrated the dangers of a safety system based on reacting to accidents. In this crash, a key feature was that one train driver went through a red light (leading to his death and that of thirty others). The inquiry discovered that several drivers had previously made complaints about the poor visibility of this signal because of its positioning but, until there was an accident, no action was taken to examine or solve the problem.

> Detailed examination of accidents in these systems consistently shows that the ability of sharp end practitioners to defend against failure in these cases depended directly and indirectly on a host of blunt end factors rather than on the isolated "error" of human practitioners. (Cook and Woods, 1998).

Reason provides a pithy list of weaknesses of the person-centred approach:

- They 'firefight' the last error rather than anticipating and preventing the next one.
- They focus on active failures rather than latent conditions.
- They focus on the personal, rather than the situational contributions to error.
- They rely heavily on exhortations and disciplinary sanctions.
- They employ blame-laden and essentially meaningless terms such as 'carelessness', 'bad attitude', 'irresponsiblity' – even in Total Quality Management.
- They do not distinguish adequately between random and systematic error-causing factors.
- They are generally not informed by current human factors knowledge regarding error and accident causation (Reason, 1997: 126).

> The more safety researchers have looked at the sharp end, the more they have realized that the real story behind accidents depends on the way that resources, constraints, incentives, and demands produced by the blunt end shape the environment and influence the behaviour of the people at the sharp end (Reason, 1997).

2.4 A systems framework

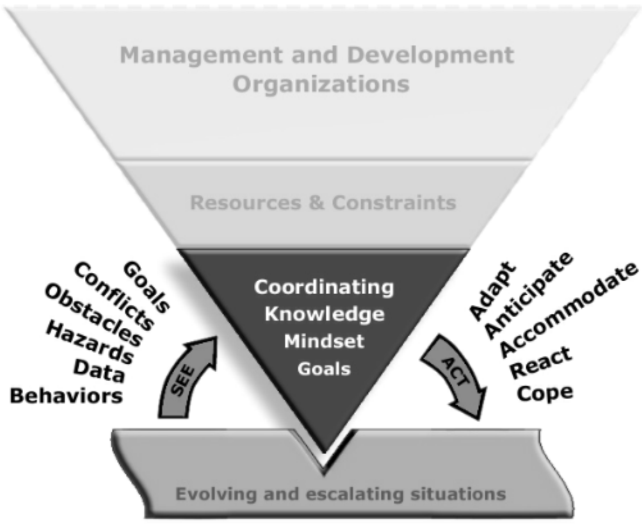
Defining a system

In place of a person-centred approach, the systems approach places individuals within the wider system(s) of which they are part. A system is a set of interrelated elements. They can be either natural (e.g. the human body) or man-made (e.g. a child protection service). In the latter, the elements are usually both people and technology – a socio-technical system. They are organised for a particular purpose or purposes. They have boundaries (some people are seen as inside the child protection service, others are outside) but the boundaries are permeable – there is movement across them so the system is responsive to outside forces. Each person within the system may be a part of other systems as well. Teachers have a role in the child protection system but also in the education system.

Woods and Hollnagel summarise the three basic premises of a systems approach:

> 1. *Interactions and emergence:* a system's behaviour arises from the relationships and interactions across the parts, and not from individual parts in isolation.
> *2. Cross-scale interactions* (multiple levels of analysis) [i.e. any system is itself part of a larger system and composed of subsystems]: understanding a system at a particular scale depends on influences from states and dynamics at scales above and below.
> *3. Perspective:* how the parts of a system and levels of analysis are defined is a matter of perspective and purpose (Woods and Hollnagel, 2006: 7).

A commonly used image of an organizational system is of a triangle, with the front line worker at the sharp end influenced by all the elements making up the rest (often referred to as the blunt end). The following diagram from (Woods and Hollnagel, 2006: 10) illustrates the structure.



Errors are predictable not random

A key premise in the systems approach is that errors are not random and unpredictable (and hence very hard to avoid) but generally take a limited number of forms so that it creates the possibility of altering the system to reduce them.

> Not only are errors much rarer than correct actions, they also tend to take a surprisingly limited number of forms, surprising, that is, when set against their possible variety. Moreover, errors appear in very similar guises across a wide range of mental activities. Thus it is possible to identify comparable error forms in action, speech, perception, recall, recognition, judgement, problem solving, decision making, concept formation and the like (Reason, 1990b: 2).

The goal of systems investigations is to build up understanding of how errors are made more or less likely depending on the factors in the task environment. This allows innovations that maximise the factors that contribute to good performance and minimise the factors that contribute to error. In re-designing the system at all levels to make it safer, the aim is 'to make it harder for people to do something wrong and easier for them to do it right' (Institute of Medicine, 1999: 2).

The example cited earlier of re-designing the structure of the flap and landing gear handles in a plane shows how the solution involved making it easy to distinguish the two so that it was harder for the pilot, preoccupied with all the tasks required at landing, to make a mistake and confuse the two.

Increasing complexity

The traditional approach is deemed to offer too simplistic a model of the work environment. This defect has been magnified in recent decades by the growing *complexity* of current socio-technical systems in which accidents occur. Rasmussen points out that, in the early days of engineering, it was a 'fairly well structured and bounded science. Technical designs could be verified and tested by quantitative models and controlled laboratory experiments' (Rasmussen, 1990: 449). It was, for example, possible to test the safety of a steam locomotive as a thermodynamic process and establish that it was, in principle, safe. How the locomotive later performed in use brought in other factors and accidents occurring in that domain were investigated by non-engineers. At this simpler stage of development: 'the analysis of industrial systems could be reasonably well decomposed into problems that could be treated separately by different professions, and specialists in human factors were primarily involved in attempts to match the interface between people and equipment to human characteristics.

However, as socio-technical systems have become more inter-connected and complex, this simple division is no longer tenable. The functioning of one element has a causal impact on the functioning of others. The role of the human operator has also changed. Many of the tasks previously performed by people are now done by machines and the human role is to supervise their functioning and intervene if something goes wrong. The pace of change is a third factor to consider: 'in a traditional work setting, the slow pace of change led to the evolution of fairly stable work procedures and it was easy to define human error with reference to normal practice' . This growing complexity leads to the need for:

> An explanation of accidents in terms of structural properties of integrated large-scale systems rather than as isolated links or conditions in a linear causal chain of events (Rasmussen, 1990: 450).

Recognition of the multi-factorial nature of causation has highlighted the importance of identifying where in the system the causal factor lies. An important distinction that has been referred to earlier is between two types of error: direct and root causes or active and latent errors. Reason's work on this distinction is the most widely cited.

> *Active errors*, whose effects are felt almost immediately, and *latent errors* whose adverse consequences may lie dormant within the system for a long time, only becoming evident when they combine with other factors to breach the system's defences. In general, active errors are associated with the performance of the 'front-line' operators of a complex system: pilots, air traffic controllers, ships' officers, control room crews and the like. Latent errors, on the other hand, are most likely to be spawned by those whose activities are removed in both time and space from the direct control interface: designers, high level decision makers, construction workers, managers and maintenance personnel (Reason, 1990b: 173).

Reason sums up the role of the front line operator in the wider causal network:

> Rather than being the main instigators of an accident, operators tend to be the inheritors of system defects created by poor design, incorrect installations, faulty maintenance, and bad management decisions. Their part is usually that of adding the final garnish to a lethal brew whose ingredients have already been long in the cooking (Reason, 1990b: 173).

Reason, expressing what has become a widely held view, makes a powerful case for focusing on latent more than on active errors in investigations. He describes latent errors as 'resident pathogens' that dwell in the system and that can combine with external factors to produce errors. He argues:

> Accidents … do not arise from single causes. They occur through the unforeseen concatenation of several distinct factors, each one necessary but singly insufficient to cause the catastrophic breakdown. This view leads to a number of general assumptions about accident causation:
> (i) the likelihood of an accident is a function of the total number of pathogens (or latent failures) resident within the system;
> (ii) the more complex, interactive, tightly coupled [i.e. tasks are tightly coupled if an error in the first task significantly increases the probability of an error in the second] and opaque [i.e the operator has not been kept informed of the processes going on in a machine] the system, the greater will be the number of resident pathogens;

(iii) the higher an individual's position within an organization, the greater is his or her opportunity for generating pathogens;
(iv) it is virtually impossible to foresee all the local triggers, though some could and should be anticipated,. Resident pathogens, on the other hand. can be assessed, given adequate access and system knowledge;
(v) it therefore follows that the efforts of safety specialists could be directed more profitably towards the pro-active identification and neutralisation of latent failures, rather than at the prevention of active failures, as they have largely been in the past.

The word 'complexity' is ambiguous. For some, it has its common usage meaning of intricate or consisting of many parts. For others, however, it is linked to complexity theory and nonlinear dynamics. This difference is discussed further in the next chapter in discussing the disputes in the theoretical foundations.

Seeing people as understandable

The key difference between the traditional and the systems inquiry is that the former is satisfied with the explanation that the accident was caused by human error while the latter sees human error as the starting point for inquiry.

A primary concern is to understand the 'local rationality' – how the situation looked to the practitioner so that the action chosen looked the sensible thing to do. Woods points out that 'no practitioner intends to make mistakes' (Woods, 2003: 27). The 'critical' points and decisions are clear with hindsight but looked different with foresight, given the demands and resources the worker had.

When considering how people acted, it is also important to have a realistic idea of human capabilities. The work on human cognitive factors, discussed in Chapter Five, aims to inform our understanding of what standards are likely to be achieved. Simon's concept of 'bounded rationality' provides a framework for studying the limitations on our cognitive skills (Simon, 1996). People use their knowledge to pursue their goals but have finite capabilities and so simplify the tasks they face in order to make them achievable. So they are rational when viewed from the locality of their knowledge, their mindset, and the multiple goals they are trying to balance.

The aim in a systems inquiry is to reconstruct the practitioner's point of view. Analysis is made of the factors that influence human performance, exploring how limited knowledge, a limited and changing mindset, and multiple interacting goals shape the behaviour of people in evolving situations. As a result of such analysis, it becomes possible to look for solutions that help the practitioner improve. For example, factors that have been found to be significant are ways to help workers, first, activate relevant knowledge, secondly, shift attention among multiple tasks in a rich changing data field, and thirdly, recognise and balance competing goals (Woods, 2003: 29).

The task of understanding local rationality is easily compromised by over-confidence. Woods and Cook warn of the need to avoid 'the psychologist's fallacy'. The phrase was originally coined by the nineteenth century psychologist William James and refers to the fallacy that occurs:

> when well-intentioned observers think that their distant view of the workplace captures the actual experience of those who perform technical work in context. Distant views can miss important aspects of the actual work situation and thus can miss critical factors that determine human performance in the field of practice'. … Understanding technical work in context requires (1) in-depth appreciation of the pressures and dilemmas practitioners face and the resources and adaptations practitioners bring to bear to accomplish their goals, and also (2) the ability to step back and reflect on the deep structure of factors that influence human performance in that setting (Woods and Cook, 2002 139).

Turner and Pidgeon highlight a failure to understand local rationality as a source of weakness in traditional inquiries since it leads to recommendations based on a false picture of the context of the accident:

> All of these recommendations [from inquiries into major tragedies] diverse though they may be, have in common the following; that they are concerned to deal with the problem which caused the disaster as it is now revealed, and not to deal with the problem as it presented itself to those involved in it before the disaster. The recommendations are designed in general to deal with the well-structured problem defined and revealed by the disaster in question, rather than with the ill-structured problem existing before the disaster' (Turner and Pidgeon, 1997: Section 4.4.8.)

People as the source of safety

The systems approach revises the role of the human operator in the process in an even more radical way. People are seen not just as a source of error (albeit an understandable or predictable error) but as a source of safety; it is their intelligence and adaptability that is able to identify and intervene when processes are going wrong.

Rasmussen made an early case for re-thinking the role of practitioners within systems (Rasmussen, 1986: 14). He pointed out that they do not just follow rules because rules alone are not enough in complex situations to fully determine what should be done. In reality, practitioners make the system work successfully as they pursue goals and match procedures to situations under resource and performance pressure; human work at the sharp end is to 'make up for gaps in designers' work'. He identified the following factors that are likely to contribute to accidents by reducing the front line workers' ability to be adaptive and create safety:

1. Removing or undermining the practitioner's ability to be resilient in the face of unanticipated variability.
2. Producing change, especially change on top of change already underway.
3. Providing them with multiple goals that interact and conflict.

Woods et al sum up the fundamental difference in the two approaches to inquiries, sketching two mutually exclusive world views:

> (1) erratic people degrade a safe system so that work on safety is protecting the system from unreliable people;
> (2) people create safety at all levels of the socio-technical system by learning and adapting to information about how we can all contribute to failure (Woods et al., 1994).

2.5 Conclusion

The shift from the traditional person-centred approach to investigating accidents to a system-centred approach represents a fundamental change in how causation is theorised. In the traditional inquiry, the front line worker who took the action that was closest in space and time to the accident has tended to take the blame unless there is evidence of technical failure that compromised his/her ability to perform the task correctly. In a systems inquiry, human error is taken as the starting point for investigation. Why did the operator consider this action was sensible? The causes of the error are looked for not just within the skills and knowledge of the individual operator but also in the many layers of causal factors that interacted to create the situation in which the operator functioned.

Woods et al summarise the key findings on understanding human error:

- the context in which incidents evolve plays a major role in human performance;
- technology can shape human performance, creating the potential for new forms of error and failure;
- the human performance in question usually involves a set of interacting people;
- the organizational context creates dilemmas and shapes trade-offs among competing goals;
- the attribution of error after-the-fact is a process of social judgment rather than an objective conclusion (Woods et al., 1994: 4).

Chapter Three

3. Theoretical foundations

Error investigation approaches have been rooted in empirical science and include theoretical assumptions about how to study the natural and social worlds. Much of the literature is concerned with the details of how to conduct an accident investigation and improve safety, whether in a traditional or systems mode. In these practical textbooks, there is little discussion of the philosophical assumptions implicit in the methodologies advocated The greatest attention to the theoretical foundations is found in the relatively recent literature where authors are arguing that a systems approach needs to recognise, as the natural sciences have, that the world is not deterministic. In the process of discussing the implications, they detail how this changes the fundamental assumptions of the preceding models.

In this chapter, the philosophical viewpoints are grouped for into three categories presented in historical sequence: the early school of investigation that was person-centred, the subsequent shift to a systems approach, and the more recent challenge to systems investigators to accept the indeterminism arising from the complexity of systems. It has to be stated, though, that this involves considerable simplification of a prolific field of practice; not all in any one category would have accepted all the assumptions listed therein and the changes in methods of investigation were more varied in practice than this chronology suggests.

Philosophical assumptions have practical consequences. In safety management, different philosophies lead to differences in whether investigators aspire to being neutral and objective or recognise that their interpretations of evidence are, to some degree, value and theory laden so that another investigator might have noted different factors as significant; whether they look for universal laws that have general application or for context-specific findings; whether they study mental processes or focus on measurable aspects of behaviour; whether they take the individual or the system as the key unit of study; and ultimately whether they aspire to creating a safe system where errors do not occur or a system that is good at detecting and learning from error.

3.1 Person-centred investigations

The man who is credited with starting the study of safety in science is (Heirich, 1931). In his widely-read textbook, he states his commitment to a scientific approach which he explains as “to find governing laws, rules, theorems … these exist. They are stated, described, illustrated. After understanding comes application.” Heinrich worked in the insurance industry and, since accidents led to insurance payouts, finding ways to reduce accidents was of great financial interest. His methodology is summarised by (Wallace and Ross, 2006: 22, Spencer et al., 2006) as “to study the various accident reports sent with insurance claims with a view to discovering the laws of accident causation, which could then be the basis for the laws of accident prevention". Heinrich spelled out the basic assumptions he was adopting:

> Whatever the situation or circumstances may be, it is still true, and as far as may be seen will remain true, that an accidental. injury is but one factor in a fixed sequence … that this event can occur only because of either a positive or a negative unsafe action of some person or the existence of an unsafe physical or material condition; that this unsafe action or condition is created only by the fault of some person; and that the fault of a person can be acquired only through inheritance or through environment (Heirich, 1931: xii).

Such a view of scientific study takes Newtonian physics as the paradigm case. Key assumptions of this form of inductivism are a realist view of the world, that the truth of observed facts can be objectively determined, and that scientists develop theories by observing the world in a neutral manner, and generalising from their observations to develop abstract laws. In the highest form, these are expressed in mathematical terms like Newton's laws of thermodynamics. Another key aspect of the methodology is reductionism: progress is made by reducing complex phenomena to their constituent parts and studying those and the causal relationships between them. Hence the focus on organizations as comprising a set of individuals with links between them rather than as a system or social group.

Heinrich's model of accident causation, which is clearly person-centred, is what has been termed 'the domino theory'. Like a row of dominoes falling, there is one root cause that triggers the next, and so on, until the accident happens. The process is fully determined and there is an identifiable first step in the chain.

One criticism of this view is the claim to scientific objectivity in identifying some factor as the root cause. (Wallace and Ross, 2006: 26), critique an example Heinrich himself offered to support his view. He described the case of an employee in a woodworking plant who fractured his skull as a result of a fall from a ladder. The ladder, which did not have nonslip safety feet, had been positioned at the wrong angle. Heinrich obtained evidence from the man's family that justified the belief that "reckless and wilful tendencies have been inherited". The accident was due factors internal to the employee. As Wallace and Ross point out, it is not self-evident that the man's personality is the 'root cause'. The failure of management to provide ladders with nonslip safety feet contributed. The faulty positioning might be due to inadequate training. If indeed the man's personality was a problem then management contributed to the accident by employing an unsuitable person for the job. To isolate one factor as the 'root cause' as if its identification were a neutral, scientific process is to ignore the importance of the investigator's own priorities, assumptions, or values in deciding on which chain of causation to focus.

Another criticism of the root cause concept is that any cause can also be seen as a consequence of prior causes and so there is no obvious point at which to stop investigating the chain of events. The empirical work by Rasmussen on stop rules, cited in the preceding chapter, provides evidence of the ways that investigators reach a decision to stop exploring further but, while these are all reasonable, they cannot be considered neutral or objective.

Within this person-centred approach, the dominant psychology was initially the orthodoxy of the time: behaviourism. In their efforts to model themselves on the natural sciences, behaviourists avoided study of mental processes as unobservable and hence unscientific. They sought explanations by reference to observable behaviour alone, excluding reference to the beliefs, intentions or motives of actors or the meaning of their actions. This led to theories explaining actions as shaped by features in the world, with particular stress on the importance of positive and negative re-enforcers. In the 1950s and 60s, this school of psychology was gradually displaced by cognitive psychology which included study of mental processes as significant variables in the causal chain leading to human actions. This shift was reflected in safety engineering both in the person-centred and, later, in the systems approach.

The objective, neutral view of science came under severe attack in the middle of the twentieth century when studies done by philosophers such as Paul Feyerabend (Feyerabend, 1975) and Thomas Khun (Kuhn, 1970) showed how scientists did not begin their study of phenomena with a open mind, making observations that were objectively true, but brought to that study a range of assumptions, beliefs and values. The assumption that there could be theory-neutral observations of the world as a solid foundation to knowledge was undermined by Karl Popper's arguments (Popper, 1963). Rather, it was argued, people to some degree construct the world they experience as a social, interactive, process. The question of to what degree reality is a social construct has since then been a major dispute. Putting the debate very simply, few would now claim that either natural or social science can achieve what Putnam calls 'a God's eye view of the world' (Putnam, 1978), an objective description that is true independent of time, location or the knower's beliefs. However, some people claim that reality is wholly a social construct so that all claims to truth are relative to a particular time and culture. Others consider that reality might exist but we have no access to it other than at a step removed through representing it as something else. Others argue that our perceptions of the world result from the combination of the concepts and beliefs that people bring to it, arising from their social environment, and the experiences the world triggers in our sensory skills so that empirical experience places some limit on what we can construct.

3.2 The shift to a systems-centred approach

The impetus to change in safety engineering seems to have come primarily from the failure of the person-centred approach to provide adequate results rather than from any philosophical unease with its theoretical foundations. There does not seem to be simply one coherent methodology within the broad school of a systems approach. Many of the methodological assumptions of the former approach were carried over. The key difference is in the range of factors analysed in theorising the causal pathway that led to the outcome.

Peter Reason, a major theorist in this field, and his co-author Hobbs, summarised the arguments against considering human error an adequate explanation in a recent publication:

> It seems obvious that a human error must have human origins. Some individual – or, quite often, a group of individuals – went wrong, and, in so doing, endangered the safety of the operation and the lives of third parties. The temptation, then, is to home in upon the individual psychological factors immediately preceding the making of an error and to do whatever seems necessary to prevent their recurrence. Such measures usually take the form of disciplinary action, writing another procedure, blaming, shaming and retraining. But this is to miss two important points about human error. First, errors are inevitable. Everyone makes them, but no one actually intends them to happen. Second, errors are consequences not just causes. They do not occur as isolated glitches in people's minds. Rather, they are shaped by local circumstances: by the task, the tools, and equipment and the workplace in general. If we are to understand the significance of these contextual factors, we have to stand back from what went on in the error-maker's head and consider the nature of the system as a whole. … systems are easier to change than the human condition – particularly when the people concerned are already well-trained and well-motivated (Reason and Hobbs, 2003: 9).

Dekker (2005) provides a summary of the main features of the systems approach. Reductionism is still the key strategy of study: reducing complex phenomena to their constituent parts and explaining them in terms of the interactions of those parts:

> In human factors and system safety, *mind* is understood as a box-like construction with a mechanistic trade in internal representations; *work* is broken into procedural steps through hierarchical task analyses; *organizations* are not organic or dynamic but consist of static layers and compartments and linkages; and *safety* is a structural property that can be understood in terms of its lower order mechanisms (reporting systems, error rates and audits, safety management function in the organizational chart, and quality systems) (Dekker, 2006: x).

He goes on to summarise three major assumptions: deconstruction, dualism and structuralism which he defines in the following way:

> *Deconstruction* is the assumption that a system's functioning can be understood exhaustively by studying the arrangement and interaction of its constituent parts. Scientists and engineers typically look at the world this way. Accident investigators deconstruct too. In order to rule

> out mechanical failure, or to locate the offending parts, accident investigators speak of ‘reverse engineering’.

The second assumption of *dualism*: ‘means that there is a distinct separation between material and human cause – between human error or mechanical failure.’ Investigations look for potentially broken human components: was the operator tired, drunk, inexperienced, or ‘fit for duty’? Human error is then explored in terms of whether the operator was fit for duty. The dualist division of mechanical and human error has become so commonplace we hardly recognise the assumptions implicit in it. Yet, many accidents arise from a combination; changing the design of the mechanical part can reduce the rate of human error: ‘the error is neither fully human, nor fully engineered’ (Dekker, 2006: 7). The dualist assumption contrasts with the alternative approach of locating the causes of accidents in ‘normal’ people doing ‘normal’ work, epitomised by Perrow’s concept of ‘normal accidents’ (discussed below).

The third characteristic of the engineering model is structuralism: ‘the language we use to describe the inner workings of successful and failed systems is a language of structures’ (Dekker, 2006: 4). Organizations, for example, are described in terms of a ‘blunt’ end and a ‘sharp’ end (Reason, 1990b).

In place of Heinrich’s domino model of accident causation, the dominant image here in systems centred investigations has been Reason’s ‘Swiss cheese’ model. He conceptualises a system as having a series of defence layers to detect and prevent error but each of those layers may have holes in them.

> The necessary condition for an organizational accident is the rare conjunction of a set of holes in successive defences, allowing hazards to come into damaging contact with people and assets. These ‘windows of opportunity’ are rare because of the multiplicity of defences and the mobility of the holes (Reason, 1997: 11).

This model of investigation has flourished for decades and is responsible for many of the lessons now generally adopted. However, since the 1990s, it has been increasingly challenged as inadequate for taking the study of errors any further.

3.3 Complex systems approach

I have coined the name ‘complex systems approach’ to distinguish recent developments within the overall framework of a systems approach that challenge the fundamental assumptions of the earlier systems model. Still sharing the aspiration of using the methods of the natural sciences, this group points out that developments in those sciences have fundamentally altered the theoretical foundations.

The nature of causality

The most significant change has been our understanding of the nature of causality. In the positivist and, to some degree, in the empiricist approaches, causality is seen as deterministic. The causal chain is linear in nature – if A occurs then B occurs - allowing prediction and control, at least in principle. This deterministic view of the world was dominant in physics until the early twentieth century when developments in quantum mechanics proved that the world was probabilistic in nature. This change in the concept of causation has gained widespread acceptance in the rest of the natural sciences. If at all, deterministic explanations are now seen as possible only in small, local applications and even then with a ceteris paribus proviso. The philosopher Ian Hacking has commented "the erosion of determinism constitutes one of the most revolutionary changes in the history of the human mind' (Hacking, 1987: 54).

An assumption of determinism in safety management fits well with modernity's project of the rational control of society through planning. The linear systems approach assumes that a system can be made safe: with sufficient care and attention, the various latent errors can be rectified so that, in theory at least, it becomes possible to design a safe system. The ideal model of a safe organization requires top-down control, specifying how all the constituent elements should interact. It is assumed that, if enough data is known, it should be possible to predict what interactions will be caused. Again, in principle at least, it should be possible to develop a set of procedures that covered all aspects of the task. However, complex systems have emergent properties. An emergent property is one which arises from the interaction of lower-level entities, none of which show it. If you mix blue and yellow paints, the colour green 'emerges'. There is no green in the original blue and yellow paints and we cannot reverse the process and divide the green back into the two other colours. This raises questions about how far it is possible to predict what will happen as elements in a system come together. Factors that, on their own, are safe may become unsafe as they form a system with others. The earlier example of alarm overload demonstrates such a problem. The designers of each alarm system did not predict what would happen if their product interacted with several others.

This emergent model of accident causation conflicts with Reason's Swiss cheese model because, for Reason, a latent error (or hole in a defensive layer) is a constant error but, in the emergent view, an error may 'emerge' from the combination of other factors none of which is necessarily a latent error in Reason's meaning of the term but only in combination with other factors does the potential for error 'emerge'.

Moving from a deterministic view of the universe to a probabilistic one has radical implications. Wallace and Ross sum up the difference: "instead of a deterministic view in which certain causes inevitably lead to certain effects, in the probabilistic view, certain causes may or may not lead to effects with differing degrees of probability" (Wallace and Ross, 2006: 17). Consequently, the ideal of a top-down control that can prescribe every action lower in the system is questioned.

The importance of the social

A consequence of viewing complex systems as, to some degree, unpredictable is that it strengthens the need to find out what went on in the run-up to an accident in a more detailed manner. Specifically, more attention is paid to the social dimension of socio-technical systems.

The study of human factors in error investigations had predominantly used cognitive psychology which paid little attention to social interactions. According to Gardner (Gardner, 1985) a more or less explicit decision was made in cognitive science to leave culture, context, history and emotion out of the early work. These were recognized as important phenomena, but their inclusion made the problem of understanding cognition very complex. In safety management, this policy was followed through for the most part but this focus on what goes on individuals' heads has been challenged on the grounds that it is the actual performance that needs to be studied, that is the actions of the person in a context:

> The common way of thinking about performance implies that there is a set of internal or mental process that determines what people see and do … [but] … even if the information processing model of human behaviour was correct – which, by the way, it is not, - this approach means that one is looking for a hypothetical intervening variable, rather than for the more manifest aspects of behaviour. … There is, however, an alternative, namely to focus on the characteristics of performance … we should not be overly concerned with the performance of (for example) the pilot per se, but rather of the pilot + aircraft, in other words, the joint pilot-aircraft system (Dekker and Hollnagel, 2004).

The significance of moving from a focus on the individual operator is particularly relevant when, in Chapter Six, the study of the human/artefact relationship is discussed. Where older traditions have conceptualised this as an individual using a tool, systems theorists argue that the focus should be on the performance of the joint human + artefact system.

Studying people in their context highlights the importance of the influence of culture on meaning. This is apparent in making sense of 'local rationality' – how people are making sense of their working environment and making decisions. If meaning is seen as, at least in part, the product of social negotiations within a group then it is important to know how those involved in the error were constructing the work environment in which they were making decisions and acting. It is argued by some that, to fully understand how the world looks to practitioners in the field, we have to recognise:

> how people use talk and action to construct perceptual and social order; how , through discourse and action, people create the environments that in turn determine further action and possible assessments, and that constrain what will subsequently be seen as acceptable discourse or rational decisions. We cannot begin to

> understand drift into failure without understanding how groups of people, through assessment and action, assemble versions of the world in which they assess and act (Dekker, 2006: xiii).

Gano, in a book detailing how to carry out a root cause analysis makes a similar point: ‘start by defining the problem and understanding that everyone’s perception is slightly different (and valid). It is a major fault to assume that we all share the same picture of reality’ (Gano, 2003: 60).

The work needed to understand the local rationality of operators is stressed by (Woods and Hollnagel, 2006: 52) who warn against the ‘psychologist’s fallacy’, discussed earlier. They add, as a corollary, a warning from (Law and Callon, 1995: 281):

> Avoid assuming that those we study are less rational or have a weaker grasp on reality than we ourselves. This rule of method, then, asks us to take seriously the beliefs, projects and resources of those whom we wish to understand.

The importance ascribed to understanding the ‘local rationality’ of the operators, how they made sense of the situation at the time and why they acted as they did, has also led to an interest in the social construction of what counts as an error and what is normal practice. Earlier approaches had taken it as possible to define ‘error’ and ‘failure’ in objective terms that are agreed by all in the organization. Cognitive psychology contains a mechanistic model of the human mind as an information-processing system separate from the phenomena it processes.

> Such a model severely restricts our ability to understand how people use talk and action to construct perceptual and social order; how, through discourse and action, people create the environments that in turn determine further action and possible assessments, and that constrain what will subsequently be seen as acceptable discourse or rational decisions (Dekker, 2006: xiii).

Accident investigations often focus on whether or not operators complied with procedures as a definition of ‘safe’ practice but, in reality, local cultures often develop in which normal practice deviates from the official manual. This is usually due to difficulties arising in the local context in trying to comply with the procedures. They are conflicting with other aspects of the work. Local, deviant cultures may, in practice, be good methods of improving the safety of the system and minimising the clumsiness or inaccuracy of the official procedures. It cannot be decided in advance that all deviations are ‘errors’.

Studying the system not just the individual

The focus of today’s performance investigations is on a socio-technical system:

> When we refer to a technological system, we invariably think of it in the context of its use. People are therefore always present in some way, not just in the sense of individuals – such as in the paradigmatic notion of the human-machine system – but also in the sense of groups and organizations, i.e. social systems'(Hollnagel and Woods, 2005: 3).

The systems focus also offers a new framework for studying cognition – moving from an individual focus to 'distributed cognition', cognition as it is observed in a system's performance. Hutchins sketches the difference:

> The "classical" vision of cognition that emerged [in cognitive psychology] was built from the inside out starting with the idea that the mind was a central logic engine. From that starting point, it followed that memory could be seen as retrieval from a stored symbolic database, that problem solving was a form of logical inference, that the environment is a problem domain, and that the body was an input device Attempts to reintegrate culture, context, and history into this model of cognition have proved very frustrating. The distributed cognition perspective aspires to rebuild cognitive science from the outside in, beginning with the social and material setting of cognitive activity, so that culture, context, and history can be linked with the core concepts of cognition.
>
> What distinguishes distributed cognition from other approaches is the commitment to two related theoretical principles. The first concerns the boundaries of the unit of analysis for cognition. The second concerns the range of mechanisms that may be assumed to participate in cognitive processes. While mainstream cognitive science looks for cognitive events in the manipulation of symbols (Newell, et.al, 1989), or more recently, patterns of activation across arrays of processing units (Rumelhart, et.al, 1986; McClelland, et.al., 1986) inside individual actors, distributed cognition looks for a broader class of cognitive events and does not expect all such events to be encompassed by the skin or skull of an individual. When one applies these principles to the observation of human activity "in the wild", at least three interesting kinds of distribution of cognitive process become apparent: cognitive processes may be distributed across the members of a social group, cognitive processes may be distributed in the sense that the operation of the cognitive system involves coordination between internal and external (material or environmental) structure, and processes may be distributed through time in such a way that the products of earlier events can transform the nature of later events. The effects of these kinds of distribution of process are extremely important to an understanding of human cognition (Hutchins, 2000).

Improving safety

Although, in the complex systems approach, systems do not have the property of linear predictability and controllability, they are far from chaotic and have mechanisms that can increase the safety of a system. They are self-organising: they have feedback loops so they are able to learn and adapt. Systems for reporting near-miss incidents can be seen as an example of such a feedback loop that permits learning and modifications to take place before a major accident occurs. The importance of such feedback loops is heightened in this approach because it is assumed that interactions lower in the system will be unexpected and senior management or designers cannot predict all that may occur.

An important debate, however, is whether systems can, even in principle, be made safe. Charles Perrow's book on 'Normal Accidents' is an early account of the argument that errors are not eradicable (Perrow, 1999). The odd term *normal accident* is meant to signal that, given the system characteristics, multiple and unexpected interactions of failures are inevitable.

In Perrow's model, the two characteristic features of a normal accident are that there are unexpected interactions of variables that, on their own, are not problematic and that the system is 'tightly coupled'. Two tasks are 'coupled' when there is a link between them such that an error in one task makes it more likely that an error will occur in the second task. Common forms of such links are that the two are closely linked In time, are part of rigidly ordered processes (as in sequence A must follow B), or there is only one way to a successful outcome. When the link makes the second error highly likely if there is an error in the first task, the tasks are described as tightly coupled.: '*tight coupling* allowing little opportunity for mitigation or defence once a fault occurs'. (Perrow, 1999: 5):

A complex system exhibits complex interactions when it has:

- Unfamiliar, unplanned, or unexpected sequences which are not visible or not immediately comprehensible;
- Design features such as branching, feedback loops;
- Opportunities for failures to jump across subsystem boundaries.

Sagan summarises the reasons why accidents will continue to happen:

- Safety is one of a number of competing objectives.
- Redundancy often causes accidents. It increases interactive complexity and opaqueness and encourages risk-taking .
- Organizational contradiction: decentralization is needed for complexity and time dependent decisions, but centralization is needed for tightly coupled systems.
- A "Culture of Reliability" is weakened by diluted accountability.
- Organizations cannot train for unimagined, highly dangerous or politically unpalatable operations.
- Denial of responsibility, faulty reporting, and reconstruction of history cripples learning efforts.
- **Accidents are inevitable in complex and tightly coupled systems.** (Sagan, 1993: 4)

Pool sums up the problems:

> Risk is unavoidable and cannot be eradicated by procedures. Complexity creates uncertainty and uncertainty demands human judgment. And complex systems can be quite sensitive to even very small changes, so that a minor mistake or malfunction can snowball into a major accident. So we cannot write out procedures for every possible situation. Operators trained to go by the book can freeze up when an unexpected situation arises' (Pool, 1997: 250).

When one thinks of the potential harm from an accident in a nuclear power plant, this view of the inevitability of error is somewhat daunting. However, there are critics who point to the existence of high reliability organizations (HROs) that have an amazing record of safety as evidence against such a depressing view. The study of such organizations has led to 'high reliability theory' which

> Proposes that with adequate procedures, safeguards, and design safety nets, accidents can be, for all intents and purposes, avoided. … The reason why most industrial systems are not safe, according to this school, is mainly political: safety costs money, and in a world of limited resources, most firms and organizations are simply not prepared to spend the money on genuine safety improvements (Wallace and Ross, 2006: 174).

High reliability organizations are discussed in more detail in Chapter Seven.

3.4 Conclusion

The study of safety management has developed over the decades, with considerable success in identifying and reducing errors and accidents. The dominant view has been that studies of safety should aspire to the methods of the natural sciences. Since the focus of study has been both on technical and human contributions to error, the social as well as the natural sciences have been drawn upon but the tendency has been to prefer those social science schools than model themselves on the natural sciences. Hence, in psychology, there has been a preference initially for behaviourism and, later, for cognitive psychology with less attention paid, until recently, to humanistic or interpretivist schools of thought. In relation to social factors, perhaps the key difference has been whether attention is given to the social dimension or whether the focus is primarily on individuals. Studying a system is not the same as studying a set of individuals.

Within this broad acceptance of the natural science model, there have been disagreements about what those methods are, reflecting developments in the philosophy of science during recent decades and these are apparent in the shift from a narrow positivist to an empiricist approach. At a philosophical level, there is a strong continuity between the traditional approach and the first versions of the systems approach in that similar methods of study are used but the focus is broadened out from the person to the person-context dyad.

The most interesting conflicts appear within the systems approach. There is an emerging school of thought arguing that the developments in natural science involving complexity and nonlinear dynamics undermine the common assumptions about linear causality and predictability. This change has major implications for the whole framework in which error investigations are conducted. More attention is paid to the social system and to understanding how operators, partially at least, construct the context within in which they operate. Therefore, investigators in this mode have broadened their range of methods to include more sociological and interpretivist methods. This school also raises questions about the overall goal of safety management: whether it is possible even in theory to create a system where errors do not occur or whether the aim should be to design systems that can detect and learn from the inevitable errors that will arise.

> *The increasing interest in studying the social dimension of a socio-technical system by a range of methods looks highly relevant to child welfare where the social dimension is the most prominent although the technical dimension is gaining increasing importance.*
>
> *Child welfare differs significantly from the other fields in which safety management has been used in that the focus of the system is also a social system: the family in its social context. The methods for studying the way that the welfare system operates may be transferable to studying how the family system operates. Indeed, it raises the question about the boundaries of the official system since it does not operate on a passive family in the way that nuclear engineers operate on a nuclear power plant but seeks to work <u>with</u> them, albeit with varying degrees of coercion.*

Chapter Four

4. Defining and classifying error

4.1 Introduction

Since investigating the cause of errors and improving safety by reducing the incidence of errors is the central aim of safety management, the definition and classification of errors are of crucial importance. However, both issues of definition and classification are problematic. This chapter begins with a discussion of attempts to define ‘error’ in various disciplines.

Classification of errors is also necessary in order to bring some order to an otherwise shapeless mass of examples of errors and to facilitate searching for the underlying patterns of errors. There are a number of such taxonomies in existence and examples are examined with the aim of clarifying what issues need to be addressed to create a taxonomy for child welfare services. Taxonomies have a central role to play in creating a near-miss or minor incident reporting system, a learning mechanism that is widely used and valued in safety management. The process of creating such a reporting system and the dangers of creating one without a reliable taxonomy are discussed.

The pilot work undertaken by SCIE on this issue is presented.

When investigating how an incident occurred, classification is needed not just of the errors themselves but of aspects of the work process that are considered to contribute to the causation of the final outcome. Subsequent chapters look in more detail at these areas but, in this chapter, attention is paid to developing a classification system for use when conducting an investigation.

4.2 Defining ‘error’

A first step in analysing errors is to define this key term but there are complexities both is formulating a definition and deciding whether a particular event or action counts as an error. (Hollnagel, 1993) points out a fundamental problem that the term ‘error’ is used in three different senses: as a cause of something, as the action itself, and as the outcome of that action.

A simple definition is:

> An erroneous action can be defined as an action which fails to produce the expected result and/or which produces an unwanted consequence (Hollnagel, 1993).

A common element in definitions is that the action must be accompanied by an intention to achieve a desired result or outcomes for an error to be said to have arisen in an action. ‘This eliminates spontaneous and involuntary action (having no prior conscious thought or intent and including random errors) (Whittingham, 2004: 5). One way of deciding whether or not an error has occurred is to focus on the outcome, not the process by which it was reached, for example:

> Error will be taken as a generic term to encompass all those occasions in which a planned sequence of mental or physical activities fails to achieve its intended outcome, and when these failures cannot be attributed to the intervention of some chance agency (Reason, 1990b: 9).

Such a definition, however, excludes the group of recovered errors, where the operator has noticed and corrected the mistake. ‘It is possible that on a subsequent occasion, the same error might occur and the recovery action not take place or be too late or ineffective, in which case the actual outcome would differ from the intended outcome. It must be true to say that the initiating error was always an error even if the final outcome was as intended’ (Whittingham, 2004: 6). Recovered errors are a very important subset and ways in which systems can recover from errors play a major role in improving safety.

The definition also seems to assume a sufficiently close temporal link between the action and the faulty outcome so that the action’s contribution to the outcome can be confidently asserted. This is a reasonable assumption in some work settings – a driver who pulls out into the stream of traffic without looking clearly contributes to the ensuing crash For practical purposes, it may be necessary to set more proximate goals in error investigations but the rationale for the choice needs to be clearly formulated.

Reason (1990: 1) draws attention to the link between error and correct performance:

> Far from being rooted in irrational or maladaptive tendencies, these recurrent error forms have their origins in fundamentally useful psychological processes. Ernst Mach (1905) put it well: ‘Knowledge and error flow from the same mental sources, only success can tell the one from the other’.

This point highlights the close links between studying human performance to identify success and to identify failure. Definitions of error assume that there is some set of rules or standards, either explicitly or implicitly defined, against which the action can be judged, that there was some kind of performance shortfall relative to these standards, and the operator could have done otherwise.

In the disciplines where safety management originated there is a considerable body of technical knowledge about the tasks. Aviation experts have a good understanding of the mechanical principles involved in getting a plane off the ground, flying it to a destination, and landing it in one piece. Against this background, flying the plane into a mountain and killing all on board is clearly an error. The relative confidence in the abstract principles of the technical dimension, however, has to be tempered with the reality that technical knowledge is not used in a vacuum but in a socio-technical system. Translating the basic theory into functional systems requires performance standards that are far more than technical protocols but deal also with organizational and individual behaviour. Then the definition becomes more problematic and complex systems theorists would question whether a consensus model of good practice at the micro level exists in an organization. This raises the question whether there can be a single definition of error in the organization.

Perhaps what we can conclude about defining error is that, in the abstract, this means some shortfall in performance that deviated from an accepted standard as prescribed by some social group and was not intended by the actor(s) .

> *All of these issues in defining error pose challenges for transferring the concept to the domain of child welfare services where the knowledge base is less developed and people are often working in circumstances where the agreed set of rules or standards leave considerable room for individual interpretation. In child welfare, there are few instances where one can confidently say 'if I do X then the outcome will be Y'. The rules or principles of good practice are less universally agreed, making the judgement that some action or omission constitutes an error highly problematic.*
>
> *Judging practice by the results it achieves is also more problematic in child welfare where the ultimate intended outcomes of the system are often long-term – that the child is safe and develops well.*

4.3 Classification of errors

Scientific progress is made by creating patterns in the mass of data available and using this simplified framework for deeper analysis. If error identification were carried out without any system of classification we would quickly be overwhelmed by the number of individual incidents that we had accumulated. Classifying x as 'like' y involves some theoretical assumption about which aspects of x and y are relevant to your study. We do not look at the world and 'see' the classifications; we create them. Social worker A might be classified as like social worker B in that both are wearing black shoes but we are unlikely to develop such a taxonomy because the colour of practitioners' shoes are not thought to have any causal significance to their performance. There are numerous attempts to classify errors but this section considers to what extent we can learn from other people's taxonomies and to what extent child welfare needs to develop its own system, albeit using ideas from other people's systems.

The two key arenas in which classification of error is used is in conducting investigations of particular incidents and in creating a reporting system to learn about minor errors or near misses from which lessons can be learned to reduce the risk of big errors.

Wallace and Ross (2006) discuss the purpose of classification and how a system is devised:

> The role of classification systems is to organize and display specific, empirical data in ways that enhance understanding. … Understanding the theoretical framework underlying a classification system is necessary to fully comprehend the data. … Taxonomists must ask themselves four questions before they begin their work. What do we want to classify? What kind of thing is it? Why do we want to classify it? What sort of classification is appropriate? (Wallace and Ross, 2006: 94-95).

When these questions are applied to the issue of classifying errors, they raise further questions about what type of error you are interested in, whether it is clear what is or is not an error, whether there needs to be evidence of an adverse outcome before it is deemed error, what is the standard against which you are judging an action a mistake.

Wallace and Ross argue that most currently used accident investigation systems and minor event reporting systems have been built in the wrong order or back to front:

> In other words, the methodology for gathering the data was set up, then the database to order the data, and then the taxonomy to order *these* data. … Certainly one needs *some* raw data at the beginning (which might be a very small number of reports, discussions with process engineers, or observations of plant activity). As soon as information starts to come in, discussions should commence with the staff who have to use the database to create a workable and database-specific taxonomy which may of course use lessons learned from other taxonomies, but the fundamentally unique aspects should never be ignored (2006: 60).

If a reporting system is set up without a foundation of a clear taxonomy then two key problems will ensue. First, criteria for making a report will be inadequately specified so that the reliability and validity of the reports received will be low. One worker's understanding of a significant near miss may vary radically from another's, making deeper analysis of the reports difficult. The second key problem is that, once a number of reports have been received, it will be difficult to start to pick out clusters because inadequate attention has been paid to identifying the relevant clusters. If the work of developing a taxonomy starts only at this point, then it is likely to be found that the existing reporting criteria and the data collection methods do not meet the needs of the taxonomy. Therefore, the sensible approach is to begin with the taxonomy and avoid amassing large quantities of ill-structured data that is hard to use.

Vincent's criticisms of the British medical incident reporting system suggest that the taxonomy was inadequately formulated before this system was put into operation. He lists its weaknesses as:

- No standardized, operational definition of incident,
- Coverage and sophistication of local incident reporting systems varies widely,
- Incident reporting in primary care is largely ignored,
- Regional incident reporting systems undoubtedly miss some serious incidents and take hardly any account of less serious incidents,
- No standardized approach to investigating serious incidents at any level,
- Current systems do not facilitate learning across the NHS as a whole (Vincent, 2006: 59).

Those who develop a reporting system without a clear, agreed, theoretically grounded, taxonomy seem to be adopting a naïve inductivist view of scientific method: the view that a scientist begins by amassing a collection of objective, neutral facts and only then starts to group the data into generalisations and theories. This overlooks the importance of the theoretical assumptions in the scientist's mind that influence what is observed, what is picked out as worth recording, and which aspects of it are described and which ignored.

Bowker and Starr's (2000) text on classification and its consequences includes a study of the development of a taxonomy in nursing that demonstrates the influence not only of theoretical assumptions in making decisions about how to classify phenomena but also the moral and political influences on the process.

> *In the highly politically and ethically charged world of child welfare, the importance of recognising the role of these factors is crucial in discussing how to classify errors in practice.*

Bowker and Starr offer sound advice:

> Seemingly pure technical issues like how to name things and how to store data in fact constitute much of human interaction and much of what we come to know as natural. We have argued that a key for the future is to produce flexible classifications whose users are aware of their political and organizational dimensions and which explicitly retain traces of their construction (2000: 326).

> *The differences in the underlying assumptions in classifications in different fields of investigation limit the scope for simply lifting a classification system developed in, say, aviation to a radically different area such as children's services. However, there does seem potential*

for drawing on some of the more abstract classifications of errors, because of their wide relevance, and on those formulated health care, because of the similarity in the two fields of work.

Genotypes and phenotypes

Within the set of errors, many authors use Hollnagel's (1993, Dekker and Hollnagel, 2004) broad taxonomy of genotypes and phenotypes. These terms taken from biology distinguish the internal blueprint or set of instructions for building a living organism (the genotype) from the outward, surface manifestation (the phenotype). In the error literature, the terms are used to differentiate (a) the surface description (phenotype of failure) and (b) underlying patterns of systemic factors: generic or genotypical patterns "that express regularities about task or psychological or human-machine system factors that shape the possibilities for erroneous actions or assessments (Woods et al., 1994: 34). For example, an individual case may show that a worker formed a view of what was happening and failed to alter it despite the evidence against it becoming apparent. This is an example of the genotype 'fixed mindset', a common lapse in human reasoning that the system needs to minimise by incorporating mechanisms to help challenge people's thinking and increase the possibility of revising judgments appropriately.

Woods et al (1994: 13) criticise many reporting systems and investigations for describing errors only in terms of phenotypes 'They do not go beyond the surface characteristics and local context of the particular episode'. Much of the research on human factors and systems has aimed to identify the genotypes of error, the more general features, the deeper patterns that contribute to accidents and may appear in slightly different manifestations in several places.

Hollnagel's Predictive Error Analysis Technique

To illustrate the difficulties in taking a classification system 'off the shelf', I am including Hollnagel's (1998) classification of phenotypes of error that has been widely used and found to have good inter-user agreement but which would take considerable development to adapt to a child welfare system.

Planning errors
Incorrect plan executed
Correct but inappropriate plan executed
Correct plan, but too soon or too late
Correct plan, but in the wrong order

Operation errors
Operation too long/too short
Operation incorrectly timed
Operation in wrong direction
Operation too little/too much

Right operation, wrong object
Wrong operation, right object
Operation omitted
Operation incomplete

Checking errors
Check omitted
Check incomplete
Right check on wrong object
Wrong check on right object
Check incorrectly timed

Retrieval errors
Information not obtained
Wrong information obtained
Information retrieval incomplete

Communication errors
Information not communicated
Wrong information communicated
Information communication incomplete

Selection errors
Selection omitted
Wrong selection made.

> *Elements of this table look possibly more relevant to child welfare than others. The sub-division of communication errors, for example. provides a useful starting point for analysis of what is often reported in inquiry reports as a single error. However, before this classification system could be employed in child welfare, considerable work would need to be done in reaching consensus on what is a 'correct' plan, which is the 'right' operation, or 'wrong selection made. In the process of undertaking such work, the researchers would probably start to form their own classification of significant points in the work process where error was likely, expressed in the language that is familiar to workers in this field and therefore more likely to be able to gain consistent usage. Existing classifications can be consulted for ideas but it seems a false hope to think that child welfare can avoid doing the basic development work itself to form a classification that is theoretically relevant to understanding and improving practice.*

Vincent's classification of medical errors

Vincent (2006) discusses the problem of using engineering definitions of error in health care and raises similar concerns about transferring the concept across to areas of uncertainty. Diagnosis in medicine has many similarities to assessment in children's services and he offers a classification of diagnostic errors that might be a useful starting point for analysing error in child and family assessments. With a caution that the categories should not be treated as rigidly distinctive, he outlines three categories of diagnostic error:

1. *Errors of uncertainty* (no-fault errors), for example, where the presentation of the disease is unusual and the doctor misinterprets the available signs and symptoms or where little is known of the disease and how to diagnose it.

2. *Errors precipitated by system factors,* (system errors), for example, overlooking a fracture in the emergency department because a radiologist was unavailable to do the tests or delay in diagnosis due to faults in co-ordinating care.

3. *Errors of thinking or reasoning* (cognitive errors), for example, mistakes arising from poor knowledge, faulty history taking, or making a decision prematurely before considering more possibilities (Vincent, 2006: 89).

Woods and Cook (Woods and Cook, 2001) offer some useful insight into the challenges of identifying the genotypes of error in health services. They point out that it draws on literature from outside health care itself but from the domains of psychology and engineering, using language and concepts that are not unique to medicine. They suggest that three key patterns of human performance that can be transferred to the health care setting are:

- patterns in human judgment;
- patterns in communication and cooperative work;
- patterns in human-computer cooperation.

They note that searching for patterns will ‘require a process of going back and forth between two different perspectives (health care and human performance) in a partnership where each party steps outside their own area of expertise to learn by looking at the situation through the other’s perspective’ (Woods and Cook, 2001: 5).

> *The points made by Woods and Cook seem highly relevant to how to begin the task of identifying patterns of practice in child welfare services. Each of their three patterns, listed above, resonate strongly with issues that are repeatedly raised in studies of child welfare practice so there is potentially a great deal to be gained by drawing on the learning acquired in related areas of human performance.*

SCIE’s work on developing defining and classifying errors

The Social Care Institute for Excellence (SCIE) has carried out pilot work on developing a way of defining and classifying errors. Learning from the health model, Phase I identified the need to develop a language in safeguarding work that would capture the range of adverse incidents that occur to include those where no harm is caused (near misses). The tentative proposition of the equivalent of a ‘patient safety incident’ in health for social care was a ‘safeguarding incident’:

> A ‘safeguarding incident’ results in harm or potential harm due to professional agencies’ failure to keep a child safe, rather than from neglect or abuse by family members, for example. Like a ‘patient safety incident’, a ‘safeguarding incident’ also covers near misses that have the potential to lead to serious harm but have been prevented or have occurred but no serious harm was caused. A ‘safeguarding incident’ refers to an action combined with a potential or actual negative outcome (Bostock et al., 2005: 17).

This study also took the first steps toward developing a grading system of the seriousness of safeguarding incidents: – no harm, harm and life threatening injury, or death. It was noted however that refinement of the model would need to “incorporate both immediate and long-term harm, as well as appreciating differences in how social workers and service users understand harm” (2005: 19) and would, therefore, require “considerable consultation and further work” (2005: 19).

Subsequently, however, and influenced by discussions in Health, paying particular attention to Vincent’s criticisms of the patient reporting safety discussed earlier, the researchers came to question how one could identify a safeguarding incident (i.e. know if it was caused by professional failures) *before* conducting any analysis. This indicated the need to prioritise analysis over reporting systems in order to develop taxonomies. Hence, this became the premise of Phase II of the work.

Chapter Five

5. Human reasoning

5.1 Introduction

The next three chapters explore different aspects of the complex interactions leading to error. This chapter looks at factors relating to people's reasoning abilities. The next chapter looks at the relationship between people and the range of artefacts or tools they use in the process of their work. The third chapter looks at the wider organizational context of practice.

In relation to human reasoning skills, the entire body of psychological research might be seen as of potential relevance. I have focused on those aspects of research that are discussed in the engineering literature and seem to have relevance to child welfare.

5.2 Human cognitive capacity

Studies of human reasoning have taken two main forms, either normative/prescriptive or naturalistic/descriptive (Hammond, 1996). The normative school has focused on measuring human performance against the standards set by formal logic and probability theory and identified how humans tend to fall short of those standards. Their studies are, for the most part, conducted in laboratories using artificial cognitive tasks. Kahneman, Slovic and Tversky (1982) are three leading researchers in this school and they pioneered the study of heuristics and biases – the cognitive shortcuts humans tend to take to reduce tasks to a manageable size and the predictable biases that these produce Knowledge of these biases is valuable for organizations in that they indicate where humans need assistance to recognise and avoid them and some are discussed later in this chapter.

> *My work on common errors of reasoning identified in child abuse inquiries shows how these biases are manifested in child protection work* (Munro, 1999).

The alternative school, the naturalists, have preferred to study cognition 'in the wild' (Hutchins, 1995) in the natural work environment where cognition takes place. Their aim is to describe, not prescribe, what humans do. This school has considerable relevance for safety management since the systems approach aims to understand human cognitive capacities not with a view to chastising workers but with a view to designing tools and systems so that the cognitive tasks demanded of the workers are generally within their competence. A well designed task is one that it is easier for a human to do correctly than to do incorrectly.

Bounded and constrained rationality

A fundamental premise that shapes the whole systems approach to understanding the role of the human operator in error causation is that studies need to be based on a realistic idea of human capacity. Therefore the descriptive approach has more direct relevance although the prescriptive approach is valuable in offering useful guidance on how reasoning tends to deviate from the ideal unless specific strategies are used to check it. The aim is to design the work environment so that maximises good reasoning. Capacity is affected by two key factors: the functioning capacities of the human brain and the constraints placed on individual choice.

On the first point, a generally accepted concept is Simon's 'bounded rationality'. Herbert Simon (1955) working on economic theory, introduced this concept: the recognition that people have limited processing power in their brains and therefore cannot live up to the optimal image then dominant in theories of rationality. The human brain places limits on our capacity to formulate and solve complex problems and in processing (receiving, storing, retrieving, transmitting) information.

> The task is to replace the rationality of economic man with a kind of rational behaviour that is compatible with the access to information and the computational capacities that are actually possessed' (Simon, 1955: 99).

A core idea is that conscious attention is *the* scarce resource. Consequently, we resort to a number of heuristics to help us simplify reasoning to a feasible level. Our 'bounded rationality' leads us to resort to simplifications when faced with complex decisions: we prefer to rely on 'a drastically simplified model of the buzzing, blooming confusion that constitutes the real world' (Simon, 1955: xxix).

Besides being limited by brain capacity, people are also limited by their work environment. Reason (1997: 61) emphasises how constrained operators' behaviour is within high risk industries compared with everyday life. Their actions are governed by managerial and regulatory controls of two main kinds:

- *external controls* made up of rules, regulations, and procedures that closely prescribe what actions may be performed and how they should be carried out. Such paper-based controls embody the system's collective wisdom on how the work should be done.
- *Internal controls* derived from the knowledge and principles acquired through training and experience.

This point serves to emphasise the importance of judging the individual's behaviour within its organizational context, not judging him or her against the picture of some ideal, unconstrained individual who was freely able to choose any course of action.

Human reliability analysis

A major and highly technical area of study is of normal performance in work settings to provide data for 'human reliability analysis': analysis of task performance and the factors that increase the risk of error. This involves precise counts being made of human actions and errors in precise circumstances. The textbooks on these types of studies are highly technical and complex so it is perhaps easiest to show what type of research they do by offering an example of the kind of findings they produce. Whittingham summarises the type of factors that shape performance. The more difficult task will be less reliable than a simple one.

> There are, however, two aspects of task complexity which need to be considered. These are:
> 1. the difficulty of the task,
> 2. the capability of the person carrying it out (Whittingham, 2004: 47).

Typical performance shaping factors are presented as follows:

Factor	Effect on human error rate
Task familiarity	Familiarity with the task being carried out will reduce human error rates at least down to the point where boredom may take over. Methods of improving vigilance may then have to be adopted.
Available time	Complex tasks may require more time to carry out than simple off-the-cuff tasks and if this time in not available, due perhaps to external constraints, then human error rates will increase.
Ergonomics	It is important that the design of the equipment adequately supports the needs of the operator and if it does not then this will cause human error rates to increase.
Fatigue and stress	Depending on the level of stress, the effects can vary from the merely distracting to the totally incapacitating. At the same time, ther is an optimal level of arousal and stimulation necessary to maintain vigilance (also see Task familiarity above). Fatigue affects the ability to perform tasks accurately and also increases the influence of other PSFs mentioned in this table.
Attentional demands	Human error rates for a single task carried out on its own may increase considerably if other tasks or distractions compete for attention. Human beings are not particularly good at multi-tasking.
Availability of plans and procedures combined with level of training	Complex tasks require both experience (see below) and information to be completed successfully. Information should be presented in the form of easily understood procedures and plans which in some cases must be memorized depending upon the task complexity.
Operator	Human error rates will depend upon whether the person

experience	carrying out the task is a novice or an expert, and this must always be taken into account in assessing Human Error Probability. To some degree, this will depend upon task complexity.

Dekker (Dekker, 2002: 125) offers a table setting out factors in the task and the coping resources that facilitate an examination of the mismatch between them in particular instances.

Problem demands	**Coping resources**
Ill-structured problems	Experience with similar problems
Highly dynamic circumstances: things changing quickly over time	Other people contributing to assessments of what is going on
Uncertainty about what is going on or about possible outcomes	Knowledge or training to deal with the circumstances
Interactions with other people that generate more investment than return (in terms of offloading)	Other people to off-load tasks or help solve problems
Organizational constraints and pressures	Organizational awareness of such pressures and constraints
Conflicts between goals	Guidance about goal priorities
High stakes associated with outcome	Knowledge there is an envelope of pathways to a safe outcome
Time pressure	Workload management skills

Task analysis and error classification

Ways of analysing the cognitive skills needed to carry out a task in a work process are a useful starting point for locating error and tracing its causation. Reason (1990a) offers a classification of the cognitive aspects of tasks that is widely used as the basis for studying error in task performance. Within it, he builds on Rasmussen's (1987) categorisation of the cognitive demands of tasks.

Slips, lapses, mistakes and violations

Reason (2001) distinguishes three groups: slips and lapses, violations, and mistakes.

Slips and lapses happen when a person knows what he or she wants to do but the action does not turn out as intended. They are failures of execution. Slips refer to observable actions, typing the letter S when intending to type D, for example. Lapses are internal events linked to failures of memory – picking up the yellow box, thinking it contains the pencils, when they are in the blue box. Slips and lapses occur usually in routine aspects of a task and are often associated with some form of distraction.

Violations: deliberate deviations from what is considered desirable practice, e.g. deliberately failing to follow a procedural guideline. The element of deliberation distinguishes violations from errors in general where we assume that the operator did not want to err. Sometimes the motivation is malicious but often people deviate from procedures because they believe it will get the job done better or it will save time with no risk of adverse consequence. In a later work, Reason (2001) offers further detail on the varieties of violation:

- routine violations basically cutting corners for one reason or another;
- necessary violations when a person considers that breaking a rule is the only way to get a job done
- optimizing violations are for personal gain.

Mistakes

Mistakes occur when the actions go as planned but the desired outcome is not achieved. Here the error lies in with the mental processing involved in planning, formulating intentions, judging and problem solving. The intention is clear, the action went as intended, but the plan was wrong. In classifying mistakes further, Reason uses Rasmussen's three-fold classification of skill, rule, and knowledge based error. This analysis is useful because, once a task has been classified, it is possible to estimate the reliability of the performance of the task within a range of human error probability. 'As performance moves from skill based through rule based to knowledge based, the reliability tends to decrease' (Whittingham, 2004: 15). The three categories are as follows.

1 *Skill-based errors* that involve slips or lapses, for example when an engineer intends, as a highly practiced action, to press the return key on the computer but presses the delete key instead. Skill based behaviour represents the most basic level of human performance and is typically used to complete familiar and routine tasks that can be carried out smoothly in an automated fashion without a great deal of conscious thought. 'This type of behaviour is generally highly reliable and when errors do occur they are either of a random nature or occur because of interruptions or intrusions into familiar patterns of activity' (Whittingham, 2004: 15). The mechanical process of driving a car, e.g. changing gear, operating the brakes, as an experienced driver is a typical example – but one only has to think back to the days of being a learner driver to realise that it takes experience to learn to need to give so little conscious attention to the task.

2. *Rule-based error*: at a slightly more complex level a person has to look at the situation and classify it into some familiar category. Although the rules may be in written procedures, it is just as likely that they are rules learned from experience or through formal training that are retrieved from memory when carrying out the task. If a situation is mis-classified, this may be called a rule-based error. For example, the

engineer presses the right switches to turn of the engine but made a mistake in deciding that the engine needed to be turned off.

3. *Knowledge-based* errors occur in the most complex cases, when people have to find the answer to a completely new problem. They do so by working out what would happen if they tried particular sets of actions. A person forms a new plan, and makes no slips in executing it, but the plan itself is mistaken and does not yield the desired result. Knowledge based behaviour involves a significant amount of feedback from the situation. However, with experience and learning, what was initially knowledge based may become more rule based.

In many task sequences, behaviour oscillates between the three types of behaviour and research has produced the principle of 'cognitive economy' (Reason, 1990a). This principle is based on limitations in mental processing (i.e. cognitive) resources which restrict their capacity to:

(a) assimilate external information,
(b) understand its implications for the current situation in terms of appropriate actions to produce a desired outcome;
(c) carry out these actions (Whittingham, 2004: 18).

As a result of the limitations in our capacity, it is always more economic and effective, when faced with any slightly complex situation requiring action, to fall back on a set of stored rules that have been acquired from past exposure to similar situations. Consequently, what might have been better treated as knowledge based behaviour is relegated to rule based behaviour:

> When confronted with a problem, human beings are strongly biased to search for and find a prepackaged solution at the RB level *before* resorting to the far more effortful KB level, even where the latter is demanded at the outset. … Only when people become aware that successive cycling around this rule-based route is failing to offer a satisfactory solution will the move down to the KB level take place (Reason, 1990a: 65).

Failure to revise assessments

One of the most persistent and important errors in cognition is people's slowness in revising their view of a situation or problem. Once they have formed a view of what is going on, there is a surprising ability to fail to notice or to dismiss evidence that challenges that picture (Kahneman et al., 1982). Becoming fixated on one assessment despite an emerging picture that conflicts with it becomes a significant source of cognitive error.

Garden path problems are 'a specific class of problems where revision is inherently difficult since early cues strongly suggest [plausible but] incorrect answers, and later, usually weaker cues suggest answers that are correct' Woods and Hollnagel, 2006, p.75). Besides being influenced by the relative strength of cues, the worker is also hampered by the fact that data tends to arrive incrementally not in a single, eye-catching batch. 'The vulnerability lies not in the initial assessment, but in whether or not the revision process breaks down and practitioners become stuck in one mindset or even become fixated on an erroneous assessment, thus missing, discounting or re-interpreting discrepant evidence' (Woods and Hollnagel, 2006: 75).

Vagabonding is a complementary vulnerability when, instead of becoming fixated on one assessment, workers respond to every new item of information and jump incoherently from one item to the next treating each superficially and never developing a coherent, coordinated response (Dorner, 1983).
A suspicious stance is the approach most associated with success. 'Expertise at revision appears to be quite special and seems to depend on taking a 'suspicious stance' toward any data that fails to fit the current assessment, despite the relative prevalence of 'red herrings' (data changes that sometimes are not significant or due to minor factors)' (Woods and Hollnagel, 2006: 77).

What makes a task difficult

Another area of task analysis that offers useful insights is research into what makes a task difficult. Feltovich et al (2004 p.90) summarise the research that has been done under the rubric of Cognitive Flexibility Theory examining learning and performance in medical education. Eleven dimensions of tasks have been identified as making them difficult and requiring mental effort:

1. Static vs. dynamic: are important aspects of a situation captured by a fixed ' snapshot' or are the critical characteristics captured only by the changes from frame to frame? Are phenomena static and scalar, or do they possess dynamic characteristics?
2. Discrete vs. continuos: do processes proceed in discernible steps or are they unbreakable continua?
3. Separable vs. interactive: do processes occur independently or with only weak interactions, or do strong interaction and interdependence exist?
4. Sequential vs. simultaneous: do processes occur one at a time or do multiple processes occur at the same time?
5. Homogeneous vs. heterogeneous: are components or explanatory schemes uniform or similar across a system, or are they diverse?
6. Single vs. multiple representations: do elements in the situation afford single or just a few interpretations, categorizations, etc, or many? Do we need multiple representations (analogies, case examples) to capture and convey the meaning of a process or situation?
7. Mechanism vs. organicism: are effects traceable to simple and direct causal agents, or are they the product of more system-wide, organic functions? Can we gain important and accurate understandings by understanding just parts of the system, or must we understand the entire system to understand even the parts well?
8. Linear vs. nonlinear: are functional relationships linear or nonlinear (that is, are relationships between input and output variables proportional or nonproportional)? Can a single line of explanation convey a concept or account for a phenomenon, or does adequate coverage require multiple overlapping lines of explanation?

9. Universal vs. conditional: do guidelines and principles hold in much the same way across different situations, or does their application require considerable context sensitivity?
10. Regular vs. irregular: does a domain exhibit a high degree of typicality or do cases differ considerably even when they have the same name?
11. Surface vs. deep: are important elements for understanding and for guiding action delineated and apparent on the surface of the situation, or are they more covert, relational, and abstract?

Research has found that when tasks exhibit the latter alternatives in this list, then learners and practitioners show common characteristics:

- they tend to interpret situations as thought they were characterized by the simpler alternatives;;
- their understandings tend to be reductive –that is, they tend to simplify;
- they tend to try to defend their simple understandings when confronted with facts that suggest that the situation is more complex than what they suppose;
- overcoming these defences requires practice, experience and mental effort.

Novices and experts

Discussions of how novices differ from experts offer some interesting insights into the nature of human performance. Reason, in analysing the sources of skill, rule and knowledge-based errors, notes that the first two categories are more predictable than knowledge-based errors because in the latter, they arise from a complex interaction of bounded rationality and incomplete or inaccurate mental models.

> Mistakes at the KB level have hit-and-miss qualities not dissimilar to the errors of beginners. No matter how expert people are at coping with familiar problems, their performance will begin to approximate that of novices once their repertoire of rules has been exhausted by the demands of a novel situation. The important differences between the novice and the expert are to be found at the SB and RB levels. Expertise consists of having a large stock of appropriate routines to deal with a wide variety of contingencies.
>
> There is considerable evidence to show that in skilled problem solving, the crucial differences between experts and novices lie in both the level and the complexity of their knowledge representation and rules. In general, experts represent the problem space at a more abstract level than nonexperts. The latter focus more on the surface features of the problem. (Reason, 1990a: 58).

Woods et al also discuss the distinctive features of an expert. They report that research has shown that mere possession of knowledge is not enough for expertise. It is also critical for knowledge to be organized so that it can be activated and used in different contexts.

Bringing knowledge to bear effectively in problem solving is a process that involves:

* content (what knowledge) – is the right knowledge there? Is it incomplete or erroneous?
* organization – how knowledge is organized so that relevant knowledge can be activated and used effectively; and
* activation - is relevant knowledge 'called to mind' in different contexts? (Woods et al., 1994: 55).

5.3 Conclusion

This chapter has aimed to give an overview of the range of ways of studying human reasoning in a work context. The final phrase is of crucial importance. What is central to a systems approach is that it does not study human reasoning in isolation or in laboratory conditions remote from the noisy, crowded environment of the typical workplace. People are seen as interacting with their environment and being influenced by it in how they reason. Understanding the strengths and limitations of human cognitive capacity is fundamental to designing systems that fit typical human levels of performance and do not inadvertently distort or encourage errors.

The errors of reasoning look very familiar to anyone who has read inquiries into child deaths. However, the more difficult task is to think how the system can be changed to support a higher level of performance. In adapting this literature to child welfare, there is a need for empirical research to study in depth the nature of the cognitive tasks that workers are expected to undertake and to place them in context so that the influence of wider systemic factors can be appreciated.

Chapter Six

6. The worker/artefact relationship

The word 'artefact' is being used here to capture the range of phenomena that are included in the topic. It is not just the worker's relationship with tools, such as screwdrivers or thermometers, but also with the numerous aids that have been designed to improve human performance, such as procedure manuals, assessment frameworks, official documentation, telephones, and computerized databases*.* A key lesson is to understand how the worker and artefact interact and how the resulting performance is shaped by the influence that each has had on the other.

> *In child welfare services, the proliferation of information-processing tools, increasingly computer-based, has been a very significant development in recent years.*

The impact of technology on human performance is complex. As Woods et al (Woods et al., 1994: 163) point out, the conventional view is that new information technology and automation creates better ways of doing the *same* task activities. However, it is more accurate to say that any new technology is a *change* from one way of doing things to another. It alters the tasks that humans are expected to perform and it can, in subtle and unexpected ways, influence and distort the way they can carry out their part of the process.

> The design of new technology is always an intervention into an ongoing world of activity. It alters what is already going on – the everyday practices and concerns of a community of people – and leads to a resettling into new practices (Flores et al, 1988, p. 154).

Designing artefacts to aid us can be seen as a key human characteristic: 'throughout history, humans have designed mechanisms to reproduce and extend the capacity of the human body as an instrument of work (Zuboff, 1988: 8).

> *In child welfare services, the main aim of artefacts tends to be to extend cognitive capacity rather than physical prowess, with computers playing an increasingly major role in everyday practice.*

With the increasing use of information technology, Woods et al (1994) stress that it is particularly important to study how these innovations shape human cognition and action in order to see how design can create latent failures. A later section in this chapter considers the human/computer relationship.

6.1 The impact of technology in high-risk industries

Several authors discuss the range of ways that innovations in technology have created both improvements in safety and sources of danger. The following section draws mainly on the work of Bainbridge (1987), Rasmussen (1988), Reason (1990a), and Perrow (1999), using headings from Reason's discussion.

Changes in technology change the tasks that the human operators are required to perform. The following are important factors that affect human performance.

Systems have become more automated

One of the most remarkable developments has been the degree of automation introduced so that operators have become increasingly remote from the processes that they nominally control. Reason (1990a: 175) cites Moray (1986: 404-405):

> There is a real sense in which the computer rather than the human becomes the central actor. For most of the time the computer will be making decisions about control, and about what to tell or ask the operator. The latter may either pre-empt control or accept it when asked to do so by the computer. But normally, despite the fact that the human defines the goal of the computer, the latter is in control.

At first thought, child welfare services may not appear to be high-technology but, when one considers the range of innovations of the past twenty years, one is struck by the fundamental changes there have been in the work environment. In the United Kingdom, for example, the social worker of the 1970s had what now seems an amazing degree of autonomy in how they worked with families while the social worker of today is located within a web of guidance, formal relationships for working with other professionals, and computerised forms. Indeed, a major consequence of the numerous traditional inquiries that have been held into child abuse deaths has been to increase the rules and the guidance under which practitioners operate and hence to reduce their autonomy. Each inquiry tends to produce a set of recommendations that add a few more rules to the list. However, as the Human Factors Study Group review of safety noted wryly, there is a strong tradition of adding rules in response to mistake but 'investigators introduce new rules without asking why the existing rules were not followed' (Human Factors Study Group, 1993: 15)*.*

Systems have become more complex and dangerous

As a result of the increasing computerisation, high-risk systems such as nuclear power plants and chemical process installations have become larger and more complex. The discussion in Chapter Three of Perrow's (1984) work on normal accidents introduced the concepts that characterise complex systems: complexity of interaction and tightness of coupling. Hollnagel and Woods (2005: 7) identify three key consequences of the growing complexity in systems:

1. It may well be that the number of accidents remains constant, but the consequences of an accident, when it occurs, will be more severe.

2. Increased coupling and interdependence means that malfunctions have consequence far beyond the narrow work environment.

3. The amount of data has increased significantly and this has created a need for better ways of describing humans, machines, and how they can work together. 'The belief that more data or information automatically leads to better decisions is probably one of the most unfortunate mistakes of the information society'.

Systems have more defences against failure

Designers have sought to provide automatic safety devices sufficient to protect the system against all the known scenarios of breakdown. For example, the design of a nuclear power station is based upon the philosophy of 'defence in depth' – several layers of defence. For a catastrophe to happen, a number of apparently unlikely events need to occur in combination during the accident sequence. But such disasters still happen. One of the most obvious reasons is that the safety systems themselves are prey to human error, particularly of the latent kind. We are thus faced with a paradox: those specialised systems designed solely to make a plant safe are also its points of greatest weakness.

Systems have become more opaque

Complex, tightly-coupled and highly defended systems have become increasingly opaque to the people who manage, maintain, and operate them. This opacity has two aspects: not knowing what is happening and not understanding what the system can do. As many accidents have demonstrated, the operators often cannot find what they need to know while, at the same time, being deluged with information they do not want or know how to interpret.

Rasmussen (1988: 3-4) points out the philosophy of 'defence in depth' adds to the problem by making it harder for operators to notice latent errors that do not lead to disaster. Humans are able to operate with a high level of reliability when slips and mistakes have immediately visible effects and can be corrected.

Compare this to working in a system designed according to 'defence in depth' principle, where several independent events have to coincide before the system responds by visible changes in behaviour. Violation of safety preconditions during work on the system will probably not result in an immediate functional response, and latent effects of erroneous acts can therefore be left in the system. When such errors are allowed to be present in a system over a longer period of time, the probability of coincidence of the multiple faults necessary for release of an accident is drastically increased.

An additional problem in an opaque system is that, when something goes wrong in an automated system and human operators have to step in to solve the problem, they do so with little understanding of how the problem evolved and this limits their problem-solving ability.

Just such a problem was created in aviation (LaPorte and Consolini, 1991, Norman, 1990, Pool, 1997). Designers progressively automated the process of flying a plane to the point where the plane was essentially flown by a computer; the pilot spent much of his/her time monitoring the instruments but taking no active role. They only had to step in if something went wrong. Unfortunately, this meant that when they needed to play an active part, they were thrust straight into the middle of the problem with limited knowledge of what had preceded it. They were therefore poorly prepared for diagnosing the problem and solving it. When this latent error designed into the automatic system became understood, the design of the cockpit was altered to make the system more visible so that pilots could stay aware of the automatic processes going on and see how difficulties evolved, thus improving their ability to intervene effectively.

The errors caused by opacity demonstrate the crucial importance of having feedback systems in place so that operators are able to monitor what is happening and to detect and correct errors (1990: 590). It is this safety requirement that drives the interest in developing learning organizations and reporting systems that are discussed in Chapter Seven.

> *In child welfare services, one source of opacity is in actuarial risk assessment instruments where workers input the data, follow the specified calculations, and receive a score. The opacity of the weighting of the variables in the probability calculation means that workers may be informed that families are classified as high risk of abuse but do not know which features of the family were significant in reaching this conclusion. Consequently, the risk assessment instrument does not offer guidance on how to manage the case in order to reduce risk.*

The ironies of automation

Bainbridge (1987) coined the phrase ‘ironies of automation’ and it has been widely quoted. Designers often take the view, discussed in Chapter Two, that humans are the main source of error in the system and need to be eradicated and controlled as much as possible.

> ‘There are two ironies of this attitude. One is that designer errors can be a major source of operating problems. Unfortunately, people who have collected data on this are reluctant to publish them, as the actual figures are difficult to interpret. … The second irony is that the designer who tries to eliminate the operator still leaves the operator to do the tasks which the designer cannot think how to automate. It is this approach that causes the problems … as it means that the operator can be left with an arbitrary collection of tasks, and little thought may have been given to providing support for them (Bainbridge, 1987: 272).

One of the key tasks left to humans in automated systems is to monitor that the automatic system is functioning properly. But this is a task that even highly motivated humans are not good at: they have trouble in maintaining effective vigilance for anything more than quite short periods.

Another key task is to take over manual control when the automatic system fails. However, because the system is automatic most of the time, the operators have little chance to practice so they become de-skilled in precisely those activities that they are needed to do in an emergency. ‘But when manual takeover is necessary something has usually gone wrong: this means that operators need to be rather and less skilled in order to cope with these atypical conditions’ (Reason, 1990a: 180).

6.2 How do artefacts help or hinder human performance?

Collins and Kusch (1998: 119) categorise machines into three types:

1. Tools – can <u>amplify</u> our ability to do what we can already do.
2. Proxies – can <u>replace</u> us.
3. Novelties – can do types of things we could never do, e.g. a fridge can cool items.

They say it is easy to mistake tools for proxies because of the ease with which we ‘repair’ the deficiencies of machines and then ‘attribute’ agency to them. A pocket calculator does not ‘do arithmetic’ (i.e. act as a proxy); it is only a tiny part of arithmetic. A lot is done by the user at the input and output stages. If the skills needed to repair the deficiencies of a tool are ones we use a lot, we tend not to notice we are using them so over-estimate the contribution of the tool.

The distinction between tools and proxies (sometimes called prostheses) is a key one in discussion of how artefacts and humans interact although most artefacts have some characteristics of each. They may indeed be used more or less as one or the other. Hollnagel and Woods (2005: 96) note that the function of an artefact may change as the user becomes more experienced. For the unskilled or unpractised user, an artefact may be a proxy, replacing their lack of competence, whereas, for the more experienced user the same artefact may be a tool that can be used in a flexible way to achieve new goals.

Consider how flexibly a skilled practitioner in child welfare may use an assessment framework when talking to a family while an unconfident student may plod systematically through the categories, impervious to the dynamics of what is happening in the interview.

Conceptualizing the worker/artefact relationship

Ways of conceptualising the worker/artefact relationship reflect the differing theoretical assumptions discussed in Chapter Three. The dualism of the engineering model makes a sharp distinction between the human and the artefact. Although it was recognised that they were both parts of a larger system: 'the distinction between the operator as an intelligent human being and the machine or the process as a technological system was never questioned' (Hollnagel and Woods, 2005: 67)0. However, growing recognition of the interdependence between them and the way that each influences the other has led to a greater focus on their relationship. Hood and Jones (1996) express such a view, describing the close interdependence of a sociotechnical system consisting of people, their social arrangements, and the technological hardware they make and use. 'People and technology interact with each other and over a period, change each other in complex and often unforeseen ways' (1996: 35).

Consequently, it is argued by some that the focus should be shifted from the human and machine as two separate units to seeing them as a joint system and studying how they function as an overall system. The relationship of significance then becomes not that between the human and machine but between the human-machine system and the environment (Hollnagel and Woods, 2005: 66). They give the simple example of a woman and a pair of scissors. Rather than taking the dualist view of the woman interacting with the scissors, she is seen as using them and, if we focus on them as a joint system, we study how well they jointly perform the function of cutting.

This argument illustrates that there are no natural systems in an absolute sense. 'Something is called a system if it is seen relative to an environment, and separated from this by a boundary' (Hollnagel and Woods, 2005: 66).

The effect of framing the human and artefact as a joint system is to draw attention to their function; it highlights the need to be clear about what they are intended to achieve before one can evaluate their level of performance.

This could be of major relevance in child welfare services where there is a rapid increase in the introduction of artefacts but it is not always adequately specified what the artefact/user is meant to achieve. For example, are data collection forms designed to meet the needs of the practitioners who are assessing the needs of the family or are they designed to provide the information needed by management to monitor *services? If they are designed to perform both tasks at once, is this compromising their function for either or both tasks. The tendency of research to date has been to take a dualist perspective and to study whether or not practitioners are using the forms as the instructions specify; little attention is paid to whether it is leading to improvements in practice.*

The accidental user

Hollnagel and Woods (2005: 109) also make the point that discussions of how artefacts are used tend to assume that the user is *motivated* to use it. However, the *accidental* user is becoming more prevalent, particularly with the spread of information technology, where the user *has* to interact with it rather than *wants to.* For example, at some train stations, one no longer has the option of buying a ticket from a person but has to use a ticket machine. They define 'accidental' to cover all instances where there is no alternative: an accidental user is 'a person who is forced to use a specific artefact to achieve an end, but who would prefer to do it in a different way if an alternative existed. From the systemic point of view of the accidental user, the system is therefore a barrier that blocks access to the goal – or which at least makes it more difficult to reach the goal'.

This concept seems to have potentially high relevance in child welfare services where many front line workers are being forced to adopt artefacts that they feel are imposed upon them. The accidental user poses a particular challenge to designers because the current orthodoxy has been that users are motivated and have the required level of knowledge and skills but it needs to be recognised that these assumptions may not apply in all cases. Hollagel and Woods suggest, slightly tongue-in-cheek, that 'the accidental user should be considered as if governed by a version of Murphy's Law, such as "Everything that can be done wrongly, will be done wrongly" (2005: 110).

The worker/IT relationship

Computers have transformed our ability to collect, process, and transmit data. They are having a major impact on the cognitive work of child welfare services. Computer technology has the potential to cause more profound changes than anything since the Industrial Revolution. It creates serious challenges to professionals and to politicians about how to manage the process of transformation of work. Therefore, its potential dangers as well as benefits are important to realise.

Woods et al (Woods et al., 1994: 125-7) lists the latent errors that have been produced by clumsy technology affecting cognitive processes – the key purpose of computer systems.

- increased demand on user memory,
- complication of situation assessment

- undermining of attentional control skills (where to focus when)
- increased workload at high-criticality high-tempo periods
- constraint on the users' ability to develop effective workload management strategies
- impairment of the development of accurate mental models of the function of the device and the underlying processes
- decreased knowledge calibration (i.e. misleading users into thinking that their mental models are more accurate than they actually are
- undermining the cognitive aspects of coordination across multiple agents.

Their strongest advice on avoiding these latent errors is to begin with understanding the human performance that you are planning to enhance with IT. It is only once we understand the factors that contribute to expertise and to failure that we can decide how to design computer systems to enhance expertise. There is no 'neutral' design. It either helps or hinders people's natural ability to express forms of expertise.

They warn of the dangers of overconfidence in the use of computers:

> There is a very ironic state of affairs associated with the clumsy use of technology. The very characteristics of computer-based devices that have been shown to complicate practitioners' cognitive activities and contribute to errors and failures, through studies of the device in context, are generally justified and marketed on the grounds that they reduce human workload and improve human performance. … When our purpose it to help create new cognitive tools, we should start, not with context-free evaluations of the current or proposed prototype computer-based devices,, but by studying and modelling the distributed cognitive system in the context of the demands of the field of activity and the constraints imposed by the organizational context (1994: 128).

> *ICT is becoming a major feature of child welfare systems so the issues around studying human-computer performance are of considerable significance. One message from this literature is that studies should not focus on whether workers are using the programmes as the designers specified but whether the usage is leading to improved services for families.*

6.3 Human-centred design

Zuboff (1988), in a very influential book, introduced the concept of 'informate' as a complement to the concept of 'automate' and argued that using the computer's potential to informate as well as automate could be very productive. An automated system does the work for the operator. An informated system, in contrast, provides the operator with information, preferably in great and varied detail, information that, without the computer, would probably not be so richly available: technology is used to inform not to take over.

Her book, entitled ‘In the age of the smart machine’, poses the question whether we want to be working for a smart machine or whether we want to have smart people working around the machine. She argues that focusing on developing the smart machine concentrates power in the hands of managers and disempowers front line workers. This, her research found, led to workers withdrawing their care – the sense of commitment to and responsibility for tasks (Zuboff, 1988: 272). In contrast, a system that focuses on informating workers reduces the power of managers and has significant impact on the hierarchical structure of the organization (1988: 290).

> Information technology re-organizes our material world, giving us choices in the conception and distribution of knowledge in the workplace. Intelligence can be lodged in the smart machine at the expense of the human capacity for critical judgment. Members then become ever more dependent, docile, and secretly cynical. The alternative is we could create a work force that can exercise critical judgment as it manages the surrounding machine systems (1988: 5).

In many cases, the automated part of the computer will be running the process and the operator just monitoring but, if he or she is given information about how it is operating then they are in a position to step in when needed which is typically when the automated system runs into difficulties.

Norman (1993) takes up this proposal in developing his argument for human-centred design:

> In an automated system, workers are relegated to the role of meter watchers, staring at automatic displays, waiting for an alarm bell that calls them to action. In an informated system, people are always active, analyzing patterns, continually able to find out the state of whatever aspect of the job is relevant at the moment. This involvement empowers the worker not only to become more interested in the job, but also to make more intelligent decisions. When the equipment fails in an automated system, workers are much more knowledgeable about the source of the problems and about possible remedies (1993: 50).

> *Readers who are working in child welfare services may think that their work environment is far removed from an automated system such as the modern plane cockpit. However, the points made about informating and automating should be borne in mind as we increasingly use computers to specify what information a worker should collect, how it should be categorised, and what decisions should be made from it.*

Norman (1993) warns that the impact of a drive to automate can lead to managers emphasizing machine intelligence and managerial control over the knowledge base at the expense of developing knowledge in the operating work force. They use technology to increase their control and sense of certainty but thereby reduce the practitioners' control.

Norman argues that the traditional approach of designers has been machine-centred, concentrating on automating the parts of the process they know how to work on and leaving it to the human operator to cope with the resulting changes. In contrast, a human-centred approach starts by considering their needs. Designers should begin by asking (i) what are humans good at, and (ii) what are machines good at? Before deciding how to help humans improve their performance. He concludes:

> We need information processing systems that complement our thinking, reasoning, and memory skills as comfortably as the calculator enhances our arithmetic skills. The best of these technologies provide us with rich information and leave control of the process, and what to do with the results, in our hands' (1993: 52).

Another dimension of human-centred design both of artefacts and of policies is having an accurate understanding of the work environment as experienced by the operators.

> Because technical work … is so poorly understood, policy makers routinely fall back on stereotypes or images of work …. in order to make sense of technical work … efforts to reduce 'error' misfire when they are predicated on a fundamental misunderstanding of the primary sources of failures in the field of practice and on misconceptions of what practitioners actually do (Barley and Orr, 1997, p.18).

As discussed in Chapter Two the Psychologist's Fallacy assumes that a distant view of the workplace captures the actual experience of those who perform work in context: 'distant views can miss important aspects of the actual work situation and thus can miss critical factors that determine human performance in that field of practice' (Hollnagel and Woods, 2005: 53).

6.4 Conclusion

There is a growing body of literature about the influence of artefacts on human performance, with computers being seen as a major factor in modern work environments. As tools and ICT are increasingly being introduced into child welfare, this is a highly relevant, albeit, highly technical, field of study. Systems theorists stress the importance of conceptualising the relationship between a worker and an artefact as a joint system with the focus of study being on whether, together, they lead to the desired level of performance.

There is also a strong message about the design of tools – the need to consider whether they are user- or tool-centred. This raises questions for child welfare about studying how workers operate so that informed decisions can be made about where they need help and what tools, if any, would provide this help.   There is also a tendency for tools to have a dual purpose:

information-gathering for senior management to monitor performance and information-gathering for front line workers to help them help families. It needs to be studied to what extent these two aims are compatible.

Chapter Seven

7. The human/organizational system relationship

The front line worker operates within a complex socio-technical system. The traditional inquiry was often satisfied by finding human error at the front line. In a systems investigation, the aim is to go beyond the error and ask why it happened. The investigation explores factors in the work environment - the latent errors - that contributed to the causation of that error. The previous two chapters looked at two specific domains of investigation – factors within the individual (human factors) and factors in the human/artefact relationship. This chapter looks at the overall context of the organizational system within which people and artefacts are situated.

The potential range of factors that could be the focus of study is vast and each factor requires specific investigatory methods to research its functioning and the contribution it makes within the organization (e.g. the impact of particular leadership styles or of inexperienced supervisors). Therefore, the aim of this chapter is to identify the organizational factors that are frequently cited as relevant but not to offer detailed guidance on how they can be studied. There are two sections. The first presents a range of models of organizational systems that have been developed to bring structure and order to the potentially vast range of factors that could be considered. The second section explores some of these factors in more depth.

7.1 The structure of the socio-technical system

Several models of the socio-technical system are available in the literature but it quickly becomes clear that they cover much the same ground although varying in the degree of detail.

The leading theorist, Peter Reason, (Reason, 1997: 120) divides the causal factors for accidents into three layers: the unsafe acts carried out by individuals that precede the accident, local workplace factors, and organizational factors. The final category is where he argues you will find the ‘parent’ factors that contribute to unsafe acts lower down in the organization and, unless these are changed, unsafe acts will continue to happen. The first layer has been covered in Chapter Five on human factors. He provides the following amplification of the other two layers.

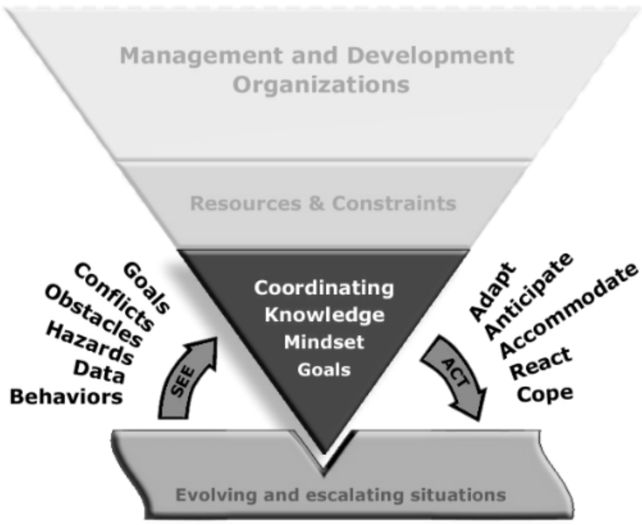
> *Local workplace factors:* One stage upstream from unsafe acts are their immediate mental and physical precursors – such things as poor workplace design, clumsy automation, inadequate tools and equipment, unworkable procedures, the absence of effective supervision, high workload, time pressure, inadequate training and experience, unsociable hours, unsuitable shift patterns, poor job planning, undermanning, badly calibrated hazard perception, inadequate personal protective equipment, poor teamwork, leadership

> shortcomings and the like. They are likely to be fewer in number than the unsafe acts they breed. As such, they are more easily managed than the human condition. But they are still only the local expression of higher-level organizational problems.
> *Organizational factors:* Only in the upper levels of the system can we begin to get to grips with the 'parent' failure types – the processes that create the downstream 'problem children'. If these remain unchanged, then efforts to improve things at the workplace and worker level will be largely in vain. The damaging effects of certain kinds of unsafea ct may be reduced and specific conditions within which the workplace improved, but the continued existence of the 'parent' failures in the upper echelons of the organization will ensure their rapid replacement by other kinds of human and work place problems. Clearly, then, organizational factors represent the priority are for process measurement.

He then offers a summary of five broad clusters of organizational factors that are important:

> - *safety-specific factors* (for example, incident and accident reporting, safety policy, emergency resources and procedures, off-the-job safety and so on)
> - *management factors* (for example, management of change, leadership and administration, communication, hiring and placement, purchasing controls, incompatibilities between production and protection and so on)
> - *technical factors* (for example, maintenance management, levels of automation, human-system interfaces, engineering controls, design, hardware and so on)
> - *procedural factors (*for example, standards, rules, administrative controls, operating procedures and so on)
> - *training* (for example, formal versus informal methods, presence of a training department, skills and competencies required to perform tasks and so on).
>
> At the core of these clusters and pervading all of them is the issue of culture … we can link cultural factors to the three Cs: commitment, competence and cognisance- but as they exist within the organization as a whole, rather than in the mind of any one senior manager.

Woods and Hollnagel (2006: 7) offer a model that identifies the key components of the socio-technical system but also details the elements of the front line workers' tasks.



Helmreich (2000) offers a structure that has been adapted to the context of health care and so has a more direct relevance to child welfare services.

Latent threats

National culture, organizational culture, professional culture, scheduling, vague policies
Threat management strategies and countermeasures
Environmental factors
Organizational factors
Individual(physician) factors

Immediate threats

Team/crew factors
Patient factors
Error detection and response induced
Patient state
Management of patient state
Adverse outcome
Further error

Behaviours that increase risk to patients in operating theatres

Communication:

- Failure to inform team of patient's problem—for example, surgeon fails to inform anaesthetist of use of drug before blood pressure is seriously affected
- Failure to discuss alternative procedures

Leadership:

- Failure to establish leadership for operating room team

Interpersonal relations, conflict:

- Overt hostility and frustration—for example, patient deteriorates while surgeon and anaesthetist are in conflict over whether to terminate surgery after pneumothorax

Preparation, planning, vigilance:

- Failure to plan for contingencies in treatment plan
- Failure to monitor situation and other team's activities—for example, distracted anaesthetist fails to note drop in blood pressure after monitor's power fails

Education and debate

7.2 Organizational factors

Moving from a consideration of frameworks for studying organizations, this section considers some of the key issues outlined in the literature.

A safety culture

Safety in organizations is ultimately linked to the priorities and values of the wider organization. Helmreich and Merritt (1998) define safety as ‘a social construct and ultimately a state of mind. It rests partly in formal structures and protocols but fundamentally in attitudes. Safety intertwines with risk and human life’ (1998: xx). They studied the influence of culture in aviation and in health care within operating theatres. Like child welfare services, both domains require high levels of interpersonal collaboration, communication and co-ordination. Teamwork is essential.

The Human Factors Study Group of the Health and Safety Executive (1993) offer the following definition of a safety culture:

> The safety culture of an organization is the product of the individual and group values, attitudes, competencies, and patterns of behaviour that determine the commitment to, and the style and proficiency of, an organization’s health and safety programmes. Organizations with a positive safety culture are characterised by communications founded on mutual trust, by shared perceptions of the importance of safety, and by confidence in the efficacy of preventative measures .

Conflicting demands and trade-offs

In reality, many operate within organizations where there is no clear, overriding priority given to safety but, instead, operators frequently have to make trade-offs between competing priorities, the most common being between efficiency and safety. A pilot who grounds a plane because of a small technical problem costs the company a significant loss of revenue. Flying despite the technical problem raises the risk of an accident but protects revenue. Pilots therefore are often faced with making the trade-off between safety and cost and, sometimes, can be seen with hindsight, to make the wrong choice.

Resources

Key elements of resource issues are how well trained and experienced are the personnel and how many of them are there? An organization operating with a high number of inexperienced, poorly trained operators at the front line with heavy workloads is setting up the scenario in which errors are more likely to occur. Inadequate training and experience increase the risk of errors in evaluating and interpreting events. Heavy caseloads lead to insufficient time to keep fully alert to what is happening and so not notice if something is out of the normal (Hollnagel and Woods, 2005: 74).

Regulatory climate

The degree of flexibility or regulation of front line work influences the ability of the worker to respond to emerging problems. It raises the question of how much work should be proceduralised. Procedures clearly have many benefits in encapsulating the wisdom acquired from experience and making it readily available to a new generation of workers. Indeed, the common conclusion of traditional inquiries has been to draw up a list of recommendations outlining new rules and procedures that should be introduced. Within such a culture, flexibility is seen as a source of danger and to be eradicated from the system as far as possible: it is 'a regulatory climate which believes that absolute adherence to procedures is the means to achieve safe operations and avoid 'human errors' (Woods et al., 1994: 88).

However, for those who argue that people should be seen as the creators of safety not just a source of error, the scope of procedures needs to be kept in check. Excessive regulation discourages successful problem solving in situations which require more than the rote following of instructions and hence restricts the ability of humans to create safety in unexpected circumstances. There comes a point where the degree of regulation becomes a latent error in the system: 'The important issue in many hazardous technologies is not *whether* to violate, but *when* to violate – or perhaps, more importantly, when to comply' (Reason, 1997: 49).

The uneven division of responsibility and authority may lead to front line workers finding themselves in the uncomfortable position where their responsibility has been removed, and they are required to follow procedures, yet they are still held responsible for bad outcomes (Woods et al., 1994: 87).

Causing the next accident by trying to prevent the last one

An issue related to regulation is the common scenario where the recommendations arising from one inquiry help prevent a precise repeat of the accident but create latent errors that contribute to new pathways to accidents.

Reason (1997: 52) cites the example of the investigation into the nuclear accident and Three Mile Island. A turbine in one of the pressurized water reactors stopped (tripped) automatically because of a leak of water through a seal. The inquiry found that one significant error of the operators was to cut back on the high-pressure injection of water into the reactor coolant system. This reduction caused serious damage to the reactor's core. As a result of this inquiry, the US nuclear regulators made it a requirement that operators should not cut back on the high-pressure injection during a recovery stage of an off-normal event. Three years later, a control room crew in another power plant were faced with an emergency in which reducing the high-pressure injection would have been an appropriate step to take. Because of the new regulation, they did not take this step and, consequently, the emergency was extended for several hours more than it need have been. As Reason points out, regulators cannot foresee all the possible scenarios of failure on complex systems.

Feedback to inform learning

A major factor that affects safety is the degree to which the organization has good feedback mechanisms in place so that it is able to identify and resolve latent errors. This issue is dealt with in more detail in the following chapter in the section on learning organizations.

7.3 High reliability organizations

High reliability organizations (HROs) can be defined as organizations which have fewer than normal accidents. The existence of such organizations has prompted considerable research to understand the secrets of their success. A team from the University of California, Berkeley (Todd La Porte, Karlene Roberts, and Gene Rochlin) have been studying groups that operate with highly complex and hazardous technological systems and achieve impressive rates of safety. The US air traffic control system was one focus of studies; it handles tens of thousands of flights a day yet, for more than a decade, none of the aircraft being monitored have collided with another. Another focus were aircraft carriers in the US navy where planes have to land on a small runway that is pitching from side to side so they are operating with very small margins of error.

This decrease in accidents occurs crucially through a change in culture. Technology has some influence but not in isolation, nor does it work without a change in the organization's culture. Research has identified some key characteristics of HROs. These include:

- organizational factors (i.e., rewards and systems that recognize costs of failures and benefits of reliability),
- managerial factors (i.e., communicate the big picture),
- and adaptive factors (i.e., become a learning organization)

More specifically, HROs actively seek:

- to know what they don't know,
- design systems to make available all knowledge that relates to a problem to everyone in the organization,
- learn in a quick and efficient manner,
- aggressively avoid organizational hubris,
- train organizational staff to recognize and respond to system abnormalities,
- empower staff to act, and
- design redundant systems [spare capacity] to catch problems early (Roberts and Bea, 2001).

In other words, an HRO expects its organization and its sub-systems will fail and works very hard to avoid failure while preparing for the inevitable so that they can minimize the impact of failure.

Research on organizations that are not HROs offers a contrasting picture. Economic pressures mean that the conflict between efficiency and safety too often prioritises efficiency so that safety is not the primary objective. Reason (1990) also shows how organizations often learn from experience but learn the wrong lessons: they learn to cut corners, that disasters are rare, and that they are not likely to be vulnerable as a consequence of their own risky enterprise.

7.4 Conclusion

Once the focus of study moves from the individual to the organization, a wide range of factors are potentially relevant. A number of models of organizational variables are available in the literature which could provide the basis for formulating models specific to the child welfare services in a particular country.

The range of potential factors makes it difficult for any one inquiry to study all in depth. This raises the question of how priorities are decided and whether the need for selection creates an opportunity for biases to creep into the inquiry. Woodcock and Smiley’s (1998) study found that the more senior the position of the safety specialist, the more likely the specialist was to provide human error-type attribution as opposed to identifying factors in the situation that contributed.

Chapter Eight

8. Methodologies of data collection

Having considered the range of factors that can be implicated in creating both safety and danger in organizational systems, the next step is to consider how to find the evidence to analyse in studying how events unfold. This chapter is divided into two main sections: methods of continuous learning within organizations and methods for investigating specific incidents.

8.1 Continuous learning processes

Mechanisms for continuous learning are in two main forms:

(i) the learning organization that incorporates systems for feedback so that there is a continual possibility for learning of emerging problems;

(ii) reporting systems for 'critical incidents' short of disaster so that there is scope for learning where there are weaknesses in the system that could contribute to a more serious outcome on some future occasion. Reporting systems were first established in aviation but are now being implemented by many health systems around the world.

(i) The learning organization

The interest in creating a learning organization comes partly from the safety literature and the realisation of the importance of having good feedback loops in order to detect latent failures. Fundamental to the process of preventing major accidents is to be able to detect minor problems that might, if coinciding with other latent errors, lead to disaster. This means that the organization needs to have in place mechanisms for learning of minor glitches and removing them.

Another source of interest, however, comes from concerns with the rate of change in modern society so that institutions need to learn continually to keep up with developments. Donald Schon (1973) wrote eloquently on this issue, arguing that we need to be adaptable 'in response to changing situations and requirements; we must invent and develop institutions which are 'learning systems', that is to say, systems capable of bringing about their own continuing transformation. (Schon, 1973: 28)

In the safety literature the merits of a learning organization are centred on the reduction in error but, for others, the concept offers gains for the individual as well as the organization. Senge (1990), for example, pays attention to the impact on the individual:

> Learning organizations [are] organizations where people continually expand their capacity to create the results they truly desire, where new

and expansive patterns of thinking are nurtured, where collective aspiration is set free, and where people are continually learning to see the whole together. (Senge, 1990: 3)

The Department of Health publication ‘An Organization with a Memory’ (2000) provides a useful summary of the relevant literature, including an overview of the barriers to learning in organizations:

- An undue focus on the immediate event rather than on the root causes of problems;
- Latching onto one superficial cause or learning point to the exclusion of more fundamental but sometimes less obvious lessons;
- Rigidity of core beliefs, values and assumptions, which may develop over time - learning is resisted if it contradicts these;
- Lack of corporate responsibility - it may be difficult, for example, to put into practice solutions which are sufficiently far-reaching;
- Ineffective communication and other information difficulties - including failure to disseminate information which is already available;
- An incremental approach to issues of risk - attempting to resolve problems through tinkering rather than tackling more fundamental change;
- Pride in individual and organizational expertise can lead to denial and to a disregard of external sources of warning - particularly if a bearer of bad news lacks legitimacy in the eyes of the individuals, teams or organizations in question;
- A tendency towards scapegoating and finding individuals to blame - rather than acknowledging and addressing deep-rooted organizational problems;
- The difficulties faced by people in 'making sense' of complex events is compounded by changes among key personnel within organizations and teams;
- Human alliances lead people to 'forgive' other team members their mistakes and act defensively against ideas from outside the team;
- People are often unwilling to learn from negative events, even when it would be to their advantage;
- Contradictory imperatives - for example communication versus confidentiality;
- High stress and low job-satisfaction can have adverse effects on quality and can also engender a resistance to change;
- Inability to recognise the financial costs of failure, thus losing a powerful incentive for organizations to change (Department of Health, 2000: 34).

]

Experience and research studies suggest that safety is likely to be a strong feature of an informed culture, which has four critical sub-components:

- **a reporting culture:** creating an organizational climate in which people are prepared to report their errors or near-misses. As part of this process data need to be properly analysed and fed back to staff making reports to show what action is being taken;
- **a just culture:** not a total absence of blame, but an atmosphere of trust in which people are encouraged to provide safety-related information - at the same time being clear about where the line is drawn between acceptable and unacceptable behaviours. An example is the airline safety system which we discuss later in this chapter;
- **a flexible culture:** which respects the skills and abilities of 'front line' staff and which allows control to pass to task experts on the spot;
- **and a learning culture:** the willingness and competence to draw the appropriate conclusions from its safety information system, and the will to implement major reforms where their need is indicated.

A combination of research and experience also suggests a number of ways in which some of the barriers to active learning can be overcome or minimised, helping to create informed cultures which can learn from and respond to failures.

What can we do to create an informed culture?

- Raise awareness of the costs of not taking risk seriously. There is a need for more routinely available data on the human and financial costs of adverse events;
- Focus on 'near misses' as well as actual incidents. This can remove the emotion from an incident and allow learning to take place more effectively. It is also easier to keep near miss data anonymous, itself a factor in encouraging reporting;
- Ensure that concerns can be reported without fear. Bearers of bad news may fear that they will be ostracised or silenced: clear rules about what must be reported, and regarding reporting as good behaviour rather than as disloyalty will all help;
- Avoid simplistic counting. Data must be analysed and synthesised to reveal their underlying lessons;
- Develop effectively-led teams as mechanisms for culture change.
- Teams need to be firmly linked into the wider management structure to ensure that alliances within them do not hamper learning. Teambased training can also be a useful tool here.
- Use external input to stimulate learning. External input can help teams to think outside established parameters and challenge assumptions about the way things are done. User involvement can be of particular value in encouraging learning;
- Ensure effective communication and feedback to front-line staff.
- Teams and organizations must operate on genuinely two-way communication, not just 'top down'. Communication systems need to be in place to allow people to see what has changed as a result of incident or near miss reporting;
- Give a high-profile lead on the issue. Make it clear both nationally and locally that safety and quality are key goals;

- Recognise staff concerns. Try hard to emphasise the personal and service benefits of change rather than just the threats.

Source: Derived from Firth-Cozens (2000).

SCIE (2004) provides a useful summary of the key characteristics of a learning organization adapted to the social care field. It lists five principal features of a learning organization (derived from Iles and Sutherland, 2001):

Organizational structure: managerial hierarchies that enhance opportunities for employee, carer and service user involvement in the organization. All are empowered to make relevant decisions. Structures support teamwork and strong lateral relations (not just vertical). Networking is enabled across organizational and hierarchical boundaries both internally and externally.

Organizational culture: strong cultures that promote openness, creativity, and experimentation among members. They encourage members to acquire, process and share information, nurture innovation, and provide the freedom to try new things, to risk failure and to learn from mistakes.

Information systems: learning organizations require information systems that improve and support practice and the move beyond those used in traditional organizations where information is generally used for control purposes. 'Transformational change' requires more sophisticated information systems that facilitate rapid acquisition, processing and sharing of risk, complex information which enables effective knowledge management.

Human resource practices: people are recognised as the creators and users of organizational learning. Accordingly, human resource management focuses on provision and support of individual learning. Appraisal and reward systems are concerned to measure long-term performance and to promote the acquisition and sharing of new skills and knowledge.

Leadership: like most interventions aimed at securing significant organizational improvement, organizational learning depends heavily on effective leadership. Leaders model the openness, risk-taking and reflection necessary for learning, and communicate a compelling vision of the learning organization, providing the empathy, support and personal advocacy needed to lead others towards it. They ensure that organizations and work groups have the capacity to learn, change and develop.

Reporting systems

The basic assumption behind reporting systems is that most problems are not just a series of random, unconnected one-off events but often have common underlying causes. If incidents are reported and analysed then it becomes possible to identify similarities and patterns in sources of error so that others can be warned of the vulnerability in their system. For example, if several doctors report that they muddled up the same two drugs, then it looks likely that similarities in packaging are a latent error in the system, making this type of mistake more likely. Changing the packaging to ensure they are very distinctive will reduce the risk of similar errors in the future. However, it would be very cumbersome for an individual hospital to try and sort this out alone – it would need to unpack and repack the incoming supplies of the drugs. The more efficient solution requires action by the manufacturer, higher up the chain of events.

Aviation in the US was the first field to develop a reporting system, influenced by the highly visible costs of plane crashes, both to human life and to profits. Its success has led to it becoming an established worldwide system creating the opportunity for all to learn of emerging problems and modify design or practices to avoid major problems.

The following summary of the system is taken from Department of Health (2000: 45). The Aviation Safety System operates internationally, though reporting of lower-level incidents in particular is better-developed in some countries than in others. The system has five principal components, which combine to provide a means of detecting, analysing and acting on actual incidents and "near misses" or other errors, along with proactive identification of issues which have the potential to pose a safety risk if left unchecked.

Components of the aviation safety system:

- Accident and serious incident investigations, governed by the International Convention on International Civil Aviation (ICAO) Accident/Incident Data Reporting Programme (ADREP). ADREP includes provision for the international dissemination of investigation reports.
- The Mandatory Occurrence Reporting Scheme (MORS), which provides a mechanism for notifying and reporting a range of adverse occurrences regardless of whether they result in an accident. MORS feeds into a database at national level for trend analysis and feedback to the industry.
- The Confidential Human Factors Incident Reporting Programme (CHIRP), which is administered by an independent body and which provides sensitive follow-up and feedback on reports of human errors that have been rendered anonymous.
- Company safety information systems, such as British Airways' BASIS system, which record all levels of safety-related incidents. Information is shared on a peer basis within systems, and staff report with an explicit reassurance that no individual will be pursued for an honest mistake.

- Operational monitoring systems, which proactively monitor crew competency through regular checks and review Flight Data Recorder information from every flight. There is management/union agreement on handling of any incidents or failures detected in this way.

The focus of the system is on detecting and learning from not only accidents and serious incidents, but also lower-level incidents or near misses, some of which might have the potential to lead to a more serious occurrence. The aviation safety system receives reports of around 600 incidents, 30 serious incidents and 10 accidents for every one fatal accident. Thus in aviation the great majority of learning is extracted not from accidents themselves but from incidents which had the potential to result in accidents.

From research on the characteristics of effective safety information systems, together with experience from the aviation industry, we can draw a number of conclusions about the characteristics of effective incident reporting systems.

Characteristics of effective incident reporting systems

- separation of collection and analysis from disciplinary or regulatory bodies
- collection of information on "near misses" as well as actual incidents
- rapid, useful, accessible and intelligible feedback to the reporting community
- ease of making a report
- standardised reporting systems within organizations
- a working assumption that individuals should be thanked for reporting incidents, rather than automatically blamed for what has gone wrong
- mandatory reporting
- standardised risk assessment - i.e. a common understanding of what factors are important in determining risk
- the potential for confidential or de-identified reporting

In setting up a reporting system in child welfare services, a key issue to be addressed is the definition of 'critical incident' or 'near miss'. This issue is discussed in more detail in the next chapter looking at developments in health care where the problems are somewhat closer to child welfare than those of aviation or the nuclear power industry.

8.2 Incident focused studies

Besides mechanisms for continual learning and improvement, there is also a need for investigations of specific incidents with serious outcomes. A major disaster is clearly an opportunity for learning and an occasion when it would be reckless not to try to learn some lessons of how the accident occurred. In addition, there are strong social factors that drive the need for a thorough examination of the sequence of events that led up to the accident.

The public inquiry has become a permanent feature of modern government as an instrument of dampening or dissembling public disquiet about a scandal or a disaster … The only viable method of allaying public fears of a cover-up by the organizations involved is to provide for a public inquiry, independent of those with a vested interest in the outcome (Blom-Cooper, 1996: 57).

When a tragic accident occurs, there is also a perhaps understandable human urge to find someone to blame. The traditional inquiry has met this urge well, often naming (usually junior) staff who made the final, fatal mistake. The systems approach is less concerned with blame but there is, in practice, often still a need to decide what degree of responsibility individuals bear for their part in the causal sequence leading to disaster.

A wide range of investigation techniques have been developed (19 are identified in the Health Techonology Assessment review of the literature (Woloshynowych et al., 2005) but their usefulness to child welfare is limited by (a) a lack of empirical evidence confirming their value in practice, and (b) many are created to deal with specific technological issues that are not pertinent to child welfare.

All the methods of accident analysis have features in common. All try to answer the three basic questions: What happened? How did it happen? Why did it happen? All rely on two prime sources of evidence; documents and interviews, although some also collect data through observation and simulation techniques. All are offering some means of putting a structure onto the wealth of data and helping to ensure that all possibilities are explored.

However, health care is a much more closely related discipline and so it seems more appropriate to discuss methods of investigating specific incidents in the next chapter, looking specifically at the way methods from other industries have been adapted for use in health care.

Chapter Nine

9. Developments in health care

The systems approach has been enthusiastically welcomed in health care and developments there are a useful model for child welfare services. This chapter covers three areas of development: the acceptance of the systems approach, the development of reporting systems, and the development of incident investigation methods.

9.1 The acceptance of the systems approach

The systems approach has received endorsements from international and national organizations in health care and led to the setting up of special organizations to help implement strategies for improving patient safety.

The World Health Organization has taken up the need to study latent as well as active errors in healthcare organizations. The international status of this organization is significant because it sees a reporting system as having the scope to produce more than local learning. Its vision includes the hope that 'it may be possible for the bad experience suffered by a patient in one part of the world to be a source of transmitted learning that benefits future patients in many countries' (WHO, 2005a) To advance the project, it has established the World Alliance for Patient Safety:.

> Thinking in terms of "systems" offers the greatest promise of definitive risk-reduction solutions, which place the appropriate emphasis on every component of patient safety, as opposed to solutions driven by narrower and more specific aspects of the problem, which tend to underestimate the importance of other perspectives.
>
> Enhancing the safety of patients includes three complementary actions: preventing adverse events; making them visible; and mitigating their effects when they occur. This requires: (a) increased ability to learn from mistakes, through better reporting systems, skilful investigation of incidents and responsible sharing of data; (b) greater capacity to anticipate mistakes and probe systemic weaknesses that might lead to an adverse event; (c) identifying existing knowledge resources, within and outside the health sector; and (d) improvements in the health-care delivery system itself, so that structures are reconfigured, incentives are realigned, and quality is placed at the core of the system. In general, national programmes are built around these principles (WHO, 2005a: 4).

The Institute of Medicine in the US was one of the first national health organizations to endorse a systems approach wholeheartedly. Its 1999 publication captures the rejection of the traditional person-centred approach in its title: To Err is Human: Building a Safer Health System. One of the report's main conclusions is a clear statement of the systems' view of error:

> The majority of medical errors do not result from individual recklessness or the actions of a particular group – this is not a "bad apple" problem. More commonly, errors are caused by faulty systems, processes, and conditions that lead people to make mistakes or fail to prevent them. … Thus mistakes can best be prevented by designing the health system at all levels to make it safer – to make it harder to do something wrong and easier for them to do it right (Institute of Medicine, 1999: 2).

To improve safety, the report recommends a four-tiered approach:

- establishing a national focus to create leadership, research, tools, and protocols to enhance the knowledge base about safety;
- identifying and learning from errors by developing a nationwide public mandatory reporting system and be encouraging health care organizations and practitioners to develop and participate in voluntary reporting systems;
- raising performance standards and expectations for improvements in safety through the actions of oversight organizations, professional groups, and group purchasers of health care;
- implementing safety systems in health care organizations to ensure safe practices at the delivery level.

In the UK, similar developments have been taking place. In June 2000, the Government accepted all recommendations made in the report of an expert group, led by Dr Liam Donaldson, Chief Medical Officer, called An Organization with a Memory. The report acknowledged that there has been little systematic learning from patient safety incidents and service failure in the NHS in the past and drew attention to the scale of the problem of potentially avoidable events that result in unintended harm to patients.

An Organization with a Memory proposed solutions based on developing a culture of openness, reporting and safety consciousness within NHS organizations. It proposed the introduction of a new national system for identifying patient safety incidents in healthcare to gather information on causes and to learn and act to reduce risk and prevent similar events occurring in future.

In 2001, the National Patient Safety Agency was established as a Special Health Authority. It describes its mission as:

> To co-ordinate the efforts of the entire country to report, and more importantly to learn from mistakes and problems that affect patient safety.
>
> As well as making sure errors are reported in the first place, the NPSA is trying to promote an open and fair culture in the NHS, encouraging

> all healthcare staff to report incidents without undue fear of personal reprimand. It will then collect reports from throughout the country and initiate preventative measures, so that the whole country can learn from each case, and patient safety throughout the NHS can be improved.

9.2 The development of reporting systems

The WHO (2005b) provides draft guidelines for adverse event reporting and learning systems and recommends the establishment of a worldwide system so that lessons can be learned and shared across countries. The following section draws on their summary of the lessons learned about setting up reporting systems because (a) it draws together lessons from many countries, and (b) it is the most recent major review of the topic.

In transferring a model from industries with a significantly different purpose, it was necessary for the WHO to redefine some of the key terms to fit the healthcare field (2005b: 8). 'Error' retains a meaning used in other industries:

> The failure of a planned action to be completed as intended (i.e. error of execution) or the use of a wrong plan to achieve an aim (i.e. error of planning). Errors may be errors of commission or omission, and usually reflect deficiencies in the systems of care'.

It defines an adverse event as:

> An injury related to medical management, in contrast to complications of disease. Medical management includes all aspects of care, including diagnosis and treatment, failure to diagnose or treat, and the systems and equipment used to deliver care'.

'Near-miss' or 'close call' is defined as:

> Serious error or mishap that has the potential to cause an adverse event but fails to do so because of chance or because it is intercepted. Also called potential adverse event.

The report summarises the key messages learned from existing reporting systems:

> Current reporting systems span a spectrum of objectives incorporating both learning and accountability considerations.

> The primary objectives of a reporting system will determine the design, for example, whether reporting is voluntary and confidential.

Reporting systems need to be clear on who reports, the scope of what is reported and how reports are made.

Reporting of incidents is of little value unless the data collected are analysed and recommendations are disseminated.

Experts who understand statistical methods, the practice concerns, clinical significance, systems issues, and potential preventive measures are essential to analyse reported incidents.

Classification and simple analytic schemes start the process of categorizing the data and developing solutions that can be generalized.

The latter message highlights the need for developing taxonomies for classifying patient safety incidents. Vincent's identification of weaknesses in the UK reporting system (discussed in Chapter Four) stemmed from an inadequately formulated taxonomy. This problem is important to remember when trying to draw lessons from reporting systems in health for use in child welfare services. Developing a theoretically-grounded taxonomy will be a major project.

The characteristics of successful reporting systems seem worth reporting in full since they contain valuable lessons (WHO, 2005b: 50-51):

Non-punitive. The most important characteristic for success of a patient safety reporting system is that it must be non-punitive. Neither reporters nor others involved in the incidents can be punished as a result of reporting. For public systems, this requirement is the most difficult to achieve, since the public often assumes an individual is to blame, and there can be strong pressure to punish the "culprit". While perhaps temporarily emotionally satisfying, this approach is doomed to fail. People will not report any errors they can hide. It is important for national systems to protect reporters from blame. The best way to do this is by keeping the reports confidential.

Confidential. The identities of the patient and reporter must never be revealed to any third party. At the institutional level, confidentiality also refers to not making public specific information that can be used in litigation. Although, historically, breach of confidentiality has not been a problem in public or private systems, concern about disclosure is a major factor inhibiting reporting for many voluntary reporting programmes

Independent. The reporting system must be independent of any authority with the power to punish the reporter or organization with a stake in the outcome. Maintaining a "firewall" between the reporting agency and the disciplinary agency in a governmental system can be difficult, but it is essential if trust in reporting is to be maintained.

Expert analysis. Reports must be evaluated by experts who understand the clinical circumstances under which the incidents occur and who are trained to recognize underlying systems causes. While it seems obvious that collecting data and not analysing it is of little value, the most common failure of governmentally run reporting systems is to require reporting but not to provide the resources needed to analyse the reports. Huge numbers of reports are collected only to sit in boxes or on computers. Expertise is a major, and essential, resource requirement for any reporting system.

Credible. The combination of independence and the use of content experts for analysis is necessary if recommendations are to be accepted and acted upon.

Timely. Reports must be analysed without delay, and recommendations must be promptly disseminated to those who need to know. When serious hazards are identified, notification should take place rapidly. For example, the Institute for Safe Medication Practice issues prompt alerts through its regular publication when new hazards in drugs are discovered.

Systems-oriented. Recommendations should focus on changes in systems, processes or products, rather than being targeted at individual performance. This is a cardinal principle of safety that must be reinforced by the nature of recommendations that come from any reporting system. It is based on the concept that even an apparently egregious individual error results from systems defects, and will recur with another person at another time if those systems defects are not remedied.

Responsive. For recommendations to result in widespread systems changes, the organization receiving reports must be capable of making and disseminating effective recommendations, and target organizations must make a commitment to implement recommendations. A good example is the National Reporting and Learning System in England and Wales which allows the National Patient Safety Agency to develop new solutions that are disseminated throughout the system.

9.3 The development of incident investigation methods.

Woloshynowych et al's (2005) literature review provides a good overview and identifies the approaches most suited to health settings. Although health differs from children's services in significant ways, the similarities between the two disciplines are very strong when contrasted with the differences between children's services and aviation or nuclear power plants. Therefore this review provides a useful filter and I have followed up the techniques it identified as the most relevant: Management Oversight and Risk Tree (MORT), Australian Incident Monitoring System (AIMS), Critical Incident Technique (CIT), Significant Event Auditing (SEA), Root Cause Analysis (RCA), Organizational Accident Causation Model (OACM), and Comparison with Standards (CWS).

(1) Management Oversight and Risk Tree

This was developed in the early 1970s for the U.S. Energy Research and Development Administration as a safety analysis method that would be compatible with complex, goal-oriented management systems. MORT is a diagram which arranges safety program elements in an orderly and logical manner. Its analysis is carried out by means of fault tree, where the top event is "Damage, destruction, other costs, lost production or reduced credibility of the enterprise in the eyes of society". The tree gives an overview of the causes of the top event from management oversights and omissions or from assumed risks or both.

The MORT tree has more than 1500 possible basic events inputted to 100 generic events which have being increasingly identified in the fields of accident prevention, administration and management. MORT is used in the analysis or investigation of accidents and events, and evaluation of safety programs. Its usefulness has been reported in the literature; normal investigations revealed an average of 18 problems (and recommendations). Further investigation with MORT analysis revealed an additional 20 contributions per case.

The user's manual can be downloaded from: http://www.nri.eu.com/toppage3.htm. Its usefulness is limited by the sheer size and cost of the task. It requires trained investigators and working through up to 1500 basic events takes considerable time.

(2) Australian Incident Monitoring System

AIMS is intended to provide a national mechanism for reporting problems in health management. It uses a classification system of software that elicits key features of the incident, places the event in context, and records the contributing factors, both system-based and human errors. Contributing factors are: management decisions; infrastructure, working conditions; communications, records; staff quantity and quality; supervision and tasking; equipment availability and/or suitability; policies, protocols and pathways. It has the weakness of being dependent on the detail provided by the person reporting the incident and has no opportunity for checking its accuracy.

(3) Critical Incident Technique

It has been described as a set of principles for gathering data rather than a rigid set of rules. They specify: the aims of the work to be studied, the incidents to be collected, methods of data collection, analysis, and interpretation. It has been applied mainly with the aim of describing and specifying the key skills involved in a particular kind of work. Its negative features are that most studies give little or no information on the methods of investigation or analysis, and it is highly reliant on the intuition and expertise of the investigators.

(4) Significant Event Auditing

SEA involves audit of a single case or event where things went badly or sometimes where things went well. It is designed as a quality improvement technique not to address safety issues. SEA meetings are held with groups of people as a work-based reflective activity. It is potentially anti-hierarchical and the effective functioning of the participants is generally accepted as a pre-requisite for successful event auditing. One possible agenda to follow is: present the case, review of acute/immediate problems, review of possibilities of prevention, plan of action and follow-up, implications for family/community, interface issues, team issues, summary and recommendations. It is widely used as an educational approach in general practice settings in the UK. Its educational value is not disputed but its capacity to promote improvement in practice has not yet been demonstrated.

(5) Root Cause Analysis

RCA was developed in industry as a methodology to investigate serious accidents but has become less used as industries have developed their own methodologies using some techniques of RCA and techniques specific to their industry. It has been widely taken up in health care. Both UK and US health services have been recommending it. However, many adapt and shorten the process so it can now refer to a number of slightly different methods. Seven steps that are usually included are:

- identify the incident to be analysed;
- organise a team to carry out the RCA;
- study the work processes;
- collect the facts;
- search for causes;
- take action;
- evaluate the actions taken.

The Joint Commission on Accreditation of Healthcare Organisatons (JCAHO) in the US provides a manual and examples of 14 RCA tools with a healthcare context.

RCA is not based on a specific theory of human error or system failure but does provide a toolbox of useful techniques and tools for use by incident investigators. Its strength is the focus on improving systems rather than blaming individuals and it provides a complete accident methodology. However, it requires training, is time-consuming, and can easily be made overly complicated and does not guarantee a complete answer.

> *One of the few examples of using a systems approach in child welfare is Rzepnicki and Johnson's (2005) use of root cause analysis to study decision making in child protection inquiries where children have died. In their work, they found that it could be adapted to the child protection*

field and offered valuable opportunities for learning. The 'errors' that they studied were the deaths of children. For each event in the chain leading to this outcome, the investigator asks 'what led to this event? What allowed it to happen?' Repeated questioning leads to the creation of complex event trees with many branches and multiple root causes. They identify three stop rules for ending the causal chain. First, a root cause is found. This is defined as 'the first point in a chain of events that can be eliminated by applying policy, practice, or procedure at the policy/management, supervisory, or individual level' (2005: 11). Second, 'a causal chain is terminated when it is determined that the situation is non-correctable e.g. the fact that 'the child was crying' is not likely to lead to changes over which the child welfare organization has any control' Third, 'a causal chain is terminated when there is insufficient data to continue with the analysis'. . In particular, they highlighted how changes introduced as a response to an earlier tragedy could, inadvertently, have negative consequences: 'the simple addition of new rules and procedures, without removing old ones or fully considering their impact, is not likely to lead to improved performance. Instead, good practice may be compromised and a culture created that encourages rule violations. RCA may be an effective tool to uncover negative consequences of well-intentioned policies.' However, one complicating factor in transferring a method developed in industry to child welfare is the complexity of the number of systems involved: child welfare investigations 'deal with multiple systems (e.g. police, schools, extended families, neighbors, courts, hospitals, etc.) which may make boundaries of the investigation ambiguous'.

(6) Organizational Accident Causation Model

This is based on Reason's model of organizational error discussed in the previous chapter. It has been adapted for a healthcare environment by Vincent and colleagues to develop a protocol for the investigation and analysis of serious incidents in healthcare. The protocol gives a detailed account of the investigation and analysis process, including who should conduct the investigation, the interviewing procedure, analysis, identifying action points and preparing the report. The method incorporates analyses from both interviews and records and assumes that much important material can only be gained from interviews.

The essential process is mirrored in the structure of interviews:

(1) to establish a chronology, including the role of the member of staff being interviewed in the incident, and their account or experience of the events;

(2) to identify the 'care management problem' (CMP) or actions or omissions made by staff or other limitations in the system that had an important role in the critical incident;

(3) to apply a framework of contributory factors to each CMP identified and discuss with the interviewee to identify the most important factors and potential methods of prevention.

The list of contributory factors offers a useful starting point for transferring the method to the field of child welfare.

FACTOR TYPES	CONTRIBUTORY INFLUENCING FACTOR
Patient Factors	Condition (complexity & seriousness)
	Language and communication
	Personality and social factors
Task and Technology Factors	
	Task design and clarity of structure
	Availability and use of protocols
	Availability and accuracy of test results
	Decision-making aids
Individual (staff) Factors	
	Knowledge and skills
	Competence
	Physical and mental health
Team Factors	
	Verbal communication
	Written communication
	Supervision and seeking help
	Team structure (congruence, consistency, leadership, etc)
Work Environmental Factors	
	Staffing levels and skills mix
	Workload and shift patterns
	Design, availability and maintenance of equipment
	Administrative and managerial support
	Environment
	Physical
Organizational & Management Factors	
	Financial resources & constraints
	Organizational structure
	Policy, standards and goals
	Safety culture and priorities
Institutional Context Factors	
	Economic and regulatory context
	National health service executive
	Links with external organizations

The strengths of this approach are that it focuses on improving systems rather than blaming individuals; it identifies a range of weakness in systems; and it is based on current accepted models of human performance. Its limitations are that some investigators have had difficulty with some terminology; models and theories have not been formally evaluated; and few papers address specific interventions.

> *In considering how to transfer incident focused investigation techniques to a child welfare setting, one question that arises is 'what counts as an incident in child welfare'? It may be possible to identify some features of practice that readily fit the term e.g. a child protection case conference deciding that the child was not at risk of significant harm. However, much of the work is long term and any one 'incident' within it may be distorted by being taken out of context.*

(7) Comparison with Standards

This uses audit and peer review to study adverse events. It is epitomised by the system of confidential inquiries. Typically efforts are made to identify all incidents of interest (usually deaths) using statutory reporting systems, voluntary notification and, in less developed countries, through additional hospital and community based enquiries. The key focus of analysis is on medical records but it can be supplemented by questionnaire enquiries, or interviews with healthcare staff or relatives. The information gathered is then appraised against explicit or implicit standards for care of such patients. A panel of experts typically conducts the appraisal and results are presented as levels of performance against expectation.

Typically, confidential enquiries focus on clinical and patho-physiological factors associated with death, with a variable emphasis on the quality of care as assessed against standards of peers. Some also look at organizational factors.

The strengths of this approach are its confidentiality and voluntary participation are reassuring for clinicians; the close involvement of professional organizations helps to endorse ownership by participants without needing statute; the complete identification of cases improves the generalisability of the findings. Its limitations are that it is only feasible to conduct serial confidential inquiries for a relatively small number of adverse outcomes; its design limits the scope for identifying emergent findings; historically, it has tended to focus more on clinical than contextual issues that might affect patient safety; and findings are remote from individual cases and influence on implementing change is usually through professional organizations and the scientific literature.

9.4 Conclusion

Healthcare is an interesting area to study because of its similarities with child welfare service. However, the resemblance must not be over-stated. It is pertinent that the most developed aspects of a systems approach are in those aspects of healthcare where the technical dimension outweighs the social dimension. The paradigm examples given of latent errors are often in relation to aspects of technical processes such as using the correct drug in the correct dosage. It is also more concentrated so far in services provided within the more controlled environment of a hospital rather than the community where professional involvement is mixed with numerous other factors in the overall experience of patients.

It is also noteworthy how recent most of the developments have been. Much of the work seems to have been done in the last ten years. However, within that time, considerable progress has been made in adapting methodologies to health care settings and to working out how key concepts such 'error' or 'critical incident' can be defined in specific healthcare settings.

A range of investigatory techniques in engineering have been scrutinised for adaptation to healthcare. However, it is important to stress Vincent's 2006, p. 121) that these have rarely been formally evaluated even within their engineering setting:

> A major concern with all the techniques discussed is the lack of formal testing and evaluation. Most of these methods are in what might be termed a 'development' phase, analogous to the initial testing of a drug in clinical populations. In one of the few reviews of these techniques, Jeremy Williams began by saying: "It must seem quite extraordinary to most scientists engaged in research into other areas of the physical and technological world that there has been little attempt by human reliability experts to validate the human reliability assessment techniques which they so freely propagate, modify an disseminate".

One final point is that the spread of a systems approach is greatly facilitated by its acceptance by the key international and national organizations that have overall responsibility for the health services. This ownership by the most senior levels of the system is perhaps not only desirable but essential before lower levels can confidently develop a learning culture instead of a blame one.

Useful websites

National Patient Safety Agency: http://www.npsa.nhs.uk

Patient Safety Research Programme, University of Birmingham: http://www.pcpoh.bham.ac.uk/publichealth/psrp/

SCIE: http://www.scie.org.uk

Chapter Ten

10. Conclusion

In many high-risk industries, including healthcare, there has been a shift from a traditional, person-centred approach to understanding error and improving safety to a systems-centred focus. This fundamental change views the complex interaction of all elements in the system as potentially contributing to both safety and error. Human error is no longer seen as an adequate explanation of a mishap but as requiring closer investigation to find what factors in the wider context contributed to that worker's action or omission. The range of factors that could be considered is immense, including factors relating to human cognition and emotion, factors in the design of technology and how it shapes human performance, and factors in the organizational culture, resources, priorities, and teams.

While all in safety management have aspired to a scientific approach, there have been differences of opinion on what this entails. Early researchers took a relatively narrow, positivist view while later researchers embrace a wider range of methods within an empirical framework. More recent writers have highlighted the distinctive features of studying systems as opposed to sets of individuals. Complexity inasmuch as it implies non linear dynamics places limits on prediction. Systems are seen as having emergent properties, arising from the interaction of lower-level entities, none of which show them. There is a greater concern to understand the 'local rationality' – how the individual workers are making sense of their environment and their tasks. This leads to a greater interest in interpretivist psychologies and the social construction of the workplace unlike their predecessors who limited the study of human actions to observable behaviour or, later, the observable elements of cognition.

Defining error is far from simple. It presupposes some standard of 'correct' performance against which a shortfall can be judged. This makes the concept problematic in a child welfare context where there is no consensus on a detailed prescription of the right way to work with families. Formal procedures offer one readily available standard but defining error relative to compliance with procedures in child welfare is risky since the procedures have not been empirically validated and may themselves make a significant contribution to adverse outcomes.

Developing a theoretical understanding of both good and problematic practice is a necessary step in formulating a taxonomy of errors, needed for both incident investigation and for setting up a reporting system. In the light of the findings by SCIE (Bostock et al., 2005), it would seem premature to consider developing a reporting system in child welfare since, without a shared theoretical understanding, there would be very low user agreement on what to call a near miss, or how to classify it when making the report. Without at least a basic shared taxonomy, the likelihood is that one would have a reporting system containing a large set of individual incidents that would pose an immense challenge to anyone attempting to analyse them and draw general lessons.

In reviewing the potential range of factors contributing to error (and to safe practice), the material was presented in three groups: factors pertaining to human reasoning, the relationship between the worker and the various artefacts that are designed to improve performance, and the overall organizational context, with a culture, sets of priorities, resources etc. For each of the variables in these clusters, it is possible to find technical books offering guidance on how to study them but the pool of such literature is so vast and so detailed that it is not feasible to do more than point to their existence in this review.

The focus on systemic vulnerabilities, not individual lapses, has led not only to changes in how accidents are investigated but to the development of strategies for continual learning in organizations so that weak areas of practice can be identified and improved before a major incident. Recognition of the complex interplay of variables undermines any simplistic, top-down model of policy implementation.

While safety management began within high-risk industries such as aviation and nuclear power, it has recently been adopted with enthusiasm by healthcare agencies. The World Health Organization and many national health agencies have all endorsed the need to move from a focus on individual blame to learning from minor mistakes to improve the functioning of the health service. The work in health has more obvious resonance with the issues facing child welfare services in that it too is concerned with helping people. One of the most impressive aspects of developments in healthcare is the acceptance of the systems approach by the highest levels of the organizations. It seems likely that this kind of high-level endorsement is required in order to start the process of moving from a blame to a learning culture. Major innovations have been the establishment of reporting systems and agencies to deal with the compilation and analysis of error reports (for example, the National Patient Safety Agency in the UK). At the same time, managements are aiming to make the needed cultural shift to encourage people to make reports.

A number of methods of accident analysis have been developed in different industries. Woloshynowych et al (2005) reviewed the available methods with a view to assessing their relevance to health care. All the methods have features in common. All try to answer the three basic questions: What happened? How did it happen? Why did it happen? All rely on two prime sources of evidence; documents and interviews, although some also collect data through observation and simulation techniques. All are offering some means of putting a structure onto the wealth of data and helping to ensure that all possibilities are explored. The model that has been most developed in healthcare is the Organizational Accident Causation Model which Vincent has adapted for investigating health care incidents.

The project of adopting a systems approach to improve services in child welfare carries the potential for significant learning and progress. Learning how to reduce errors and learning how to improve practice are two ways of looking at the same central issue. A system that is designed to maximise safety is also one that has good feedback on what is or is not working well. However, this literature review has shown the scale of the project. Moving from a person-centred to a systems approach is a fundamental change in the conceptual framework within which we frame individual and group behaviour. In this final chapter, I wish to raise some issues that seem to pose particular challenges for child welfare.

10.1 Changing the blame culture

The blame culture is widespread within child welfare services in many countries. This is not only apparent in the tendency of child abuse inquiries to conclude by blaming individuals for the tragic outcome but also in day-to-day practice where many practitioners report feeling vulnerable to blame if anything adverse happens in one of the families with whom they are working. Consequently many adopt a defensive attitude to their work and stick to procedures regardless of their appropriateness in a specific situation.

The systems approach goes beyond human error to seek deeper causes of adverse outcomes and therefore is less inclined to blame the individual who was proximate to the outcome because the investigation draws attention to how many other people or factors were implicated. Moreover, the approach requires practitioners to feel free to report minor problems in order to provide the learning needed by the organization to recognise latent errors in the system.

To what extent is it either possible or desirable to alter this blame culture and move toward a reporting and safety culture? Changing an entrenched culture of blame is a major task and the example set by health care is informative. The drive to creating a safety culture is being led by the most senior international and national organizations. It is difficult to imagine how lower levels could change their culture if still operating within a wider system where they knew that blame would be the first response of seniors to any adverse outcome.

Dekker (2006) suggests that change from a blame culture is, in some respects, in conflict with modern culture where:

> 'individualism is still crucial to self-identity in modernity. The idea that it takes teamwork, or an entire organization, or an entire industry to break a system (as illustrated by cases of drift into failure) is too unconventional relative to our inherited cultural preconceptions' (2006: xiii).

Despite the persuasive arguments of a systems approach for a no-blame approach, it has to be accepted that there is a point at which society will hold individuals to account for their actions. Some-one like the English doctor Harold Shipman who killed over a hundred patients cannot be excused as a system error: he deserved the life sentences for murder that he received. However, it can be questioned whether system factors hindered the detection of his crimes. The difficulty lies in deciding where the boundary lies or what degree of culpability an individual carries within a faulty system.

The UK National Patient Safety Agency has done some work on this problem and, drawing on Reason's work on a more general 'culpability matrix' have created the 'Incident Decision Tree' (downloadable from www.npsa.org.) After the incident has been investigated, a series of questions is asked: Were the actions intentional? If yes, was there an intention to cause harm or not? Is there any evidence of a medical condition? Was there any departure from agreed protocols? There are areas of particular difficulty when the 'correct' action is not clear cut and professional judgment is needed to decide what to do.

There is a need to address the issues of culpability and blame at all levels of responsibility from senior management to front line workers in child welfare services before the systems approach can be fully implemented.

10.2 Where do families fit in the system?

The systems approach as it was developed in engineering conceptualises the organizational system as one system interacting with what is often called 'the managed process' – the plane that is being flown, the nuclear power plant that is generating electricity.   What happens to this division when the concepts are transferred to child welfare?

There seem to be two major options and the choice between them is perhaps essentially a political one. Adopting the same approach as engineering, children and parents can be seen as objects to be managed. Their existence as subjects as well as objects (their ability to think and act independently) would need to be factored in in terms of the complications this creates for the interactions between practitioners in the organizational system and individuals in the managed family system.

Alternatively, children and parents can be seen as active participants in the system, not outside it. This resonates with professional guidance stressing partnership and listening to families in practice and with ethical concerns about self-determination. However, this raises complications for the organizational system in terms of how to involve them. Much of the systems literature stresses the need to understand and value the front line workers' perception of events and processes. In a child welfare system, the same degree of attention would need to be given to the experience of families. Constructing mechanisms for doing this will be a challenging process.

The earlier work done by SCIE (Bostock et al., 2005) included input from discussions with families to identify ‘safeguarding incidents’ – lapses that did not, on this occasion, lead to an adverse outcome. This work is encouraging in that it reports families to be able to understand the rationale of the systems approach and to be willing to participate. This raises the question whether, if a reporting system is developed, it should be possible for service users to make reports, not just practitioners.

10.3 Error in a poorly developed knowledge area

What is error in field where there is a relatively poor knowledge base and practitioners have relatively little scope to control the whole environment where change is sought? Compared with engineering and health, child welfare services have far fewer processes where we can confidently say ‘this is the correct course of action’ and so there will be more occasions where there are difficult decisions to make about what should or should not have been done.

On a more positive note, a systems approach with its emphasis on feedback offers a potentially rich mechanism for learning about what does or does not work in child welfare but this leads on to my next issue of feedback.

10.4 Feedback in a long-term service

What type of feedback is needed when the ultimate aim of the service is to have a long term effect on a child’s welfare? The examples from engineering and healthcare focus on outcomes in the near future – whether the plane landed safely or not, or whether the patient survived the operation.

In the audit and inspection systems developed in child welfare services in many countries in recent years, measurement of the product of services has usually been severely limited in value. The focus is often on service outputs (how quickly the case was dealt with, whether a case conference was held) not on user outcomes (whether the child was safe and well) (Munro, 2004). In adopting a systems approach, there is a danger of overemphasising compliance with procedures because this is easily measurable and overlooking the outcomes for children and families that are much more long-term, contested, and hard to measure.

10.5 Progress so far

Initial work has been done on developing the systems approach in child welfare services. Articles by Rzepnicki and Johnson, 2005; Munro (2005a, Munro, 2005b) and Lachman and Bernard (2006) have set out how the approach might be used. Research by the Social Care Institute for Excellence has developed the concept of ‘near misses’ as conceptualised by both practitioners and families (Bostock et al., 2005). A current project being conducted by the UK Social Care Institute for Excellence is piloting a systems approach to investigating adverse outcomes in specific cases (see web address for details).

10.6 Key lessons from the systems approach

Woods and Cook (2002) offer nine steps to move forward from error. These steps provide a simple summing up of the key points covered in this review and therefore seem an appropriate way to end:

1. Pursue second stories beneath the surface to discover multiple contributors.
2. Escape the hindsight bias.
3. Understand work as performed at the sharp end of the system.
4. Search for systemic vulnerabilities.
5. Study how practice creates safety.
6. Search for underlying patterns.
7. Examine how change will produce new vulnerabilities and paths to failure.
8. Use new technology to support and enhance human expertise.
9. Tame complexity through new forms of feedback.

11. References

BAINBRIDGE, L. (1987) Ironies of Automation. *New Technology and Human Errorq*, 271-283.

BLOM-COOPER, L. (1996) Some Reflections on Public Inquiries. IN PEAY, J. (Ed.) *Inquiries after Homicide.* London, Duckworth.

BOEING PRODUCT SAFETY ORGANIZATION (1993) Statistical summary of commercial jet aircraft accidents: Worldwide operations, 1959-1992. Seattle. WA, Boeing Commercial Airplanes.

BOSTOCK, L., BAIRSTOW, S., FISH, S. & MACLEOD, F. (2005) Managing risk and minimising mistakes in services to children and families. London, SCIE.

BOWKER, G. & STARR, S. (2000) *Sorting Things Out,* Cambridge, MA, MIT.

CAPLAN, R., POSNER, K. & CHENEY, F. (1991) Effect of outcome on physician judgments of appropriateness of care. *Journal of the American Medical Association,* 265**,** 1957-1960.

COLLINS, H. & KUSCH, M. (1998) *The Shape of Actions: What Humans and Machines Can Do,* Cambridge MA, MIT.

COOK, R. & WOODS, D. (1998) A Tale of Two Stories: Contrasting Views of Patient Safety. Washington, DC, National Patient Safety Foundation.

COOPER, J., NEWBOWER, R. & KITZ, R. (1984) An analysis of major errors and equipment failures in anesthesia management: conditions for prevention and detection. *Anasthesiology,* 60**,** 42-43.

DEKKER, S. (2002) *The Field Guide to Human Error Investigations,* Aldershot, Ashgate.

DEKKER, S. (2006) *Ten Questions about Human Error,* London, Lawrence Erlbaum Associates.

DEKKER, S. & HOLLNAGEL, E. (2004) Human factors and folk models. *Cognition, Technology & Work,* 6**,** 79.

DEPARTMENT OF HEALTH (1991) Child abuse: A study of inquiry reports, 1980-89. London, HMSO.

DEPARTMENT OF HEALTH (2000) *An Organization with a Memory,* London, The Stationery Office.

DEPARTMENT OF HEALTH AND SOCIAL SECURITY (1974) Report of the committee of inquiry into the care and supervision provided in relation to Maria Colwell. London, HMSO.

DEPARTMENT OF HEALTH AND SOCIAL SECURITY (1982) *Child abuse: a study of inquiry reports, 1973-81,* London, HMSO.

DORNER, D. (1983) Heuristics and cognition in complex systems. IN GRONER, R., GRONER, M. & BESCHOF, W. (Eds.) *Methods of Heuristics.* Hillsdale, New Jersey, Lawrence Erlbaum.

EASTMAN, N. (1996) Towards an Audit of Inquiries: Enquiry not inquiries. IN PEAY, J. (Ed.) *Inquiries after homicide.* London, Duckworth.

FEYERABEND, P. (1975) *Against Method,* London, New Left Books.

FIRTH-COZENS, J. (2000) Teams, Culture and Managing Risk. IN VINCENT, C. (Ed.) *Clinical Risk Management.* London, BMJ Books.

FISCHHOFF, B. (1975) Hindsight-foresight: The effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance,* 1**,** 288-299.

FISCHHOFF, B. (1977) Perceived informativeness of facts. *Journal of Experimental Psychology: Human Perception and Performance,* 3**,** 349-358.

FISCHHOFF, B. (1982) For those condemned to study the past: Heuristics and biases in hindsight. IN KAHNEMAN, D., SLOVIC, P. & TVERSKY, A. (Eds.) *Judgment under Uncertainty.* Cambridge, Cambridge University Press.

FITTS, P. & JONES, R. (1947) Analysis of factors contributing to 460 'pilot error' experiences in operating aircraft controls. *Memorandum Rep.* Ohio, Aero Medical Laboratory.

FRASER, J., SMITH, P. & SMITH, J. (1992) A catalog of errors. *International Journal of Man-Machine Studies,* 37**,** 265-393.

GANO, D. (2003) *Apollo Root Cause Analysis,* Yakima, Washington, Apollonian Publications.

GARDNER, H. (1985) *The mind's new science: A history of the cognitive revolution,* New York, Basic Books.

HACKING, I. (1987) *The Taming of Chance,* Cambridge, Cambridge University Press.

HAMMOND, K. (1996) *Human Judgement and Social Policy,* Oxford, Oxford University Press.

HEIRICH, W. H. (1931) *Industrial Accident Prevention,* New York, McGraw-Hill.

HELMREICH, R. L. (2000) On Error Management: Lessons from Aviation. *British Medical Journal,* 320**,** 781-5.

HELMREICH, R. L. & MERRITT (1998) *Culture at work in aviation and medicine*.

HOCH, S. & LOWENSTEIN, G. (1989) Outcome feedback: Hindsight and information. *Journal of Experimental Psychology: Learning, Memory and Cognition,* 15**,** 605-619.

HOLLNAGEL, E. (1993) *Human Reliability Analysis: Context and control,* London, Academic Press.

HOLLNAGEL, E. (1998) *Cognitive reliability and error analysis method,* Oxford, Elsevier.

HOLLNAGEL, E. & WOODS, D. (2005) *Joint Cognitive Systems: Foundations of Cognitive Systems Engineering,* Poca Raton, Taylor & Francis.

HOME OFFICE (1945) *Report by Sir William Monckton on the Circumstances Which Led to the Boarding Out of Dennis and Terence O'Neill at Bank Farm, Masterley and the Steps Taken to Supervise Their Welfare,* London, HMSO.

HOOD, C. & JONES, D. (Eds.) (1996) *Accident and Design, contemporary debates in risk management,* London, Routledge.

HUMAN FACTORS STUDY GROUP (1993) Third report: Organising for Safety. London, Health and Safety Executive.

HUTCHINS, E. (1995) Cognition in the Wild. *The MIT Press (Publishers)*.

HUTCHINS, E. (2000) Distributed Cognition. Available at http://eclectic.ss.uci.edu/~drwhite/Anthro179a/DistributedCognition.pdf.

INSTITUTE OF MEDICINE (1999) *To Err is Human: Building a Safer Health System,* Washington, DC, National Academic Press.
KAHNEMAN, D., SLOVIC, P. & A, T. (Eds.) (1982) *Judgement under uncertainty: heuristics and biases,* Cambridge, Cambridge University Press.
KUHN, T. (1970) *The Structure of Scientific Revolutions,* Chicago, University of Chicago Press.
LACHMAN, P. & BERNARD, C. (2006) Moving from blame to quality: How to respond to failure in child protection services. *Child Abuse & Neglect,* 30**,** 963-968.
LAPORTE, T. & CONSOLINI, P. (1991) Working in practice but not in theory: Theoretical challenges of "high reliability organisations". *Journal of Public Administration,* January**,** 19-47.
LAW, J. & CALLON, M. (1995) Engineering and sociology in a military aircraft project: A network analysis of technological change. IN STAR, S. (Ed.) *Ecologies of Knowledge: Work and politics in science and technology.* Albany, State University of New York Press.
MUNRO, E. (1999) Common errors of reasoning in child protection. *Child Abuse & Neglect,* 23**,** 745-758.
MUNRO, E. (2004) The impact of child abuse inquiries since 1990. IN MANTHORPE, J. & STANLEY, N. (Eds.) *The Age of the Inquiry.* London, Routledge.
MUNRO, E. (2005a) Improving practice: child protection as a systems problem. *Children and Youth Service Review,* 27**,** 375-391.
MUNRO, E. (2005b) A system's approach to investigating child abuse deaths. *British Journal of Social Work,* 35**,** 531-546.
NORMAN, D. (1990) The 'problem' with automation: inappropriate feedback and interaction, not 'over-automation'. *Philosophical Transactions of the Royal Society of London,* Series B. Biological 327**,** 585-593.
NORMAN, D. (1993) Toward Human-Centered Design. *Technology Review,* 30**,** 47-53.
PERROW, C. (1999) *Normal Accidents: Living with High-Risk Technologies,* Princeton, New Jersey, Princeton University Press.
POOL, R. (1997) *Beyond Engineering: How society shapes technology,* Oxford, Oxford University Press.
POPPER, K. (1963) *Conjectures and Refutations,* London, Routledge and Kegan Paul.
POWER, M. (1997) *The Audit Society,* Oxford, Oxford University Press.
PUTNAM, H. (1978) *Meaning and the Moral Sciences,* London, Routledge and Kegan Paul.
RASMUSSEN, J. (1985) Trends in human reliability analysis. *Ergonomics,* 28**,** 1185-1196.
RASMUSSEN, J. (1986) *Information processing and human-machine interaction: An approach to cognitive engineering,* New York, North-Holland.
RASMUSSEN, J. (1987) The Definition of Human Error. *New Technology and Human Errorq***,** 23-29.
RASMUSSEN, J. (1988) Interdisciplinary workshops to develop a multi-disciplinary research programme based on a holistic system approach

to safty and management of risk in large-scale techonological operations. Washington, DC, World Bank.
RASMUSSEN, J. (1990) Human error and the problem of causality in analysis of accidents. *Philosophical Transactions of the Royal Society of London,* 327**,** 449-460.
REASON, J. (1990a) The contribution of latent human failures to the breakdown of complex systems. *Philosophical Transactions of the Royal Society of London,* 327**,** 475-484.
REASON, J. (1997) *Managing the Risks of Organizational Accidents,* Aldershot, Hants, Ashgate.
REASON, J. (2001) Understanding adverse events: the human factor. IN VINCENT, C. (Ed.) *Clinical risk management: enhancing patient safety.* 2nd ed. London, BMJ Publications.
REASON, J. & HOBBS, A. (2003) *Managing Maintenance Error,* Aldershot, Aldgate.
REASON, P. (1990b) *Human Error,* Cambridge, Cambridge University Press.
ROBERTS, K. & BEA, R. (2001) Must accidents happen? *Academy of Management Executive,* 15**,** 70-79.
RZEPNICKI, T. & JOHNSON, P. (2005) Examining decision errors in child protection: A new application of root cause analysis. *Children and Youth Service Review,* 27**,** 393-407.
SAGAN, S. (1993) *The Limits of Safety Organisations: Accidents, and Nuclear Weapons,* Princeton, NJ, Princeton University Press.
SCHON, D. (1973) *Beyond the Stable State: Public and private learning in a changing society,* London, Harmondsworth, Penguin.
SCIE (2004) *Improving the use of research in social care practice,* Bristol, Policy Press.
SENGE, P. M. (1990) *The Fifth Discipline: The art and practice of the learning organization,* London, Random House.
SIMON, H. (1955) A behavioural model of rational choice. *The Quarterly Journal of Economics,* LXIX**,** 99-118.
SIMON, H. (1996) *The Sciences of the Artificial,* Cambridge, Massachussets, The MIT Press.
SPENCER, N., WALLACE, A., SUNDRUM, R., BACCHUS, C. & LOGAN, S. (2006) Child abuse registration, fetal growth, and preterm birth: a population based study. *Journal of Epidemiological Community Health,* 60**,** 337-340.
TURNER, B. & PIDGEON, N. (1997) Man-made Disasters. Oxford, Butterworth Heinemann.
VINCENT, C. (2006) *Patient Safety,* Edinburgh, Elsevier.
VON WINTERFELLDT, D. & EDWARDS, E. (1986) *Decision analysis and behavioural research,* Cambridge, Cambridge University Press.
WALLACE, B. & ROSS, A. (2006) *Beyond Human Error: Taxonomies and Safety Science,* Boca Raton, Taylor & Francis.
WHITTINGHAM, R. (2004) *The Blame Machine: Why Human Error Causes Accidents,* Oxford, Elsevier Butterworth-Heinemann.
WHO (2005a) World Alliance for Patient Safety Forward Programme. Geneva, Switzerland.
WHO (2005b) WHO Draft Guidelines for Adverse Event Reporting and Learning Systems. Geneva, Switzerland, World Health Organisation.

WILDAVSKY, A. (1988) *Searching for Safety,* New Brunswick, NJ, Transaction Publishers.

WOLOSHYNOWYCH, M., ROGERS, S., TAYLOR-ADAMS, S. & VINCENT, C. (2005) The investigation and analysis of critical incidents and adverse events in healthcare. London, Health Technology Assessment.

WOOD, G. (1978) The 'knew-it-all-along' effect. *Journal of Experimental Psychology: Human Perception and Performance,* 4**,** 345-353.

WOODCOCK, K. & SMILEY, A. (1998) Organizational pressures and accident investigation. *Proceedings of the Human Factors Association of Canada.* Mississiauga, Ontario, Human Factors Association of Canada.

WOODS, D. (2003) Discovering how distributed cognition systems work. IN HOLLNAGEL, E. (Ed.) *Handbook of cognitive task design.* HIllsdale, NJ, Lawrence Erlbaum.

WOODS, D. & COOK, R. (2001) From counting failure to anticipating risk: possible futures for patient safety. IN ZIPPERER, L. & CRUSHMAN, S. (Eds.) *Lessons in patient safety: a primer.* Chicago, National Patient Safety Foundation.

WOODS, D. & COOK, R. (2002) Nine Steps to Move Forward from Error. *Cognition, Technology & Work,* 4**,** 137-144.

WOODS, D. & HOLLNAGEL, E. (2006) *Joint Cognitive Systems: Patterns in Cognitive Systems Engineering,* Boca Raton, Taylor & Francis.

WOODS, D., JOHANNESEN, L., COOK, L. & SARTER, N. (1994) *Behind Human Error: Cognitive Systems, Computers and Hindsight,* Wright-Patterson Air Force Base, Ohio, CSERIAC.

WRIGHT, D., MACKENZIE, S., BUCHAN, I., CAIRNS, C. & PRICE, L. (1991) Critical incidents in the intensive therapy unit. *Lancet,* 388**,** 676-678.

ZUBOFF, S. (1988) *In the Age of the Smart Machine: The Future of Work and Power,* New York, Basic Books.